ABSTRACT

Title of dissertation: Common Randomness Principles of Secrecy

Himanshu Tyagi, Doctor of Philosophy, 2013

Dissertation supervisor: Professor Prakash Narayan
Department of Electrical and Computer Engineering
Institute for Systems Research

This dissertation concerns the secure processing of distributed data by multiple terminals, using interactive public communication among themselves, in order to accomplish a given computational task. In the setting of a probabilistic multiterminal source model in which several terminals observe correlated random signals, we analyze secure distributed data processing protocols that harness the correlation in the data. The specific tasks considered are: computing functions of the data under secrecy requirements; generating secretly shared bits with minimal rate of public communication; and securely sharing bits in presence of a querying eavesdropper. In studying these various secure distributed processing tasks, we adopt a unified approach that entails examining the form of underlying common randomness (CR) that is generated at the terminals during distributed processing. We make the case that the exact form of established CR is linked inherently to the data processing task at hand, and its characterization can lead to a structural understanding of the associated algorithms. An identification of the underlying CR and its decomposition into independent components, each with a different operational significance, is

a recurring fundamental theme at the heart of all the proofs in this dissertation. In addition to leading to new theoretical insights, it brings out equivalences between seemingly unrelated problems. Another distinguishing feature of this work is that it considers interactive communication protocols. In fact, understanding the structure of such interactive communication is a key step in proving our results.

We make the following contributions. First, we propose a new information theoretic formulation to study secure distributed computing using public communication. The parties observing distributed data are trusted but an eavesdropper has access to the public communication network. We examine distributed communication protocols that allow the trusted parties to accomplish their required computation tasks while giving away negligible information about a specified portion of the data to an eavesdropper with access to the communication. Our theoretical results provide necessary and sufficient conditions that characterize the feasibility of various secure computing tasks; in many cases of practical importance, these conditions take a simple form and can be verified easily. When secure computing is feasible, we propose new algorithms in special cases.

Next, we revisit the problem of generating shared secret keys (SKs). We investigate minimum communication requirements for generating information theoretically secure SKs of maximum rates from correlated observations using interactive public communication. In particular, our approach allows us to examine the role of interaction in such communication. On the one hand, we find that interaction is not needed when the observed correlated bits are symmetrically correlated and therefore, in this case, simple noninteractive protocols are the most efficient means of

generating optimum rate SKs. On the other hand, we illustrate that interactive protocols can require a strictly lower rate of overall communication than noninteractive protocols.

Finally, we consider the task of ensuring security against an eavesdropper who makes queries about a portion of the distributed data that the terminals share by communicating over a public network. We introduce an alternative notion of secrecy which requires rendering the task of a querying eavesdropper as onerous as possible. Our main contribution in this part is the development of a new technique for proving converse results for secrecy problems involving CR with interactive communication, which is employed then to obtain an upper bound for the maximum number of queries that can be inflicted on the eavesdropper for any CR and corresponding communication. Surprisingly, there is an equivalence between this notion of secrecy and that of information theoretic security, which leads to new theoretical results for SK generation; for instance, we prove a strong converse for the SK capacity.

We conclude by hypothesizing the basic principles of secrecy generation that emerge from the results developed in this dissertation.

Common Randomness Principles of Secrecy

by

Himanshu Tyagi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Dissertation Committee:
Professor Prakash Narayan, Advisor
Professor Alexander Barg
Professor P. S. Krishnaprasad
Professor Sennur Ulukus
Professor Armand Makowski
Professor Jonathan Katz

PREFACE

It is different, they say, from knowledge;
It is different, they say, from ignorance.
- Isha Upanishad

This dissertation started over a coffee conversation with my advisor Professor Prakash Narayan about his results with Professor Imré Csiszár on secret key generation. He verbally described the main result: Multiple terminals can generate an optimum rate secret key by sharing the combined observations of all the terminals. The prior art, for a two terminal setup, involved generating optimum rate secret keys by sharing the observations of one of the terminals. However, an extension of this asymmetric scheme to the multiterminal case was not available. The alternative scheme of Csiszár and Narayan first recovered the entire data at all the terminals, using the least rate of public communication, and then extracted an optimum rate secret key from the recovered data. This scheme has an intriguing interpretation. In coffee shop parlance, if the observation of each terminal is represented by a slice of the "randomness cake," the overall cake can be cut into two (almost) independent parts: one corresponding to the secret key and the other to the interactive public communication used to share it. It is clear that such decompositions of *common randomness* must always underlie secure data processing protocols requiring (almost) independence of the observations of the eavesdropper and the secure portion of the data. The question is what constitutes the hypothetical randomness cake – what is the total common randomness available for decomposition?

This question is the starting point of our research. For various secure data processing tasks, we characterize the form of common randomness allowed and make the case that it is connected inherently to the structure of the underlying protocols. Our presentation is reverse chronological – starting with the technical results, we conclude by hypothesizing our basic common randomness principles of secrecy that led to them.

This dissertation could not have been completed without the help of so many of my friends and well-wishers, and the support of U. S. National Science Foundation under grants CCF0635271, CCF0830697, and CCF1117546. I express my deepest gratitude to my guru and advisor Professor Prakash Narayan for teaching me how to read, write, critique, discover, and invent. He patiently heard my hyperbolic assertions and pointed out flaws in my ambitious proofs. Also, my graduate learning experience was enriched by long discussions with the faculty at UMD. I am especially grateful to Professor Alexander Barg, Professor P. S. Krishnaprasad, Professor Steve Marcus, and Professor Patrick Fitzpatrick for their mentorship. It is my honor to thank Professor Imré Csiszár for hosting me at the Rényi Institute of Mathematics, Budapest for a semester and sharing with me his wealth of knowledge and experience. I am indebted to Dr. Piyush Gupta for serving as my mentor during my summer internship at Bell Labs, Murray Hill, where a part this dissertation was completed. My intellectual experience at UMD would have been incomplete without my fellow graduate students. In particular, I would like to thank Ersen Ekrem for exploring with me the nuances of the physical and the metaphysical. Thanks are due also to Shalabh Jain, Osman Yağan, and all my other friends at UMD who made Maryland my home away from home.

I am grateful to my partners in crime Arvind, Mukesh, Achint, Shalabh, Sameen, and my sister Jeevika for paying no heed to my incessant blathering about my research and always amusing me with their antics. I thank my parents for their unconditional love and unreasonable support. Their reassuring presence facilitated every difficult decision of my graduate life. Last and the most, I thank my loving wife Pooja for sharing this journey with me. Her love for the absurd and clarity of thought have inspired every aspect of my research.

Himanshu Tyagi
College Park, Maryland.

To maaji, the woman who defied time

Table of Contents

List of Abbreviations and Notations

ASK	aided secret key
BSS	binary symmetric sources
CI	common information
CR	common randomness
i.i.d.	independent and identically distributed
mcf	maximum common function
pmf	probability mass function
rv	random variable
SK	secret key
WSK	wiretap secret key

Notation. Let $X_1, \ldots, X_m$, $m \geq 2$, be rvs with finite alphabets $\mathcal{X}_1, \ldots, \mathcal{X}_m$, respectively, and a known joint pmf. For any nonempty set $A \subseteq \mathcal{M} = \{1, \ldots, m\}$, we denote $X_A = (X_i,\ i \in \mathcal{A})$. Similarly, for real numbers $R_1, \ldots, R_m$ and $A \subseteq \mathcal{M}$, we denote $R_A = (R_i,\ i \in A)$. Let $\mathcal{A}^c$ be the set $\mathcal{M} \backslash A$. We denote n i.i.d. repetitions of $X_{\mathcal{M}} = (X_1, \ldots, X_m)$ with values in $\mathcal{X}_{\mathcal{M}} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_m$ by $X_{\mathcal{M}}^n = (X_1^n, \ldots, X_m^n)$ with values in $\mathcal{X}_{\mathcal{M}}^n = \mathcal{X}_1^n \times \ldots \times \mathcal{X}_m^n$. Given $\epsilon > 0$, for rvs U, V, we say that U is *ϵ-recoverable* from V if $\mathrm{P}\left(U \neq f(V)\right) \leq \epsilon$ for some function $f(V)$ of V. Denote by $\|U\|$ the size of the range space of the rv U. All logarithms and exponentials are with respect to the base 2.

CHAPTER 1

Introduction

This dissertation concerns the secure processing of distributed data by multiple terminals, using interactive public communication among themselves, in order to accomplish a given task. In many applications, data is collected and stored at physically or temporally separated locations. Examples of the former include data grids and sensor networks. The distributed data is assumed to be correlated, where the correlation can arise from shared copies of the same data, as in data grids [72]; or from the nature of the data itself, as in distributed video coding [30, 55] and in sensor networks [14, 25]. Instances of temporally separated locations with correlated data arise in biometric authentication [56] and hardware authentication [28], where the distributed data consists of the original and the noisy versions of signatures recorded at the registration and the authentication stages, respectively. In all such settings, the entities at these locations are provided with a communication infrastructure to exchange information in order to facilitate various tasks such as sharing distributed data or computing functions of their collective data. For instance, sensor nodes can communicate over a wireless network to compute average or extreme values of their measurements, or "helper" data in a biometric security system can be stored publicly to correct any errors in subsequent recordings of a biometric signature. We study the following general question:

If the shared communication is public, how can we guarantee security?

In the setting of a probabilistic multiterminal source model in which different terminals observe correlated random signals [58, 20], we analyze secure distributed data processing protocols that harness the correlation in the data. The specific tasks considered are: computing functions of the data under secrecy requirements (Chapters 3, 4); generation of secretly shared bits with minimal rate of public communication (Chapter 5); and securely sharing bits in presence of a querying eavesdropper (Chapter 6).

For the first task, we propose a new information theoretic formulation to study secure distributed computing using public communication. The parties observing distributed data are trusted but an eavesdropper can access the public communication network. We examine distributed communication algorithms that allow the trusted parties to accomplish their required computation tasks while giving away negligible information about the computed value to an eavesdropper with access to the communication. This proposed setup is general and provides a unified framework for studying problems of secure computing over sensor networks, secure distributed storage and secure computing over a data grid. Our theoretical results provide necessary and sufficient conditions characterizing the feasibility of various secure computing tasks; in many cases of practical importance, these conditions take a simple form and can be verified easily. This characterization of secure computing provides a basic test that must be undertaken before attempting to construct algorithms. Furthermore, in special cases, we propose new algorithms for secure computing when it is feasible.

The second task concerns the problem of generating shared *secret keys* (SKs), a common objective in many security applications. Most of the current cryptosystems rely on the availability of such shared SKs. In the context of biometric and hardware security, biometric signatures [37] and physically unclonable functions[54], respectively, constitute such shared SKs that have been generated from noisy recordings. We investigate minimum communication requirements for generating information theoretically secure SKs of maximum rates from correlated observations using interactive public communication. In particular, our approach allows us to examine the role of interaction in such communication. For instance, we find that interaction is not needed when the observed correlated bits are symmetrically correlated and therefore, in this case, simple noninteractive protocols are the most efficient for generating optimum rate SKs.

The last task considered entails ensuring security against an eavesdropper who makes queries about a portion of the distributed data that the terminals share by communicating over a public network. We introduce a new notion of secrecy which requires rendering the task of a querying eavesdropper as onerous as possible. Applications include systems secured using biometric authentication where a portion of the recorded biometric information is stored as (public) helper data to correct any errors that occur in subsequent recordings of a biometric signature. How many attempts must a malicious party, with access to this helper data, make in order to enter the system? We present a general formulation to answer such questions. Furthermore, as a surprise, we find that this new notion of secrecy is in essence equivalent to the notion of information theoretic security. We leverage this connection to obtain new theoretical results for SK generation; for instance, we prove a new strong converse for the SK capacity.

In studying these various secure distributed processing tasks, we follow a unified approach that entails examining the form of underlying *common randomness* (CR) (i.e., shared bits [2]) that is generated at the terminals during distributed processing. In this dissertation, we make the case that the exact form of established CR is linked inherently to the data processing task at hand, and its characterization can lead to a structural understanding of the associated algorithms. An identification of the underlying CR and its decomposition into independent components, each with a different operational significance, is a recurring fundamental theme underlying all the proofs in this dissertation. In addition to leading to new theoretical results, it brings out equivalences between seemingly unrelated problems with a common feature that the same CR is established at the terminals. Previously, Csiszár and Narayan had observed such an equivalence between multiterminal SK generation and multiterminal data compression [20]. Specifically, a duality was shown between the generation of an SK of maximum rate and the problem of attaining "omniscience", i.e., recovering the entire data at the SK-seeking terminals. This duality, led to a characterization of multiterminal SK capacity, which is the largest rate of nearly uniformly distributed CR that meets the security requirement of being nearly independent of the communication used to generate it. Furthermore, it enabled new algorithms for SK generation [81, 83]. In the same spirit, here, too, equivalences of tasks are uncovered, leading to new structural results and new algorithms for secure function computation and efficient SK generation.

Another distinguishing feature of this work is that it considers interactive communication protocols. The communication from a terminal can be any (randomized) function of its own observed signal and of all previous communication. Throughout this dissertation we assume that the communication is authenticated, i.e., the "honest but curious" eavesdropper can only observe passively the communication but cannot tamper with it. Furthermore, it is assumed that the communication is in a broadcast mode – each terminal observes the communication from every other terminal. Understanding the structure of such interactive communication is a key step in proving our results. For instance, the results of Chapter 6 rely on showing that if the observations of two terminals are independent to begin with, then they remain independent when conditioned upon the value of an interactive communication. Also, we show that for many tasks complex interactive protocols are not needed and simple noninteractive communication protocols suffice (see the next section for specific examples).

In the concluding chapter of this dissertation, we hypothesize basic principles of secrecy generation that have emerged from the precise results as well as heuristics developed in our work. These principles have important engineering implications and can serve as guidelines for the design of secure protocols. The theoretical results presented in this dissertation support these principles, and we conjecture that even in much broader settings, the proposed principles must hold for appropriately chosen notions of security. For instance, in recent work we examine the validity of these principles in nonasymptotic regime by considering general descriptions of the observed data with fixed length [70].

1.1 Main contributions

This dissertation makes three main technical contributions, which are summarized below.

1.1.1 Secure computation

Suppose that the terminals in $\mathcal{M} = \{1, \ldots, m\}$ observe correlated signals, and that a subset $\mathcal{A}$ of them are required to compute "securely" a (single-letter) function g of all the signals. To this end, the terminals in $\mathcal{M}$ are allowed to communicate interactively over a public noiseless channel of unlimited capacity, with the communication being observed by all the terminals. The terminals in $\mathcal{A}$ seek to compute g in such a manner as to keep its value information theoretically secret from an eavesdropper that observes the public interterminal communication. A typical application arises in a wireless network of colocated sensors which seek to compute a given function of their correlated measurements using public communication that does not give away the value of the function. In contrast to the classic notion of secure computing in cryptography [80], we assume that the terminals are trustworthy but their public communication network can be accessed by an eavesdropper.

We formulate a new Shannon theoretic multiterminal source model that addresses the elemental question: *When can a function g be computed so that its value is independent of the public communication used in its computation*?

The study of problems of function computation, with and without secrecy requirements, has a long and varied history in multiple disciplines, to which we can make only a skimpy allusion here. Examples include: algorithms for exact function computation by multiple parties (cf. e.g., [40, 79, 27, 29]); algorithms for asymptotically accurate (in observation length) function computation (cf. e.g., [53, 43]); and problems of oblivious transfer [50, 3]. In contrast, our requirement of secure computation[1] is to protect the value of a given function; an instance is [52] where exact function computation with secrecy was sought.

We establish that the answer to the question posed above is connected innately to a problem of SK generation for terminals in $\mathcal{M}$, when, in addition to the public communication, side information is provided to the decoders at the terminals in $\mathcal{A}^c$ in the form of the value of g, and can be used only for recovering the key. Such a key, termed an aided secret key (ASK), constitutes a modification of the original notion of an SK in [48, 1, 20, 21]. The largest rate of such an ASK for $\mathcal{M}$ is the ASK capacity C.

Clearly, a function g that is securely computable for $\mathcal{A}$ can be recovered securely by all the terminals in $\mathcal{M}$ when its value is provided as side information for decoding to the terminals outside[2] $\mathcal{A}$, and so, it will yield an ASK for $\mathcal{M}$ of rate

[1] Unlike in [79] and allied literature, no key is available *apriori* for secure computation but can be devised as a part of the computation procedure.

[2] We do not assume that this value is provided to the terminals in $\mathcal{A}^c$ in the actual secure computing protocol. This is an artifice that is used to derive a necessary condition for secure

equal to the entropy H of g. Therefore, g necessarily must satisfy $H \leq C$. In Theorem 3.4, we show that surprisingly, $H < C$ is a sufficient condition for the existence of a protocol for the secure computation of g by $\mathcal{A}$. When all the terminals in $\mathcal{M}$ seek to compute g securely, the corresponding ASK capacity reduces to the standard SK capacity for $\mathcal{M}$ [20, 21]. Furthermore, under this sufficient condition, our proof exhibits a secure computing protocol that uses noninteractive communication. Therefore, *although interaction was allowed, it is not needed to accomplish secure computing.* As a side result of independent interest, we show that a function that is securely computed by $\mathcal{A}$ can be augmented by residual secret CR to yield an SK for $\mathcal{A}$ of optimum rate.

In proving the sufficient condition above, our main technical tool is a new version of the "balanced coloring lemma" [2, 20]. The latter is an important basic result that is used to show the existence of (nontrivial) mappings h of a given rv U such that $h(U)$ is (almost) independent of another rv[3] V, where U and V are correlated. In Section 2.7.1, we present a new balanced coloring lemma, which builds on and extends the version given in [20].

We also present the capacity for a general ASK model involving *arbitrary* side information at the secrecy-seeking set of terminals; such side information is not available for communication and can be used for key recovery alone. Its capac-

computability of g by $\mathcal{A}$.

[3] In spirit, the same purpose is served by the "generalized privacy amplification" result of Bennett, Brassard, Crépeau and Maurer [7]. Indeed, an alternative proof of some of our results based on generalized privacy amplification was presented in [11].

ity is characterized in terms of the classic concept of *maximum common function* (mcf) [26]. Although this result is not needed in full dose for characterizing secure computability, it remains of independent interest.

Next, we consider a generalization where different terminals seek to compute different functions, without giving away the value of a private function of the data[4]. Specifically, the terminals $\{1, ..., m\}$ wish to compute functions $g_1, ..., g_m$, respectively, of their collective data using communication that must not reveal the value of a specified private function g_0 of the data. If such a communication protocol exists, the functions $g_0, g_1, ..., g_m$ are said to be securely computable.

A characterization of securely computable functions for this general setup remains open. The simplest case of interest when the terminals in a subset $\mathcal{A}$ of $\mathcal{M}$ compute only the private function g_0 and those not in $\mathcal{A}$ perform no computation is settled in Chapter 3 and was discussed above. For this simple case, our results can be reinterpreted as follows: If g_0 is securely computable (by the terminals in $\mathcal{A}$), then

$$H\left(X_{\mathcal{M}}|G_0\right) = H\left(X_{\mathcal{M}}\right) - H\left(G_0\right) \geq R^*, \tag{1.1}$$

and g_0 is securely computable if

$$H\left(X_{\mathcal{M}}|G_0\right) > R^*, \tag{1.2}$$

where R^* has the operational significance of being the minimum overall rate of

[4] For instance, in a variant of Yao's millionaire problem [80], two millionaires communicate to determine the richer between them and they want an eavesdropper not to learn their combined wealth.

communication needed for a specific multiterminal source-coding task that involves the recovery of entire data at all the terminals in $\mathcal{M}$ when the terminals in $\mathcal{A}^c$ are provided the value of g_0 as side information; *this task does not involve any security constraint.* Loosely speaking, denoting the collective data of the terminals by the random variable (rv) $X_{\mathcal{M}}$ and the random value of the function g_0 by the rv G_0, the maximum rate of randomness (in the data) that is independent of G_0 is $H\left(X_{\mathcal{M}}|G_0\right)$. The conditions above imply, in effect, that g_0 is securely computable if and only if this residual randomness of rate $H\left(X_{\mathcal{M}}|G_0\right)$ contains an interactive communication, of rate R^*, for the mentioned source-coding task.

In Theorem 4.1, for a broad class of settings involving the secure computation of multiple functions, we establish necessary and sufficient conditions for secure computation of the same form as (1.1) and (1.2), respectively. The rate R^* now corresponds to, roughly, the minimum overall rate of communication that allows each terminal to:

(i) accomplish its required computation task, and,

(ii) recover the entire data (i.e., attain omniscience) when its decoder alone is also given the value of the private function.

Using the sufficient condition (1.2), we present a specific secure computing protocol with communication of rate R^*. For the simple case of a single function $g = g_0$ discussed above, under (1.2), the secure computing scheme recovers the entire data, i.e., the collective observations of all the terminals, at the (function-seeking) terminals in $\mathcal{A}$ using communication that is independent of G_0. In fact, we observe that this is a special case of the following more general principle: A terminal that computes the private function g_0, can recover the entire data without affecting the conditions for secure computability. This exhibits a structural equivalence between securely computing g_0 at a terminal and recovering the entire data at that terminal without giving away the value of g_0 to an eavesdropper observing the public communication used.

In general, a single-letter formula for R^* is not known. Nevertheless, conditions (1.1) and (1.2) provide a structural characterization of securely computable functions in a broader setting. Also, a general recipe for single-letter characterization is presented, and for the cases in which a single-letter characterization is available, the aforementioned heuristic interpretation of R^* is precise.

1.1.2 Communication for optimum rate secret keys

Consider SK generation by a pair of terminals that observe i.i.d. repetitions of two finite-valued rvs of known joint pmf. The terminals communicate over a noiseless public channel of unlimited capacity, interactively in multiple rounds, in order to agree upon the value of an SK which is required to be (almost) independent of the public communication. The maximum rate of such an SK, termed the SK capacity, was characterized in [48, 1].

In the works of Maurer and Ahlswede-Csiszár [48, 1], SK generation of maximum rate entailed both the terminals recovering the observations of any one of the terminals, using the least rate of communication required to do so. Later, it was shown by Csiszár-Narayan [20] that a maximum rate SK can be generated also by the terminals recovering the observations of both the terminals. Clearly, the latter scheme requires more communication than the former. We address the following question, which was raised in [20, Section VI]:

What is the minimum overall rate of interactive communication R_{SK} required to establish a maximum rate SK?

Curtailing the rate of communication used in SK generation to a minimum is an important design objective, especially when engineering lightweight cryptography systems such as secure sensor networks with limited transmission power available at sensor nodes [23]. The basic question above is a first step towards understanding the tradeoff between the rate of communication used and the rate of SK generated. We answer this question by characterizing the form of CR that the terminals must establish in order to generate a maximum rate SK; two examples of such CR are the observations of any one terminal [48, 1] and of both terminals [20]. While our main result does not yield a single-letter characterization of the minimum rate of communication above, it nonetheless reveals a central link between secrecy generation and Wyner's notion of *common information* (CI) between two dependent rvs X_1 and X_2 [77]. Wyner defined CI as the minimum rate of a function of i.i.d. repetitions of two correlated rvs X_1 and X_2 that enabled a certain distributed source coding task. Alternatively, it can be defined as the minimum rate of a function of i.i.d. repetitions of X_1 and X_2 such that, conditioned on this function, the i.i.d. sequences are (almost) independent; this definition, though not stated explicitly in [77], follows from the analysis therein. We introduce a variant of this notion of CI called the *interactive CI* where we seek the minimum rate of CR, established using interactive communication, that renders the mentioned sequences conditionally independent. Clearly, interactive CI cannot be smaller than Wyner's CI, and can exceed it. Our main contribution in this section is to show a one-to-one correspondence between the CR corresponding to interactive CI and the CR established for generating an optimum rate SK. This correspondence is used to characterize the minimum rate of communication R_{SK} required for generating a maximum rate SK in Theorem 5.1. It is shown that, in fact, R_{SK} is simply interactive CI minus the SK capacity.

Finding a single-letter expression for interactive CI remains an open problem. However, when the number of rounds of interaction are bounded, we do obtain a single-letter formula for interactive CI, which in turn yields a single-letter expression for R_{SK} in Theorem 5.3. Using this expression for R_{SK}, we show that for generating an SK of maximum rate, an interactive communication scheme can have lower rate than a noninteractive one, in general. However, interaction offers no advantage for binary symmetric sources. The expression for R_{SK} in Theorem 5.3 also illustrates the role of sufficient statistics in SK generation. We further explore this issue and show that many CI quantities of interest remain unchanged if the rvs are replaced by their corresponding sufficient statistics (with respect to each other)[5].

[5] Interestingly, the effect of substitution by sufficient statistics has been studied in the context of a rate-distortion problem for a remote source in [24, Lemma 2], and recently, for the lossy and lossless distributed source coding problems in [78].

1.1.3 Querying common randomness

A set of terminals observing correlated signals agree on CR, by communicating interactively among themselves. What is the maximum number of queries of the form "Is CR $= l$?" with yes-no answers, that an observer of (only) the communication must ask in order to resolve the value of the CR?[6] As an illustration, suppose that two terminals observe, respectively, n i.i.d. repetitions of the finite-valued rvs X_1 and X_2. The terminals agree on CR X_1^n with terminal 1 communicating to terminal 2 a Slepian-Wolf codeword of rate $H\left(X_1 \mid X_2\right)$ obtained by random binning. An observer of the bin index can ascertain the value of CR with large probability in approximately $\exp\left[nI\left(X_1 \wedge X_2\right)\right]$ queries (corresponding to bin size). Our results show that more queries cannot be incurred by any other form of CR and associated interactive communication.

In a general setting, terminals $1, ..., m$ observe, respectively, n i.i.d. repetitions of the rvs $X_1, ..., X_m$, and communicate interactively to create CR, say L, for the terminals in a given subset $\mathcal{A} \subseteq \{1, ..., m\}$. For appropriate CR L and interactive communication, the number of queries of the form "Is $L = l$?" that an observer of the communication must ask to resolve L is exponential in n. In Theorem 6.1, we find a single-letter formula for the largest exponent E^*. Remarkably, this formula coincides

[6] This general setup includes the aforementioned biometric application mentioned earlier. When a user is authenticated, the two versions of the biometric signatures at registration and authentication match, and they constitute a CR. Here the helper data is a proxy for the communication. This view is adapted to construct efficient biometric authentication schemes in, for instance, [22].

with the SK capacity for a multitermial source model with underlying rvs $X_1, ..., X_m$ [20, 21]. While it is to be expected that E^* is no smaller than SK capacity, the less-restricted E^* may seem *a priori* to be larger. But it is not so. The coincidence brings out, in effect, an equivalence between inflicting a maximum number of queries on an observer of communication on the one hand, and imposing the explicit secrecy constraint requiring (almost) independence of the SK and the communication on the other hand. In fact, as in the achievability proof of SK capacity in [20], the exponent E^* is achieved by the terminals in $\mathcal{A}$ attaining omniscience, i.e., by generating CR $L = (X_1^n, ..., X_m^n)$ for $\mathcal{A}$, using a communication of minimum rate.

Alternatively, E^* can be interpreted as the smallest rate of a list of CR values produced by an observer of the communication which contains the CR value with large probability.

Our main contribution in this section is a new technique for proving converse results for security problems involving CR with interactive communication, which is employed here to obtain an upper bound on E^*. It relies on query strategies for the CR given the communication that do not depend explicitly on the form of the CR or the communication, and do not require the rvs $(X_{1t}, ..., X_{mt})_{t=1}^n$ to be finite-valued or i.i.d. In fact, our converse results hold even when the underlying alphabets are arbitrary, but under mild technical assumptions. Jointly Gaussian rvs are treated as a special case. Furthermore, our converses are strong in that the characterization of E^* does not depend on the probability of recovery of the CR. This, in turn, leads to a new strong converse result for the SK capacity of the multiterminal source model [20], [21], showing the maximum rate of SK that can be generated does not depend on the probability of recovery of the SK (at the terminals). A byproduct of our technique is a simple lossless block coding result for general finite sources, in terms of Rényi entropies. A particularization recovers the classic lossless block coding result for i.i.d. sources [58] without recourse to the asymptotic equipartition property. The technique[7] is recorded separately in Section 2.7.2.

The number of queries above can be interpreted as a measure of the correlation among the random signals observed by the terminals: A stronger correlation necessitates more queries for resolving the CR that can be generated by them. Such a measure of correlation is in the spirit of the body of work on "guessing" the value of an rv based on a correlated observation [47, 4, 5, 34].

1.2 Organization of the dissertation

The basic multiterminal source model and the notions of CR and SK, along with pertinent known results are given in Chapter 2. In the same chapter, we include a discussion on various measures of CI and point out an interesting invariance property satisfied by these CI quantities. The last section of Chapter 2 contains two important technical tools that have been introduced in this dissertation, namely a new version of the balanced coloring lemma and an estimate of the size of large probability sets in terms of Rényi entropy. These are of independent interest, too.

The secure computing problem is presented in two parts, with Chapter 3 containing the case of a single computed function and Chapter 4 the general case of

[7] Recently, it was brought to our attention [75] that alternative forms of this result exist in prior literature; for instance [59, 13].

multiple functions. Chapter 5 addresses the problem of minimum communication requirements for generating an optimum rate SK. This is followed in Chapter 6 by the problem of querying the value of CR. We conclude in Chapter 7 by hypothesizing the basic principles of secrecy generation that emerge from the results developed in this dissertation.

CHAPTER 2

Classical Concepts and New Tools

2.1 Synopsis

We formulate basic concepts that will be of relevance throughout this dissertation. For a multiterminal source model, the notions of common randomness, omniscience, and secret key capacity are defined. Also, measures of common information of two rvs due to Gács-Körner and Wyner are described, and a new invariance property is established for these measures. In particular, it is shown that for a two-terminal setup, these common information quantities remain unchanged if the two rvs are replaced by their respective sufficient statistics (with respect to each other). Finally, new technical tools are described, which emerge in this dissertation and underlie our proofs. A key tool used in Chapter 6, estimating the size of a large probability set in terms of Rényi entropy, is interpreted separately, too, as a lossless block coding result for general sources. As a specific instance, it yields the classic result for a discrete memoryless source.

Section 2.2 gives the basic set-up of the multiterminal source model and interactive communication that will be used throughout the dissertation. This is followed by Sections 2.3 and 2.4 on definitions and preliminary results for common randomness and secrecy generation. In Section 2.5, we define various common information quantities and establish a new invariance property for them in Section 2.6. The final Section 2.7 formulates technical tools that will be used in this dissertation. Specifically, a new version of the "balanced coloring lemma" is established, which is an important tool to extract almost independent rvs, and a new connection between Rényi entropy and lossless source coding rate is provided. The results in Sections 2.6, 2.7.1 and 2.7.2 are contained, respectively, in [63], [68] and [65].

2.2 Multiterminal source model and interactive communication

Consider a set of terminals $\mathcal{M} = \{1, ..., m\}$ that observe, respectively, the sequences $X_1^n, ..., X_m^n$ of length n. Unless stated otherwise, we assume that the rvs $(X_{1t}, ..., X_{mt})$, $t = 1, ..., n$, are i.i.d. with known distribution $\mathrm{P}_{X_\mathcal{M}}$. This basic multiterminal source model was introduced in [20] in the context of SK generation with public transaction.

The terminals have access to a noiseless public communication network of unlimited capacity over which they can communicate interactively. The communication is authenticated and it is assumed that each terminal observes the communication from every other terminal. Randomization at the terminals is permitted; we assume that terminal i generates a rv $U_i,\ i \in \mathcal{M}$, such that $U_1, \ldots, U_m$ and $X_\mathcal{M}^n$ are mutually independent. While the cardinalities of range spaces of $U_i, i \in \mathcal{M}$, are unrestricted, we assume that $H\left(U_\mathcal{M}\right) < \infty$.

Definition 2.1. *(Interactice Communication)* Assume without any loss of generality that the communication of the terminals in $\mathcal{M}$ occurs in consecutive time slots in r rounds; such communication is described in terms of the mappings

$$f_{11}, \ldots, f_{1m}, f_{21}, \ldots, f_{2m}, \ldots, f_{r1}, \ldots, f_{rm},$$

with f_{ji} corresponding to a message in time slot j from terminal i, $1 \leq j \leq r$, $1 \leq i \leq m$; in general, f_{ji} is allowed to yield any function of (U_i, X_i^n) and of previous communication

$$\phi_{ji} = \{f_{kl} : k < j,\ l \in \mathcal{M} \text{ or } k = j,\ l < i\}.$$

The corresponding rvs representing the communication will be depicted collectively as

$$\mathbf{F} = \{F_{11}, \ldots, F_{1m}, F_{21}, \ldots, F_{2m}, \ldots, F_{r1}, \ldots, F_{rm}\},$$

where $\mathbf{F} = \mathbf{F}^{(n)}(U_{\mathcal{M}}, X_{\mathcal{M}}^n)$; the rv corresponding to ϕ_{ji} is denoted by Φ_{ji}. A special form of such communication will be termed *noninteractive communication* if $\mathbf{F} = (F_1, ..., F_m)$, where $F_i = f_i\,(U_i, X_i^n)$, $i \in \mathcal{M}$. The overall rate of all such communication is given by

$$\frac{1}{n} \log \|\mathbf{F}\|.$$

2.3 Common randomness

It is known from the pioneering work of Gács-Körner [26] (also, see [74]) that correlation does not result in shared bits, in general. Nevertheless, as the terminals communicate with each other they are able to share bits. In fact, if the observations of the terminals are correlated, the rate of the shared bits is greater than the rate of the communication. The concept of CR introduced by Csiszár-Ahlswede [2] formalizes this idea.

Definition 2.2. *(Common Randomness [2])* Given interactive communication $\mathbf{F}$ as in Definition 2.1, an rv $L = L^{(n)}\left(X_{\mathcal{M}}^n\right)$ is ϵ-*common randomness* (ϵ-CR) for $\mathcal{A}$ from[1] $\mathbf{F}$ if there exist rvs $L_i = L_i^{(n)}\left(X_i^n, \mathbf{F}\right)$, $i \in \mathcal{A}$, satisfying

$$\mathrm{P}\left(L_i = L,\ i \in \mathcal{A}\right) \geq 1 - \epsilon. \tag{2.1}$$

The rv L_i will be called an estimate of L at terminal $i \in \mathcal{A}$.

2.4 Secret keys and secret key capacity

Shared SKs lie at the heart of all cryptographic applications. Maurer [48] proposed a framework for studying the generation of (information theoretically secure) SKs as *secret* CR from correlated observations at two terminals. As mentioned above, if the observations of the terminals are correlated, the rate of the *overall* CR generated by the communication is greater than the rate of the communication. Heuristically, this gain in the rate is the root of the generated SK rate. The largest rate of such an SK that can be generated, the SK capacity, was characterized in [48, 1].

[1] The rv L is ϵ-recoverable from $(X_i^n, \mathbf{F})$ for every $i \in \mathcal{A}$ (see the "List of Abbreviations and Notations" before Chapter 1) but not necessarily from $\mathbf{F}$ alone. The deliberate misuse of the terminology "recoverable from $\mathbf{F}$" economizes our presentation.

This standard concepts of SK and SK capacity were extended to multiple terminals in [20, 21]; we will present these general concepts below.

Definition 2.3. *(SK capacity [20, 21])* For $\epsilon_n > 0, n \geq 1$, a function K of $X_{\mathcal{M}}^n$ is an ϵ_n-*secret key* (ϵ_n-SK) for (the terminals in) a given set[2] $\mathcal{A}' \subseteq \mathcal{M}$ with $|\mathcal{A}'| \geq 2$, achievable from observations of length n, randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F} = \mathbf{F}^{(n)}(U_{\mathcal{M}}, X_{\mathcal{M}}^n)$ as above if

(i) K is ϵ_n-recoverable from $(U_i, X_i^n, \mathbf{F})$ for every $i \in \mathcal{A}'$;

(ii) K satisfies the "strong secrecy" condition [20, 21]

$$s_{in}(K, F) \triangleq \log |\mathcal{K}| - H(K \mid \mathbf{F}) = \log |\mathcal{K}| - H(K) + I(K \wedge \mathbf{F}) \leq \epsilon_n, \tag{2.2}$$

where $\mathcal{K} = \mathcal{K}^{(n)}$ denotes the set of possible values of K; The terminology *perfect* SK will be used for a 0-SK.

The SK capacity $C(\mathcal{A}')$ for $\mathcal{A}'$ is the largest rate $\limsup_n (1/n) \log H(K)$ of ϵ_n-SKs for $\mathcal{A}'$ as above,[3] such that $\lim_n \epsilon_n = 0$.

Remark. The secrecy condition (2.2) is tantamount jointly to a nearly uniform distribution for K (i.e., $\log |\mathcal{K}| - H(K)$ is small) and to the near independence of K and $\mathbf{F}$ (i.e., $I(K \wedge \mathbf{F})$ is small).

A single-letter characterization of the SK capacity $C(\mathcal{A}')$ is provided in [20, 21].

[2] For reasons of notation that will be apparent later, we distinguish between the secrecy seeking set $\mathcal{A}' \subseteq \mathcal{M}$ and the set $\mathcal{A} \subseteq \mathcal{M}$ pursuing secure computation.

[3] In [20, 21], a secret key was defined, in general, as $K = K(U_{\mathcal{M}}, X_{\mathcal{M}}^n)$ and SK capacity was shown to be achieved by a function of $X_{\mathcal{M}}^n$. Also, in view of (2.2), SK rate can be defined as $\limsup_n \frac{1}{n} \log |\mathcal{K}^{(n)}|$.

Theorem 2.1. (Characterization of SK Capacity [20, 21]) *The SK capacity* $C(\mathcal{A}')$ *equals*

$$C(\mathcal{A}') = H(X_{\mathcal{M}}) - R_{CO}(\mathcal{A}'), \tag{2.3}$$

where

$$R_{CO}(\mathcal{A}') = \min_{R_{\mathcal{M}} \in \mathcal{R}(\mathcal{A}')} \sum_{i=1}^{m} R_i \tag{2.4}$$

with

$$\mathcal{R}(\mathcal{A}') = \left\{ R_{\mathcal{M}} : \sum_{i \in B} R_i \geq H(X_B \mid X_{B^c}), \quad B \subsetneq \mathcal{M}, \mathcal{A}' \nsubseteq B \right\}. \tag{2.5}$$

Furthermore, the SK capacity can be achieved with noninteractive communication and without recourse to randomization at the terminals in $\mathcal{M}$*.*

Remarks. (i) We recall from [20] that $R_{CO}(\mathcal{A}')$ has the operational significance of being the smallest rate of "communication for omniscience" for $\mathcal{A}'$, namely the smallest rate $\lim_n (1/n) \log \|\mathbf{F}^{(n)}\|$ of suitable communication for the terminals in $\mathcal{M}$ whereby $X_{\mathcal{M}}^n$ is ϵ_n-recoverable from $(U_i, X_i^n, \mathbf{F}^n)$ at each terminal $i \in \mathcal{A}'$, with $\lim_n \epsilon_n = 0$; here $\|\mathbf{F}^{(n)}\|$ denotes the cardinality of the set of values of $\mathbf{F}^{(n)}$. Thus, $R_{CO}(\mathcal{A}')$ is the smallest rate of communication among the terminals in $\mathcal{M}$ that enables every terminal in $\mathcal{A}'$ to reconstruct with high probability all the sequences observed by all the other terminals in $\mathcal{M}$, with the cooperation of the terminals in $\mathcal{M}/\mathcal{A}'$. The resulting omniscience for $\mathcal{A}'$ corresponds to total CR of rate $H(X_{\mathcal{M}})$.

(ii) For the trivial case $|\mathcal{A}'| = 1$, say with $\mathcal{A}' = \{1\}$, we have that $C(\{1\}) = H(X_1)$. Clearly, $K = X_1^n$ attains $C(\{1\})$. On the other hand, if $K = K(X_{\mathcal{M}}^n)$ is an SK for terminal 1, it is also an SK for a relaxed model where terminal 1 remains the same while terminals $2, ..., m$ coalesce and have additional access to X_1^n. The SK capacity for the latter model with two terminals, which is no smaller than $C(\{1\})$, equals $I\left(X_1 \wedge X_{\mathcal{M}}\right) = H\left(X_1\right)$ [48, 1]. Hence, $C\left(\{1\}\right) = H\left(X_1\right)$.

(iii) The SK capacity $C(\mathcal{A}')$ is not increased if the secrecy condition (2.2) is replaced by the following weaker requirement [48, 20]:

$$\frac{1}{n} I(K \wedge \mathbf{F}) \leq \epsilon_n. \tag{2.6}$$

In fact, the "weak secrecy" criterion above was first introduced in [48, 1]. Subsequently, it was noted in [49, 16] that a capacity achieving SK can be generated that satisfies the following stronger secrecy criterion:

$$I(K \wedge \mathbf{F}) \leq \epsilon_n.$$

(iv) An alternative security criterion is based on the variational distance:

$$s_{var}(K, \mathbf{F}) \triangleq \left\| \mathrm{P}_{K,\mathbf{F}} - U_{\mathcal{K}} \times \mathrm{P}_F \right\|_1, \tag{2.7}$$

where $U_{\mathcal{K}}$ denotes a uniform distribution on the set $\mathcal{K}$. Note that the security index s_{in} in (2.2) can be expressed as

$$s_{in}(K, \mathbf{F}) = D\left(\mathrm{P}_{K,\mathbf{F}} \| U_{\mathcal{K}} \times \mathrm{P}_F\right),$$

and so by Pinsker's inequality [18]

$$s_{var}(K, \mathbf{F}) \leq \sqrt{\frac{1}{2} s_{in}(K, \mathbf{F})}.$$

A weaker secrecy criterion than (2.2), which is also widely used in the literature (c.f. [36] and the follow-up work based on "leftover hash lemma"), is the following:

$$s_{var}(K, \mathbf{F}) \leq \epsilon_n. \tag{2.8}$$

Also, it was observed in [20, Lemma 1] that

$$s_{in}(K, \mathbf{F}) \leq s_{var}(K, \mathbf{F}) \log \frac{|\mathcal{K}|}{s_{var}(K, \mathbf{F})}.$$

Therefore, if

$$ns_{var}(K, \mathbf{F}) \to 0 \text{ as } n \to 0, \tag{2.9}$$

then $s_{in}(K, \mathbf{F}) \to 0$. In fact, the achievability scheme in [20] ensures (2.9) by driving $s_{var}(K, \mathbf{F})$ to 0 exponentially rapidly in n. In Chapter 6, we shall establish a new "strong converse" for SK capacity under (2.9)[4].

(v) The weak secrecy criterion in (2.6) does not imply the security criterion in (2.8). Also, the former is not implied by (2.9).

The expression for the SK capacity $C(\mathcal{A}')$ in 2.3 can be expressed alternatively using a (linear programming) dual expression for $R_{CO}(\mathcal{A}')$. Let

$$\mathcal{B} = \{B \subsetneq \mathcal{M} : B \neq \emptyset, \mathcal{A} \nsubseteq B\}. \tag{2.10}$$

Let $\Lambda(\mathcal{A})$ be the set of all collections $\lambda = \{\lambda_B : B \in \mathcal{B}\}$ of weights $0 \leq \lambda_B \leq 1$, satisfying

$$\sum_{B \in \mathcal{B}: B \ni i} \lambda_B = 1, \quad i \in \mathcal{M}. \tag{2.11}$$

[4] The proofs in Chapter 6 can be modified to show a strong converse under (2.8) [73].

Every $\lambda \in \Lambda(\mathcal{A})$ is called a *fractional partition* of $\mathcal{M}$ (see [21, 44, 45, 46]). An equivalent expression for $C(\mathcal{A}')$ is

$$C(\mathcal{A}') = H\left(X_{\mathcal{M}}\right) - \max_{\lambda \in \Lambda(\mathcal{A})} \sum_{B \in \mathcal{B}} \lambda_B H\left(X_B \mid X_{B^c}\right), \qquad 0 < \epsilon < 1. \tag{2.12}$$

Denoting

$$\lambda_{\mathtt{sum}} = \sum_{B \in \mathcal{B}} \lambda_B, \tag{2.13}$$

the expression (2.12) can be written also as

$$C(\mathcal{A}') = \min_{\lambda \in \Lambda(\mathcal{A})} \left[\sum_{B \in \mathcal{B}} \lambda_B H\left(X_{B^c}\right) - \left(\lambda_{\mathtt{sum}} - 1\right) H\left(X_{\mathcal{M}}\right) \right], \tag{2.14}$$

For the case $\mathcal{A} = \mathcal{M}$, the expression above simplifies further to

$$C(\mathcal{M}) = \min_{\pi} \frac{1}{|\pi| - 1} D\left(\mathrm{P}_{X_{\mathcal{M}}} \| \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi_i}} \right), \tag{2.15}$$

where the minimum is over all (nontrivial) partitions $\pi = (\pi_1, ..., \pi_k)$ of $\mathcal{M}$ with $|\pi| = k$ parts, $2 \leq k \leq m$ [12] (see also [20, Example 4]).

Depending on the task at hand, we shall use these expressions for $C(\mathcal{A}')$ interchangeably in this dissertation. Finally, for $m = 2$, the expression for $C(\mathcal{M})$ reduces to the omnipresent mutual information.

Theorem 2.2. [48, 1] *The SK capacity for $\mathcal{A}' = \mathcal{M} = \{0, 1\}$ is*

$$C(\{1, 2\}) = I(X_1 \wedge X_2).$$

For further discussion on SKs, see Section 6.7.

2.5 Common information quantities

The first notion of CI for two rvs was given by Gács and Körner in their seminal work [26]. One interpretation of the Gács-Körner CI is as the maximum rate of a CR that can be established by two terminals observing i.i.d. repetitions of two correlated rvs X_1 and X_2, *without any communication.* Formally,

Definition 2.4. A number $R \geq 0$ is an achievable Gács-Körner CI rate if for every $0 < \epsilon < 1$ there exists an $n \geq 1$ and a (finite-valued) rv $L = L\left(X_1^n, X_2^n\right)$ of rate $(1/n)H(L) \geq R$ such that L is ϵ-recoverable from X_1^n and ϵ-recoverable from X_2^n.

The supremum over all achievable Gács-Körner CI rates is called the Gács-Körner CI of X_1 and X_2, denoted $CI_{GK}(X_1 \wedge X_2)$.

For characterizing their CI, Gács and Körner specified the maximal common function of X_1 and X_2, denoted here as $\text{mcf}(X_1, X_2)$, as separate functions of X_1 and X_2 that agree with probability 1, such that any other common function of X_1 and X_2 is a function of $\text{mcf}(X_1, X_2)$.

Theorem 2.3. [26] *The Gács-Körner CI of the rvs X_1, X_2 is*

$$CI_{GK}(X_1 \wedge X_2) = H(\text{mcf}(X_1, X_2)).$$

In Chapter 3, we introduce a new multiterminal version of mcf in Definition 3.4.

Subsequently, Wyner defined CI as the minimum rate of a function of i.i.d. repetitions of two correlated rvs X_1 and X_2 that facilitated a specific distributed source coding task [77]. Alternatively, it can be defined as the minimum rate of a function of i.i.d. repetitions of X_1 and X_2 such that, conditioned on this function, the i.i.d. sequences are (almost) independent; this definition, though not stated explicitly in [77], follows from the analysis therein. Formally,

Definition 2.5. A number $R \geq 0$ is an achievable Wyner CI rate if for every $0 < \epsilon < 1$ there exists an $n \geq 1$ and a (finite-valued) rv $L = L\left(X_1^n, X_2^n\right)$ of rate $(1/n)H(L) \leq R$ that satisfies the property:

$$\frac{1}{n} I\left(X_1^n \wedge X_2^n \mid L\right) \leq \epsilon. \tag{2.16}$$

Obvious examples of such an rv L are $L = \left(X_1^n, X_2^n\right)$ or X_1^n or X_2^n. The infimum of all achievable CI rates, denoted $CI_W(X_1 \wedge X_2)$, is called the Wyner CI of X_1 and X_2. The following theorem characterizes $CI_W(X_1 \wedge X_2)$.

Theorem 2.4. [77] *The Wyner CI of the rvs X_1, X_2 is*

$$CI_W(X_1 \wedge X_2) = \min_U I(X_1, X_2 \wedge U), \tag{2.17}$$

where the rv U takes values in a (finite) set $\mathcal{U}$ with $|\mathcal{U}| \leq |\mathcal{X}_1||\mathcal{X}_2|$ and satisfies the Markov condition $X_1 -\!\!\circ\!\!- U -\!\!\circ\!\!- X_2$.

The direct part follows from [77, equation (5.12)]. The proof of the converse is straightforward. The following inequality ensues [26, 77]:

$$CI_{GK}(X_1 \wedge X_2) \leq I(X_1 \wedge X_2) \leq CI_W(X_1 \wedge X_2).$$

2.6 Invariance of common information

The concepts and the results reviewed above, which are standard in multiterminal information theory, will be used throughout this dissertation. In this section, we present a new invariance property of CI quantities.

Since any good notion of CI between rvs X_1 and X_2 measures the correlation between X_1 and X_2, it is reasonable to expect the CI to remain unchanged if X_1 and X_2 are replaced by their respective sufficient statistics. The following theorem establishes this for the quantities $H(\text{mcf}(X_1, X_2)), I(X_1 \wedge X_2)$ and $CI_W(X_1 \wedge X_2)$.

Theorem 2.5. *For rvs X_1 and X_2, let functions g_1 of X_1 and g_2 of X_2 be such that $X_1 -\!\!\circ\!\!- g_1(X_1) -\!\!\circ\!\!- X_2$ and $X_1 -\!\!\circ\!\!- g_2(X_2) -\!\!\circ\!\!- X_2$. Then the following relations hold:*

$$H(\text{mcf}(X_1, X_2)) = H\left(\text{mcf}\left(g_1(X_1), g_2(X_2)\right)\right),$$

$$I(X_1 \wedge X_2) = I\left(g_1(X_1) \wedge g_2(X_2)\right),$$

$$CI_W(X_1 \wedge X_2) = CI\left(g_1(X_1) \wedge g_2(X_2)\right),$$

Remark. A new notion of CI, termed interactive CI, is introduced in Chapter 5 and a similar invariance property is established for it in Theorem 5.8

Proof. First note that

$$I(X_1 \wedge X_2) = I\left(g_1(X_1) \wedge X_2\right) = I\left(g_1(X_1) \wedge g_2(X_2)\right).$$

Next, we consider the Gács-Körner CI. Note that any common function of $g_1(X_1)$ and $g_2(X_2)$ is also a common function of X_1 and X_2. Consequently,

$$H(\text{mcf}(X_1, X_2)) \geq H(\text{mcf}(g_1(X_1), g_2(X_2))). \tag{2.18}$$

For the reverse inequality, observe that for an rv U such that $H(U|X_2) = H(U|X_1) = 0$ we have

$$U \mathrel{-\!\!\circ\!\!-} X_1 \mathrel{-\!\!\circ\!\!-} g_1(X_1) \mathrel{-\!\!\circ\!\!-} X_2.$$

Thus, $H\left(U|g_1(X_1)\right) \leq H(U|X_2) = 0$, and similarly, $H\left(U|g_2(X_2)\right) = 0$. In particular, it holds that

$$H\left(\mathrm{mcf}(X_1, X_2)|g_1(X_1)\right) = H\left(\mathrm{mcf}(X_1, X_2)|g_2(X_2)\right) = 0,$$

and so,

$$H(\mathrm{mcf}(X_1, X_2)) \leq H(\mathrm{mcf}(g_1(X_1), g_2(X_2))),$$

which along with (2.18) yields

$$H(\mathrm{mcf}(X_1, X_2)) = H(\mathrm{mcf}(g_1(X_1), g_2(X_2))).$$

Finally, we consider Wyner's CI and claim that this, too, remains unchanged upon replacing the rvs with their respective sufficient statistics (for the other rv). It suffices to show that

$$CI_W(X_1 \wedge X_2) = CI_W(g(X_1) \wedge X_2),$$

for a function g such that $X_1 \mathrel{-\!\!\circ\!\!-} g(X_1) \mathrel{-\!\!\circ\!\!-} X_2$. Consider an rv U for which $X_1 \mathrel{-\!\!\circ\!\!-} U \mathrel{-\!\!\circ\!\!-} X_2$ is satisfied. We have

$$0 = I(X_1 \wedge X_2 \mid U) \geq I\left(g(X_1) \wedge X_2 \mid U\right).$$

It follows from (2.17) that

$$CI_W(X_1 \wedge X_2) \geq CI_W\left(g(X_1) \wedge X_2\right). \tag{2.19}$$

On the other hand, for an rv $L = L\left(g^n\left(X_1^n\right), X_2^n\right)$ we have

$$\frac{1}{n} I\left(X_1^n \wedge X_2^n \mid L\right) = \frac{1}{n} I\left(g^n\left(X_1^n\right) \wedge X_2^n \mid L\right),$$

since

$$\begin{aligned} I\left(X_1^n \wedge X_2^n \mid L, g^n\left(X_1^n\right)\right) &\leq I\left(X_1^n \wedge X_2^n, L \mid g^n\left(X_1^n\right)\right) \\ &= I\left(X_1^n \wedge X_2^n \mid g^n\left(X_1^n\right)\right) = 0. \end{aligned}$$

Thus, from the definition of $CI_W(g(X_1) \wedge X_2)$ we get

$$CI_W(X_1 \wedge X_2) \leq CI_W(g(X_1) \wedge X_2),$$

so that, by (2.19),

$$CI_W(X_1 \wedge X_2) = CI_W(g(X_1) \wedge X_2).$$

□

2.7 Two basic tools

In this section, we present two technical tools that have been developed in this dissertation and may be of independent interest.

2.7.1 Balanced coloring lemma

Since our security criterion involves almost independence, all our achievability schemes rely on the existence of a mapping ϕ of an rv U that is almost independent of another rv V correlated with U. For instance, in the SK generation problem, with the established CR in the role of U and the eavesdropper's observations (including the public communication) in the role of V, ϕ is used to extract an SK. In the secure computing problem, with the local observations at a terminal in the role of U and the private function value in the role of V, ϕ constitutes a communication from the terminal which is almost independent of the private function value. One basic tool for showing the existence of such a mapping ϕ is the "balanced coloring lemma" of Ahlswede and Csiszár [2] stated below[5].

Lemma 2.6. [2, Lemma 3.1] *Let $\mathcal{P}$ be any family of N pmfs on a finite set $\mathcal{U}$, and let $d > 0$ be such that $P \in \mathcal{P}$ satisfies*

$$P\left(\left\{u : P(u) > \frac{1}{d}\right\}\right) \leq \epsilon, \tag{2.20}$$

for some $0 < \epsilon < (1/9)$. Then the probability that a randomly selected mapping $\phi : \mathcal{U} \rightarrow \{1, ..., r\}$ fails to satisfy

$$\sum_{i=1}^{r} \left| \sum_{u:\phi(u)=i} P(u) - \frac{1}{r} \right| < 3\epsilon, \tag{2.21}$$

simultaneously for each $P \in \mathcal{P}$, is less than $2Nr \exp\left(-\frac{\epsilon^2 d}{3r}\right)$.

Note that for $\mathcal{P} = \left\{\mathrm{P}_{U|V=v}, v \in \mathcal{V}\right\}$, the left side of (2.21) is $s_{var}(\phi(U), V)$, where s_{var} is defined in (2.7). Therefore, we can find a mapping ϕ such that $s_{var}(\phi(U), V) < 3\epsilon$ if $2Nr \exp\left(-\frac{\epsilon^2 d}{3r}\right) < 1$. This gives an estimate of the size r of the range of ϕ to ensure almost independence of $\phi(U)$ and V. In fact, the bound in (2.20) need not be satisfied for every $\mathrm{P}_{U|V=v}$, and it suffices to require (2.20) for $v \in \mathcal{V}_0$ with $\mathrm{P}_V(\mathcal{V}_0)$ close to 1. As an illustration, consider rvs U^n and V^n, the n

[5] An alternative tool for the same purpose is the "generalized privacy amplification" result of Bennett, Brassard, Crépeau and Maurer [7].

i.i.d. repetitions of U, V. Then, the foregoing approach guarantees the existence of a mapping ϕ of U^n of rate (approximately) $H(U \mid V)$ such that $s_{var}(\phi(U^n), V^n) \to 0$ as n goes to ∞.

A typical application of the lemma above is to the case where V includes the value of a mapping $h : \mathcal{U} \to \{1, \ldots, r'\}$. In [20, Lemma B.2], Csiszár and Narayan proved a version of the "balanced coloring lemma" that was tailored to handle this case. When applied to the i.i.d. illustration above, it implied that for a mapping $h(U^n)$ of rate R, there exists a mapping $\phi(U^n)$ of rate (approximately) $H(U \mid V) - R$ such that for $Z_n = (V^n, h(U^n))$,

$$s_{var}(\phi(U^n), Z_n) \to 0 \quad \text{as} \quad n \to \infty.$$

In other words, there is a loss equal to R in the rate of the constructed mapping ϕ if the eavesdropper additionally knows $h(U^n)$ of rate R. As an application, consider the problem of generating an SK for $\mathcal{M}$. If a CR $X_{\mathcal{M}}^n$ is established using a communication of rate R, then an SK of rate $H(X_{\mathcal{M}}) - R$ can be generated. Therefore $H(X_{\mathcal{M}}) - R_{CO}(\mathcal{M})$ is an achievable SK rate for $\mathcal{M}$, which is optimal by Theorem 2.1.

In this dissertation we need a further generalization of [20, Lemma B.2] when the bound in (2.20) holds not for U but for another rv U' that differs from U with probability close to 0, i.e., we call for a balanced coloring of U while we have the bound (2.20) holding for U'.

Specifically, consider rvs U, U', V with values in finite sets $\mathcal{U}, \mathcal{U}', \mathcal{V}$, respectively, where U' is a function of U, and a mapping $h : \mathcal{U} \to \{1, \ldots, r'\}$. For $\lambda > 0$, let $\mathcal{U}_0$

be a subset of $\mathcal{U}$ such that

(i) $\mathrm{P}\left(U \in \mathcal{U}_0\right) > 1 - \lambda^2$;

(ii) given the event $\{U \in \mathcal{U}_0, h(U) = j, U' = u', V = v\}$, there exists $u = u(u') \in \mathcal{U}_0$ satisfying

$$\begin{aligned} &\mathrm{P}\left(U' = u' \mid h(U) = j, V = v, U \in \mathcal{U}_0\right) \\ &= \mathrm{P}\left(U = u \mid h(U) = j, V = v, U \in \mathcal{U}_0\right), \end{aligned} \tag{2.22}$$

for $1 \leq j \leq r'$ and $v \in \mathcal{V}$. Then the following holds.

Lemma 2.7. *Let the rvs U, U', V and the set $\mathcal{U}_0$ be as above. Further, assume that*

$$\mathrm{P}_{UV}\left(\left\{(u, v) : \mathrm{P}\left(U = u \mid V = v\right) > \frac{1}{d}\right\}\right) \leq \lambda^2. \tag{2.23}$$

Then, a randomly selected mapping $\phi : \mathcal{U}' \to \{1, \ldots, r\}$ fails to satisfy

$$\begin{aligned} &\sum_{j=1}^{r'} \sum_{v \in \mathcal{V}} \mathrm{P}\left(h(U) = j, V = v\right) \times \\ &\sum_{i=1}^{r} \left| \sum_{\substack{u' \in \mathcal{U}': \\ \phi(u') = i}} \mathrm{P}\left(U' = u' \mid h(U) = j, V = v\right) - \frac{1}{r} \right| < 14\lambda, \end{aligned} \tag{2.24}$$

with probability less than $2rr'|\mathcal{V}| \exp\left(-\frac{c\lambda^3 d}{rr'}\right)$ for a constant $c > 0$.

Remark. Note that the quantity on the left side of (2.24) is $s_{var} = s_{var}(\phi(U), (h(U), V))$. By [20, Lemma 1] it holds that

$$s_{in}(\phi(U), (h(U), V)) \leq s_{var} \log \frac{r}{s_{var}},$$

where s_{in} is as in (2.2). Since the function $f(x) = x \log(r/x)$ is increasing for $0 < x < r/e$, it follows from (2.24) that

$$s_{in}(\phi(U), (h(U), V)) \leq 14\lambda \log \frac{|\mathcal{U}|}{14\lambda}. \tag{2.25}$$

Therefore, the "balanced coloring lemma" suffices to show security in the sense of (2.2), too.

The proof of Lemma 2.7 is a variation of the proof of [20, Lemma B.2] and is given in Appendix C.

2.7.2 Rényi entropy and sets with large probability

The next result relates the cardinalities of large probability sets to Rényi entropy. The first part is used in the converse proofs in Chapter 6. The mentioned result is of independent interest and is shown below to yield an elementary alternative proof of the source coding theorem for an i.i.d. (finite-valued) source.

Definition 2.6. *[57]* Let μ be a nonnegative measure on $\mathcal{U}$. For $0 \leq \alpha \neq 1$, the *Rényi entropy of order* α of μ is defined as

$$H_\alpha(\mu) = \frac{1}{1-\alpha} \log \sum_{u \in \mathcal{U}} \mu(u)^\alpha.$$

Lemma 2.8. *(i) For every* $0 < \delta < \mu(\mathcal{U})$*, there exists a set* $\mathcal{U}_\delta \subseteq \mathcal{U}$ *such that*

$$\mu\left(\mathcal{U}_\delta\right) \geq \mu(\mathcal{U}) - \delta, \tag{2.26}$$

and

$$|\mathcal{U}_\delta| \leq \delta^{-\alpha/(1-\alpha)} \exp\left(H_\alpha(\mu)\right), \qquad 0 \leq \alpha < 1. \tag{2.27}$$

(ii) Conversely, for $\delta, \delta' > 0$*,* $\delta + \delta' < \mu(\mathcal{U})$*, any set* $\mathcal{U}_\delta \subseteq \mathcal{U}$ *with* $\mu\left(\mathcal{U}_\delta\right)$ *as in (2.26) must satisfy*

$$|\mathcal{U}_\delta| \geq \left(\delta'\right)^{1/(\alpha-1)} \left(\mu(\mathcal{U}) - \delta - \delta'\right) \exp\left(H_\alpha(\mu)\right), \qquad \alpha > 1. \tag{2.28}$$

Proof. (i) For $0 \leq \alpha < 1$, defining $\mathcal{U}_\delta = \left\{u \in \mathcal{U} : \mu(u) > \delta^{\frac{1}{1-\alpha}} \exp\left[-H_\alpha(\mu)\right]\right\}$, we get

$$\mu(\mathcal{U}) = \mu(\mathcal{U}_\delta) + \sum_{u:\, \mu(u) \leq \delta^{\frac{1}{1-\alpha}} \exp[-H_\alpha(\mu)]} \mu(u).$$

Writing the summand in the right-side above as $\mu(u) = \mu(u)^\alpha \mu(u)^{1-\alpha}$, we obtain

$$\begin{aligned}\mu(\mathcal{U}) &\leq \mu(\mathcal{U}_\delta) + \delta \exp\left[-(1-\alpha)H_\alpha(\mu)\right] \sum_{u \in \mathcal{U}} \mu(u)^\alpha \\ &= \mu\left(\mathcal{U}_\delta\right) + \delta, \end{aligned} \tag{2.29}$$

which is (2.26). Furthermore,

$$\begin{aligned}\exp\left[(1-\alpha)H_\alpha(\mu)\right] &= \sum_{u \in \mathcal{U}} \mu(u)^\alpha \\ &\geq \sum_{u \in \mathcal{U}_\delta} \mu(u)^\alpha \\ &\geq |\mathcal{U}_\delta| \delta^{\frac{\alpha}{1-\alpha}} \exp\left[-\alpha H_\alpha(\mu)\right], \end{aligned} \tag{2.30}$$

which gives (2.27).

(ii) By following the steps in the proof of (i), for $\alpha > 1$, it can shown that the set

$$\mathcal{U}_0 = \left\{u \in \mathcal{U} : \mu(u) < (\delta')^{1/(1-\alpha)} \exp[-H_\alpha(\mu)]\right\} \tag{2.31}$$

has

$$\mu(\mathcal{U}_0) > \mu(\mathcal{U}) - \delta',$$

which, with (2.26), gives

$$\mu(\mathcal{U}_0 \cap \mathcal{U}_\delta) > \mu(\mathcal{U}) - \delta - \delta'.$$

Since by (2.31)

$$\mu(\mathcal{U}_0 \cap \mathcal{U}_\delta) < |\mathcal{U}_0 \cap \mathcal{U}_\delta| \, (\delta')^{1/(1-\alpha)} \exp[-H_\alpha(\mu)],$$

(2.28) follows. □

Lemma 2.8 relating the cardinalities of large probability sets to Rényi entropy can be interpreted as a *source coding* result for a general source with finite alphabet $\mathcal{U}$. Furthermore, it leads to the following asymptotic result.

Consider a sequence of probability measures μ_n on finite sets $\mathcal{U}_n$, $n \geq 1$. For $0 < \delta < 1$, R is a δ-achievable (block) source coding rate if there exists sets $\mathcal{V}_n \subseteq \mathcal{U}_n$ satisfying

$$\mu_n(\mathcal{V}_n) \geq 1 - \delta,$$

for all n sufficiently large, and

$$\limsup_n \frac{1}{n} \log |\mathcal{V}_n| \leq R.$$

The optimum source coding rate $R^*(\delta)$ is the infimum of all such δ-achievable rates.

Proposition 2.9. *For each* $0 < \delta < 1$*,*

$$\lim_{\alpha \downarrow 1} \; \limsup_n \frac{1}{n} H_\alpha(\mu_n) \leq R^*(\delta) \leq \lim_{\alpha \uparrow 1} \; \limsup_n \frac{1}{n} H_\alpha(\mu_n). \tag{2.32}$$

Corollary 2.10. *If* μ_n *is an i.i.d. probability measure on* $\mathcal{U}_n = \mathcal{U} \times ... \times \mathcal{U}$*, then*

$$R^*(\delta) = H(\mu_1), \qquad 0 < \delta < 1.$$

Proof. The Proposition is a direct consequence of Lemma 2.8 upon taking appropriate limits in (2.27) and (2.28) with $\mathcal{U}_n$ in the role of $\mathcal{U}$. The Corollary follows since for i.i.d. μ_n,

$$H_\alpha(\mu_n) = nH_\alpha(\mu_1) \text{ and } \lim_{\alpha\to 1} H_\alpha(\mu_1) = H(\mu_1).$$

□

Note that the Corollary above is proved without recourse to the "asymptotic equipartition property". Moreover, it contains a strong converse for the *lossless coding theorem* for an i.i.d. source. In general, Proposition 2.9 implies a strong converse whenever the lower and upper bounds for $R^*(\delta)$ in (2.32) coincide. This implication is a special case of a general source coding result in [32, Theorem 1.5.1], [33], where it was shown that a strong converse holds iff for rvs U_n with pmfs μ_n, the "lim-inf" and "lim-sup" of $Z_n = \frac{1}{n}\log\frac{1}{\mu_n(U_n)}$ in μ_n-probability coincide, i.e.,

$$\sup\left\{\beta : \lim_n \mu_n(Z_n < \beta) = 0\right\} = \inf\left\{\beta : \lim_n \mu_n(Z_n > \beta) = 0\right\}. \tag{2.33}$$

In fact, a straightforward calculation shows that the lower and upper bounds for $R^*(\delta)$ in (2.32) are admissible choices of β on the left- and right-sides of (2.33), respectively.

CHAPTER 3

Secure Computation

3.1 Synopsis

A subset of a set of terminals that observe correlated signals seek to compute a given function of the signals using public communication. It is required that the value of the function be concealed from an eavesdropper with access to the communication. We show that the function is securely computable if and only if its entropy is less than the capacity of a new secrecy generation model, for which a single-letter characterization is provided.

The main results in Section 3.3 are organized in three parts: capacity of a (new) aided secret key model; characterization of the secure computability of a function g; and a decomposition result for the total entropy of the model, which lies at the heart of our technical approach. Proofs are provided in Section 3.4 and concluding remarks in Section 3.5. The results of this chapter were reported in [67, 69, 68].

3.2 Formulation: Secure function computation by public communication

Terminals $1, \ldots, m$ observe, respectively, the sequences $X_1^n, \ldots, X_m^n$, of length n. Let $g : \mathcal{X}_{\mathcal{M}} \to \mathcal{Y}$ be a given mapping, where $\mathcal{Y}$ is a finite alphabet. For $n \geq 1$, the mapping $g^n : \mathcal{X}_{\mathcal{M}}^n \to \mathcal{Y}^n$ is defined by

$$g^n(x_{\mathcal{M}}^n) = (g(x_{11}, \ldots, x_{m1}), \ldots, g(x_{1n}, \ldots, x_{mn})),$$
$$x_{\mathcal{M}}^n = (x_1^n, \ldots, x_m^n) \in \mathcal{X}_{\mathcal{M}}^n.$$

For convenience, we shall denote the rv $g^n\left(X_{\mathcal{M}}^n\right)$ by $G^n, n \geq 1$, and, in particular, $G^1 = g\left(X_{\mathcal{M}}\right)$ simply by G. The terminals in a given set $\mathcal{A} \subseteq \mathcal{M}$ wish to "compute securely" the function $g^n(x_{\mathcal{M}}^n)$ for $x_{\mathcal{M}}^n$ in $\mathcal{X}_{\mathcal{M}}^n$. To this end, the terminals are allowed to communicate over a noiseless public channel, possibly interactively in several rounds. Randomization at the terminals is permitted; we assume that terminal i generates a rv $U_i,\ i \in \mathcal{M}$, such that $U_1, \ldots, U_m$ and $X_{\mathcal{M}}^n$ are mutually independent. While the cardinalities of range spaces of $U_i, i \in \mathcal{M}$, are unrestricted, we assume that $H\left(U_{\mathcal{M}}\right) < \infty$ (see Definition 2.1).

Definition 3.1. For $\epsilon_n > 0, n \geq 1$, we say that g is ϵ_n-*securely computable* (ϵ_n- SC) by (the terminals in) a given set $\mathcal{A} \subseteq \mathcal{M}$ with $|\mathcal{A}| \geq 1$ from observations of length n, randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F} = \mathbf{F}^{(n)}$, if

(i) g^n is ϵ_n- recoverable from $(U_i, X_i^n, \mathbf{F})$ for every $i \in \mathcal{A}$, i.e., there exists $\widehat{g}_i^{(n)}$

satisfying

$$\mathrm{P}\left(\widehat{g}_i^{(n)}(U_i, X_i^n, \mathbf{F}) \neq G^n\right) \leq \epsilon_n, \quad i \in \mathcal{A}, \tag{3.1}$$

and

(ii) g^n satisfies the strong secrecy condition [49, 16, 19].

$$I(G^n \wedge \mathbf{F}) \leq \epsilon_n. \tag{3.2}$$

By definition, an ϵ_n-SC function g is recoverable (as g^n) at the terminals in $\mathcal{A}$ and is effectively concealed from an eavesdropper with access to the public communication $\mathbf{F}$.

Definition 3.2. We say that g is *securely computable* by $\mathcal{A}$ if g is ϵ_n- SC by $\mathcal{A}$ from observations of length n, suitable randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F}$, such that $\lim_n \epsilon_n = 0$.

Figure 3.1 shows our setup for secure computing.

3.3 When is a function securely computable?

We consider first the case when all the terminals in $\mathcal{M}$ wish to compute securely the function g, i.e., $\mathcal{A} = \mathcal{M}$. Our result for this case will be seen to be linked inherently to the standard concept of SK capacity for a multiterminal source model described in the previous chapter (see Definition 2.3), and serves to motivate our approach to the general case when $\mathcal{A} \subseteq \mathcal{M}$.

A comparison of the conditions in (3.2) and (2.6) that must be met by a securely computable g and an SK K, respectively, shows for a given g to be securely

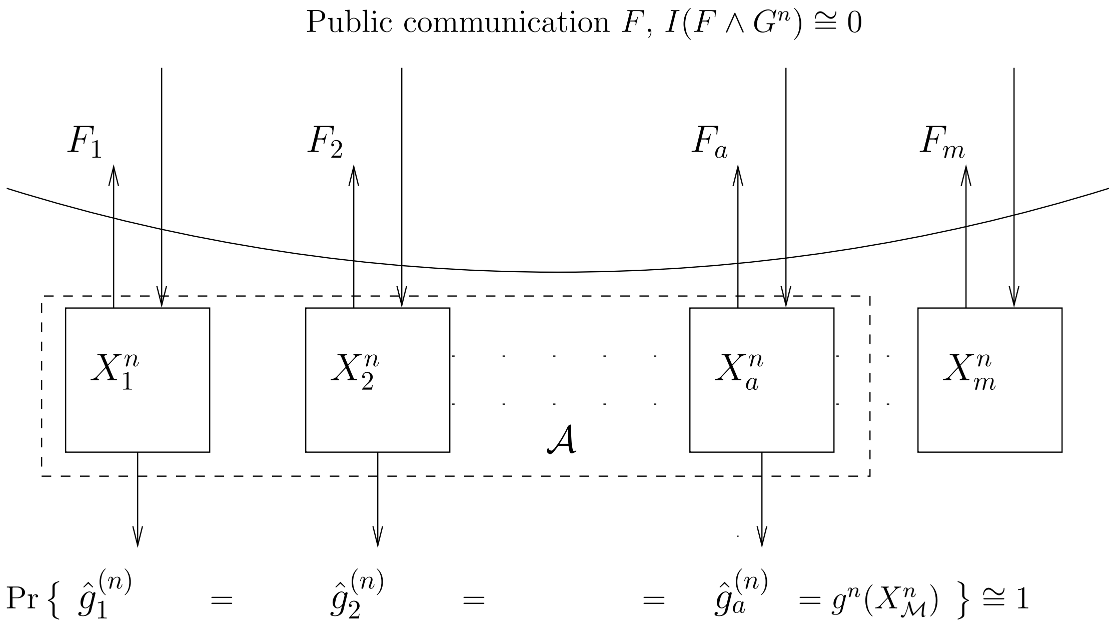
Public communication F, $I(F \wedge G^n) \cong 0$
F_1
F_2
F_a
F_m
X_1^n
X_2^n
X_a^n
X_m^n
$\mathcal{A}$
$\Pr\{\ \hat{g}_1^{(n)} = \hat{g}_2^{(n)} = \ = \hat{g}_a^{(n)} = g^n(X_{\mathcal{M}}^n)\ \} \cong 1$

Figure 3.1: Secure computation of g

computable, it is necessary that

$$H(G) \leq C(\mathcal{M}). \tag{3.3}$$

Remarkably, it transpires that $H(G) < C(\mathcal{M})$ is a sufficient condition for g to be securely computable, and constitutes our first result.

Theorem 3.1. *A function g is securely computable by $\mathcal{M}$ if*

$$H(G) < C(\mathcal{M}). \tag{3.4}$$

Conversely, if g is securely computable by $\mathcal{M}$, then $H(G) \leq C(\mathcal{M})$.

Theorem 3.1 is, in fact, a special case of our main result in Theorem 3.4 below.

Example 3.1. **Secure Computation of Parity.** Let $m = 2$, and let X_1 and X_2 be

$\{0,1\}$-valued rvs with

$$P_{X_1}(1) = p = 1 - P_{X_1}(0), \quad 0 < p < 1,$$

$$P_{X_2|X_1}(1 \mid 0) = P_{X_2|X_1}(0 \mid 1) = \delta, \quad 0 < \delta < \frac{1}{2};$$

such rvs X_1 and X_2 give rise to *binary symmetric sources* (BSS). Let $g(x_1, x_2) = x_1 + x_2 \mod 2$.

From Theorem 2.1, $C(\{1,2\}) = h(p*\delta) - h(\delta)$, where $p*\delta = (1-p)\delta + p(1-\delta)$. Since $H(G) = h(\delta)$, by Theorem 3.1 g is securely computable if

$$2h(\delta) < h(p * \delta). \tag{3.5}$$

We give a simple scheme for the secure computation of g when $p = 1/2$, that relies on Wyner's well-known method for Slepian-Wolf data compression [76] and a derived SK generation scheme in [82], [81]. When $p = 1/2$, we can write

$$X_1^n = X_2^n + G^n \mod 2 \tag{3.6}$$

with G^n being independent separately of X_2^n and X_1^n. We observe as in [76] that there exists a binary linear code, of rate $\cong 1 - h(\delta)$, with parity check matrix $\mathbf{P}$ such that X_1^n, and so G^n, is ϵ_n-recoverable from (F_1, X_2^n) at terminal 2, where the Slepian-Wolf codeword $F_1 = \mathbf{P}X_1^n$ constitutes public communication from terminal 1, and where ϵ_n decays to 0 exponentially rapidly in n. Let $\widehat{G^n}$ be the estimate of G^n thereby formed at terminal 2. (We can take $\widehat{G^n}$ to have been compressed losslessly to rate $H(G)$.) Further, let $K = K(X_1^n)$ be the location of X_1^n in the coset of the standard array corresponding to $\mathbf{P}$. By the previous observation, K

too is ϵ_n-recoverable from (F_1, X_2^n) at terminal 2. From [82], [81], K constitutes a "perfect" SK for terminals 1 and 2, of rate $\cong I(X_1 \wedge X_2) = 1 - h(\delta)$, and satisfying

$$I(K \wedge F_1) = 0. \tag{3.7}$$

Also, observe from (3.6) that $K = K(X_1^n) = K(X_2^n + G^n)$ and $F_1 = F_1(X_1^n) = F_1(X_2^n + G^n)$. Since G^n is independent of X_2^n, it follows that conditioned on each fixed value $G^n = g^n$, the (common) argument of K and F_1, namely $X_2^n + G^n$, has a conditional pmf that equals the pmf of $X_2^n + g^n$ which, in turn, coincides with the pmf of $X_1^n + g^n$, i.e., a permutation of the pmf of X_1^n. Hence by (3.7),

$$I(K \wedge F_1, G^n) = I(K \wedge F_1 \mid G^n) = 0, \tag{3.8}$$

since $I(K \wedge G^n) \leq I(X_1^n \wedge G^n) = 0$.

Then terminal 2 communicates $\widehat{G^n}$ in encrypted form as

$$F_2 = \widehat{G^n} + K \mod 2$$

(all represented in bits), with encryption feasible since

$$H(G) = h(\delta) < 1 - h(\delta) \cong \frac{1}{n} H(K),$$

by the sufficient condition (3.5). Terminal 1 then decrypts F_2 using K to recover $\widehat{G^n}$. The computation of g^n is secure since

$$I(G^n \wedge F_1, F_2) = I(G^n \wedge F_1) + I(G^n \wedge F_2 \mid F_1)$$

is small; specifically, the first term equals 0 since $I(G^n \wedge F_1) \leq I(G^n \wedge X_1^n) = 0$, while the second term is bounded according to

$$\begin{aligned} I(G^n \wedge F_2 \mid F_1) &= H(\widehat{G^n} + K \mid F_1) - H(\widehat{G^n} + K \mid F_1, G^n) \\ &\leq H(K) - H(G^n + K \mid F_1, G^n) + \delta_n, \\ &\qquad\qquad\qquad\qquad \text{with } \delta_n \to 0 \\ &= I(K \wedge F_1, G^n) + \delta_n = \delta_n, \end{aligned}$$

where the intermediate step uses Fano's inequality and the exponential decay of ϵ_n to 0, and the last equality is by (3.8). □

Example 3.2. Consider the setup of Example 3.1 for the case $p = 1/2$, but now with terminal 1 alone seeking to compute g. Since G^n is independent of X_2^n, secure computation of g at terminal 1 is possible with terminal 2 simply communicating X_2^n, even when X_1 and X_2 are independent. Note that

$$H(G) = h(\delta) \leq C(\{1\}) = H(X_1) = 1,$$

for $0 \leq \delta \leq 1/2$. □

We now turn to the general model for the secure computability of g by a given set $\mathcal{A} \subseteq \mathcal{M}$. Again in the manner of (3.3), it is clear that a necessary condition is

$$H(G) \leq C(\mathcal{A}).$$

In contrast, when $\mathcal{A} \subsetneq \mathcal{M}$, the condition $H(G) < C(\mathcal{A})$ is *not* sufficient for g to be securely computable by $\mathcal{A}$ as seen by the following simple example.

Example 3.3. **Omniscience is Forbidden.** Let $m = 3$, $A = \{1, 2\}$ and consider rvs X_1, X_2, X_3 with $X_1 = X_2$, where X_1 is independent of X_3 and $H(X_3) < H(X_1)$.

Let g be defined by $g(x_1, x_2, x_3) = x_3$, $x_i \in \mathcal{X}_i$, $1 \leq i \leq 3$. Clearly, $C(\{1, 2\}) = H(X_1)$. Therefore, $H(G) = H(X_3) < C(\{1, 2\})$. However, for g to be computed by the terminals 1 and 2, its value must be conveyed to them necessarily by public communication from terminal 3. Thus, g is not securely computable. □

We observe in Example 3.2 that if the value of G^n is given to terminal 2 after it has communicated X_2^n to terminal 1, then both terminals attain omniscience, with terminal 1 doing so from communication that is independent of G^n. Terminal 1 then computes G^n from its omniscience. Interestingly, the secure computability of g can be examined in terms of a new SK generation problem that contains these features and is formulated next.

3.3.1 Secret key aided by side information

We consider an extension of the SK generation problem in Definition 2.3, which involves additional side information $Z_{\mathcal{A}'}^n$ that is correlated with $X_{\mathcal{M}}^n$ and is provided to the terminals in $\mathcal{A}'$ for use in *only the recovery stage* of SK generation; however, the public communication $\mathbf{F}$ remains as in Definition 2.1. Formally, the extension is described in terms of generic rvs $(X_1, \ldots, X_m, \{Z_i, i \in \mathcal{A}'\})$, where the rvs Z_i too take values in finite sets $\mathcal{Z}_i$, $i \in \mathcal{A}'$. The full force of this extension will not be needed to characterize the secure computability of g; an appropriate particularization will suffice. Nevertheless, this concept is of independent interest.

Definition 3.3. A function K of $(X_{\mathcal{M}}^n, Z_{\mathcal{A}'}^n)$ is an ϵ_n- secret key aided by side information $Z_{\mathcal{A}'}^n$ (ϵ_n-ASK) for the terminals $\mathcal{A}' \subseteq \mathcal{M}$, $|\mathcal{A}'| \geq 2$, achievable from observations of length n, randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F} = \mathbf{F}(U_{\mathcal{M}}, X_{\mathcal{M}}^n)$ if it satisfies the conditions in Definition 2.3 with $(U_i, X_i^n, Z_i^n, \mathbf{F})$ in the role of $(U_i, X_i^n, \mathbf{F})$ in condition (i). The corresponding ASK capacity $C(\mathcal{A}', Z_{\mathcal{A}'})$ is defined analogously as in Definition 2.3.

In contrast with the omniscience rate of $H(X_{\mathcal{M}})$ that appears in the passage following Theorem 2.1, now an underlying analogous notion of omniscience will involve total CR of rate exceeding $H(X_{\mathcal{M}})$. Specifically, the enhanced CR rate will equal the entropy of the mcf of the rvs $(X_{\mathcal{M}}, Z_i)_{i \in \mathcal{A}}$, introduced for a pair of rvs in [26] (see also [17, Problem 3.4.27] and Chapter 2 above).

Definition 3.4. [26] For two rvs Q, R with values in finite sets $\mathcal{Q}, \mathcal{R}$, the equivalence relation $q \sim q'$ in $\mathcal{Q}$ holds if there exist $N \geq 1$ and sequences $(q_0, q_1, \ldots, q_N)$ in $\mathcal{Q}$ with $q_0 = q$, $q_N = q'$ and $(r_1, \ldots, r_N)$ in $\mathcal{R}$ satisfying $\mathrm{P}\left(Q = q_{l-1}, R = r_l\right) > 0$ and $\mathrm{P}\left(Q = q_l, R = r_l\right) > 0$, $l = 1, \ldots, N$. Denote the corresponding equivalence classes in $\mathcal{Q}$ by $\mathcal{Q}_1, \ldots, \mathcal{Q}_k$. Similarly, let $\mathcal{R}_1, \ldots, \mathcal{R}_{k'}$ denote the equivalence classes in $\mathcal{R}$. As argued in [26], $k = k'$ and for $1 \leq i, j \leq k$,

$$\mathrm{P}\left(Q \in \mathcal{Q}_i \mid R \in \mathcal{R}_j\right) = \mathrm{P}\left(R \in \mathcal{R}_j \mid Q \in \mathcal{Q}_i\right) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

The mcf of the rvs Q, R is a rv $\mathrm{mcf}(Q, R)$ with values in $\{1, \ldots, k\}$, defined by

$$\mathrm{mcf}(Q, R) = i \quad \text{iff} \quad Q \in \mathcal{Q}_i, R \in \mathcal{R}_i, \quad i = 1, \ldots, k.$$

For rvs $Q_1, ..., Q_m$ taking values in finite alphabets, we define the $\mathrm{mcf}(Q_1, ..., Q_m)$

recursively by

$$\text{mcf}(Q_1, ..., Q_m) = \text{mcf}\big(\text{mcf}(Q_1, ..., Q_{m-1}), Q_m\big) \tag{3.9}$$

with $\text{mcf}(Q_1, Q_2)$ as above.

Definition 3.5. With Q^n denoting n i.i.d. repetitions of the rv Q, we define

$$\text{mcf}^n(Q_1, ..., Q_m) = \left\{\text{mcf}\left(Q_{1t}, ..., Q_{mt}\right)\right\}_{t=1}^n . \tag{3.10}$$

Note that $\text{mcf}^n(Q_1, ..., Q_m)$ is a function of *each* individual $Q_i^n, i = 1, ..., m$.

Remark. As justification for the Definition (3.9), consider a rv ξ that satisfies

$$H(\xi \mid Q_i) = 0, \quad i = 1, ..., m \tag{3.11}$$

and suppose for any other rv ξ' satisfying (3.11) that $H(\xi) \geq H(\xi')$. Then Lemma 3.2 below shows that ξ must satisfy $H(\xi) = H(\text{mcf}(Q_1, ..., Q_m))$.

The following result for the mcf of $m \geq 2$ rvs is a simple extension of the classic result for $m = 2$ [26, Theorem 1].

Lemma 3.2. *Given $0 < \epsilon < 1$, if $\xi^{(n)}$ is ϵ-recoverable from Q_i^n for each $i = 1, ..., m$, then*

$$\limsup_n \frac{1}{n} H\left(\xi^{(n)}\right) \leq H(\text{mcf}(Q_1, ..., Q_m)). \tag{3.12}$$

Proof: The proof involves a recursive application of [26, Lemma, Section 4] to $\text{mcf}(Q_1, ..., Q_m)$ in (3.9), and is provided in Appendix A.

We are now in a position to characterize ASK capacity. In a manner analogous to Theorem 2.1, this is done in terms of $H(\text{mcf}(X_{\mathcal{M}}, Z_i)_{i \in \mathcal{A}'})$ and the smallest rate of communication $R_{CO}(\mathcal{A}', Z_{\mathcal{A}'})$ for each terminal in $\mathcal{A}'$ to attain omniscience that corresponds to n i.i.d. repetitions of mcf $((X_{\mathcal{M}}, Z_i)_{i\in\mathcal{A}'})$.

Theorem 3.3. *The ASK capacity $C(\mathcal{A}'; Z_{\mathcal{A}'})$ is given by*

$$
\begin{aligned}
&C(\mathcal{A}'; Z_{\mathcal{A}'}) = H(\mathrm{mcf}((X_{\mathcal{M}}, Z_i)_{i\in\mathcal{A}'})) - R_{CO}(\mathcal{A}'; Z_{\mathcal{A}'}), \ \textit{where}\\
&R_{CO}(\mathcal{A}'; Z_{\mathcal{A}'}) = \min_{R_{\mathcal{M}}\in\mathcal{R}(\mathcal{A}'; Z_{\mathcal{A}'})} \sum_{i\in\mathcal{M}} R_i, \ \textit{with}\\
&\mathcal{R}(\mathcal{A}'; Z_{\mathcal{A}'}) = \left\{ R_{\mathcal{M}} : \sum_{i\in B} R_i \geq \max_{j\in B^c\cap\mathcal{A}'} H(X_B \mid X_{B^c}, Z_j), B \subsetneq \mathcal{M}, \mathcal{A}' \nsubseteq B \right\}. \quad (3.13)
\end{aligned}
$$

The proof of Theorem 3.3 is along the same lines as that of Theorem 2.1 [20] and is provided in Appendix B.

The remark following Theorem 2.1 also applies to the ASK capacity $C(\mathcal{A}'; Z_{\mathcal{A}'})$, as will be seen from the proof of Theorem 3.3.

3.3.2 Characterization of secure computability

If g is securely computable by the terminals in $\mathcal{A}$, then G^n constitutes an ASK for $\mathcal{M}$ under the constraint (2.6), of rate $H(G)$, with side information in the form of G^n provided only to the terminals in $\mathcal{A}^c$ in the recovery stage of SK generation. Thus, a necessary condition for g to be securely computable by $\mathcal{A}$, in the manner of (3.3), is

$$H(G) \leq C(\mathcal{M}; Z_{\mathcal{M}}), \tag{3.14}$$

where $Z_{\mathcal{M}} = Z_{\mathcal{M}}(\mathcal{A}) = \{Z_i\}_{i \in \mathcal{M}}$ with

$$Z_i = \begin{cases} 0, & i \in \mathcal{A} \\ \\ G, & i \in \mathcal{A}^c. \end{cases} \tag{3.15}$$

By particularizing Theorem 3.3 to the choice of $Z_{\mathcal{M}}$ as above, the right side of (3.14) reduces to

$$C(\mathcal{M}; Z_{\mathcal{M}}) = H(X_{\mathcal{M}}) - R_{CO}(\mathcal{M}; Z_{\mathcal{M}}), \text{ where} \tag{3.16}$$

$$R_{CO}(\mathcal{M}; Z_{\mathcal{M}}) = \min_{R_{\mathcal{M}} \in \mathcal{R}(\mathcal{M}; Z_{\mathcal{M}})} \sum_{i \in \mathcal{M}} R_i, \text{ with}$$

$$\mathcal{R}(\mathcal{M}; Z_{\mathcal{M}}) = \left\{ R_{\mathcal{M}} : \sum_{i \in B} R_i \geq \begin{cases} H(X_B \mid X_{B^c}), & B \subsetneq \mathcal{M}, \mathcal{A} \not\subseteq B \\ H(X_B \mid X_{B^c}, G), & B \subsetneq \mathcal{M}, \mathcal{A} \subseteq B \end{cases} \right\}.$$

Our main result says that the necessary condition (3.14) is tight. Consider a protocol that enables the terminals in $\mathcal{M}$ to attain omniscience using communication that is independent of G^n, when G^n is provided only as "decoder side information" to the terminals in $\mathcal{A}^c$ but cannot be used for communication. Our proof shows that condition (3.17) below is sufficient for such a protocol to exist. Clearly, this protocol also serves for the secure computation of g by the terminals in $\mathcal{A}$ upon disregarding the decoding tasks in $\mathcal{A}^c$ (so that the protocol does not depend on a knowledge of G^n).

Theorem 3.4. *A function g is securely computable by $\mathcal{A} \subseteq \mathcal{M}$ if*

$$H(G) < C(\mathcal{M}; Z_{\mathcal{M}}). \tag{3.17}$$

Furthermore, under the condition above, g is securely computable with noninteractive communication and without recourse to randomization at the terminals in $\mathcal{M}$.

Conversely, if g is securely computable by $\mathcal{A} \subseteq \mathcal{M}$, then $H(G) \leq C(\mathcal{M}; Z_{\mathcal{M}})$.

Remarks. (i) As in the proof of achievability of SK capacity in [20], our proof of the sufficiency of (3.17) for the secure computability of g holds with ϵ_n in (3.1), (3.2) decaying to zero exponentially rapidly in n.

(ii) It is easy to see that $C(\mathcal{M}) \leq C\left(\mathcal{M}; Z_{\mathcal{M}}\right) \leq C(\mathcal{A})$, where $Z_{\mathcal{M}}$ is as in (3.15). In particular, the second inequality holds by noting that an SK for $\mathcal{M}$ is also an SK for $\mathcal{A}$, and that the side information for recovery $Z_{\mathcal{M}}$ in (3.15) is not provided to the terminals in $\mathcal{A}$.

(iii) Observe in Example 3 that $C\left(\mathcal{M}; Z_{\mathcal{M}}\right) = C(\mathcal{M}) = 0$ and so, by Theorem 3.4, g is not securely computable as noted earlier.

Example 3.4. **Secure Computing Using an SK.** In certain practical applications, different terminals observe mutually independent data and each seeks to securely compute a function g of the totality of all the observations. To enable this, they share a perfect SK, say K, of rate R. Then, since the SK capacity for this model is equal to R, by Theorem 3.1 a protocol for securely computing g exists if $H(G) < R$, and only if $H(G) \leq R$. Therefore, the terminals must share an SK of rate larger than $H(G)$ to accomplish secure computing.

Concretely, consider the case $m = 2$ with terminals 1 and 2 observing, respectively, random independent bits B_1 and B_2. Each terminal wishes to compute securely $B_1 \oplus B_2$. Furthermore, assume that the terminals share a one-bit SK K, which is independent of (B_1, B_2); thus, $X_1 = (B_1, K)$ and $X_2 = (B_2, K)$. Then, the following simple protocol ensures secure computing: $F_1 = B_1$, $F_2 = K \oplus B_1 \oplus B_2$.

In fact, this same protocol can be repeated n times to securely compute the parity for n i.i.d. pairs of independent random bits observed by the two terminals. But is this optimal or can we make do with less than n bits of SK? Heuristically, secure computing is feasible owing to the advantage that the legitimate parties have over the eavesdropper due to the correlation in their observations. In this simple example, this correlation corresponds to a shared SK. Therefore, the question raised above is, in effect, an "inverse" problem where we wish to quantify the minimum correlation needed to ensure secure computing. Specifically, since the SK capacity in this model is equal to the rate R of the SK, secure computing is feasible only if $R \geq H(B_1 \oplus B_2) = 1$ and so, the number of bits of SK cannot be (asymptotically) less than n for secure computing. Hence, the simple protocol above is asymptotically optimal. □

Example 3.5. **Secure Auction.** In an online auction, $m-1$ bidders acting independently of each other, randomly place one of k bids on a secure server. After a period of independent daily bidding, the server posts a cryptic message on a public website. We shall see that such a message exists from which each bidder can deduce securely the highest daily bids, but for $m > k+1$ no message exists to allow any of them to identify securely the daily winners.

Indeed, here $\mathcal{A} = \{1, ..., m-1\}$ and $X_1, ..., X_{m-1}$ are i.i.d. rvs distributed uniformly on $\{1, ..., k\}$, while $X_m = (X_1, ..., X_{m-1})$. Let $g_1(x_1, ..., x_m) = \max\limits_{1 \leq i \leq m-1} x_i$ and $g_2(x_1, ..., x_m) = \arg \max\limits_{1 \leq i \leq m-1} x_i$. Then, straightforward computation yields that $H(G_1) < \log k$, and for both g_1, g_2 that $C\left(\mathcal{M}; Z_{\mathcal{M}}\right) = C(\mathcal{M})$, where, by Theorem

2.1,

$$C(\mathcal{M}) = H(X_{\mathcal{M}}) - R_{CO}(\mathcal{M}) = (m-1)\log k - (m-2)\log k = \log k. \qquad (3.18)$$

Hence, by Theorem 3.4, g_1 is securely computable. Since $H(G_2) = \log(m-1)$, g_2 is securely computable if $k > m-1$. However, for $k < m-1$, g_2 is not securely computable by *any* terminal $i \in \{1, ..., m-1\}$. This, too, is implied by Theorem 3.4 upon noting that for each $i \in \{1, ..., m-1\}$ and a restricted choice $\mathcal{A} = \{i\}$ and $Z_{\mathcal{M}}$ as in (3.15),

$$C\left(\mathcal{M}; Z_{\mathcal{M}}\right) = H(X_i) = \log k < \log(m-1) = H(G_2),$$

where the first equality is a consequence of remark (ii) following Theorem 3.4, (3.18) and remark (i) following Theorem 2.1. □

3.3.3 A decomposition result

The sufficiency condition (3.17) prompts the following two natural questions: Does the difference $C\left(\mathcal{M}; Z_{\mathcal{M}}\right) - H(G)$ possess an operational significance? If g is securely computable by the terminals in $\mathcal{A}$, clearly G^n forms an SK for $\mathcal{A}$. Can G^n be augmented suitably to form an SK for $\mathcal{A}$ of maximum achievable rate?

The answers to both these questions are in the affirmative. In particular, our approach to the second question involves a characterization of the minimum rate of communication for omniscience for $\mathcal{A}$, under the additional requirement that this communication be independent of G^n. Specifically, we show below that for a securely computable function g, this minimum rate remains $R_{CO}(\mathcal{A})$ (see (2.4), (2.5)).

Addressing the first question, we introduce a rv $K_g = K_g^{(n)}$ such that $K = (K_g, G^n)$ constitutes an ϵ_n-ASK for $\mathcal{M}$ with side information $Z_{\mathcal{M}}$ as in (3.15) and satisfying the additional requirement

$$I\left(K_g \wedge G^n\right) \leq \epsilon_n. \tag{3.19}$$

Let the largest rate $\lim_n (1/n) H\left(K_g\right)$ of such an ASK be $C^g\left(\mathcal{M}; Z_{\mathcal{M}}\right)$. Observe that since K is required to be nearly independent of $\mathbf{F}$, where $\mathbf{F}$ is the public communication involved in its formation, it follows by (3.19) that K_g is nearly independent of $(G^n, \mathbf{F})$.

Turning to the second question, in the same vein let K_g' be a rv such that $K' = \left(K_g', G^n\right)$ constitutes an ϵ_n-SK for $\mathcal{A} \subseteq \mathcal{M}$ and satisfying (3.19). Let $C^g(\mathcal{A})$ denote the largest rate of K_g'. As noted above, K_g' will be nearly independent of $(G^n, \mathbf{F}')$, where $\mathbf{F}'$ is the public communication involved in the formation of K'.

Proposition 3.5. *If g satisfies (3.17), for $\mathcal{A} \subseteq \mathcal{M}$ it holds that*

$$\begin{aligned} &(i) \quad C^g\left(\mathcal{M}; Z_{\mathcal{M}}(\mathcal{A})\right) = C\left(\mathcal{M}; Z_{\mathcal{M}}(\mathcal{A})\right) - H(G), \\ &(ii) \qquad\qquad C^g(\mathcal{A}) = C(\mathcal{A}) - H(G). \end{aligned}$$

Remarks. (i) For the case $\mathcal{A} = \mathcal{M}$, both (i) and (ii) above reduce to $C^g(\mathcal{M}) = C(\mathcal{M}) - H(G)$.

(ii) Theorem 2.1 and Proposition 3.5 (ii) lead to the observation

$$H(X_{\mathcal{M}}) = R_{CO}(\mathcal{A}) + H(G) + C^g(\mathcal{A}),$$

which admits the following heuristic interpretation. The "total randomness" $X_{\mathcal{M}}^n$ that corresponds to omniscience decomposes into three "nearly mutually independent" components: a minimum-sized communication for omniscience for $\mathcal{A}$ and the independent parts of an optimum-rate SK for $\mathcal{A}$ composed of G^n and K_g'.

3.4 Proofs of main results

Proof of Theorem 3.4.

The necessity of (3.14) follows by the comments preceding Theorem 3.4.

The sufficiency of (3.17) will be established by showing the existence of *non-interactive* public communication comprising source codes that enable omniscience corresponding to $X_{\mathcal{M}}^n$ at the terminals in $\mathcal{A}$, and thereby the computation of g. Furthermore, the corresponding codewords are selected so as to be simultaneously independent of G^n, thus assuring security.

First, from (3.17) and (3.16), there exists $\delta > 0$ such that $R_{CO}(\mathcal{M}; Z_{\mathcal{M}}) + \delta < H(X_{\mathcal{M}}|G)$, using $G = g(X_{\mathcal{M}})$. For each i and $R_i \geq 0$, consider a (map-valued) rv J_i that is uniformly distributed on the family $\mathcal{J}_i$ of all mappings $\mathcal{X}_i^n \to \{1, \ldots, \lceil \exp(nR_i) \rceil\}$, $i \in \mathcal{M}$. The rvs $J_1, ..., J_m, X_{\mathcal{M}}^n$ are taken to be mutually independent.

Fix ϵ, ϵ', with $\epsilon' > m\epsilon$ and $\epsilon + \epsilon' < 1$. It follows from the proof of the general source network coding theorem [17, Lemma 3.1.13 and Theorem 3.1.14] that for all sufficiently large n,

$$\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : X_{\mathcal{M}}^n \text{ is } \epsilon_n\text{-recoverable from } \left(X_i^n, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^n\right), Z_i^n\right), i \in \mathcal{M}\right\}\right) \geq 1 - \epsilon, \quad (3.20)$$

provided $R_{\mathcal{M}} = (R_1, ..., R_m) \in \mathcal{R}(\mathcal{M}; Z_{\mathcal{M}})$, where ϵ_n vanishes exponentially rapidly in n. This assertion follows exactly as in the proof of [20, Proposition 1, with $A = \mathcal{M}$] but with $\tilde{X}_i$ there equal to (X_i, Z_i) rather than X_i, $i \in \mathcal{M}$. In particular, we shall choose $R_{\mathcal{M}} \in \mathcal{R}(\mathcal{M}; Z_{\mathcal{M}})$ such that

$$\sum_{i=1}^{m} R_i \leq R_{CO}(\mathcal{M}; Z_{\mathcal{M}}) + \frac{\delta}{2}. \tag{3.21}$$

Below we shall establish that

$$\mathrm{P}\left(\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : I\left(j_{\mathcal{M}}(X_{\mathcal{M}}^n) \wedge G^n\right) \geq \epsilon_n\}\right) \leq \epsilon', \tag{3.22}$$

for all n sufficiently large, to which end it suffices to show that

$$\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : I\left(j_i(X_i^n) \wedge G^n, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^n\right)\right) \geq \frac{\epsilon_n}{m}\right\}\right) \leq \frac{\epsilon'}{m}, \quad i \in \mathcal{M}, \tag{3.23}$$

since

$$\begin{aligned}
&I\left(j_{\mathcal{M}}\left(X_{\mathcal{M}}^n\right) \wedge G^n\right) \\
&= \sum_{i=1}^{m} I\left(j_i\left(X_i^n\right) \wedge G^n \mid j_1\left(X_1^n\right), \ldots, j_{i-1}\left(X_{i-1}^n\right)\right) \\
&\leq \sum_{i=1}^{m} I\left(j_i\left(X_i^n\right) \wedge G^n, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^n\right)\right).
\end{aligned}$$

Then it would follow from (3.20), (3.22) and definition of $Z_{\mathcal{M}}$ in (3.15) that

$$\begin{aligned}
\mathrm{P}\Bigg(\Bigg\{ j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : G^n \text{ is } \epsilon_n\text{-recoverable from } \left(X_i^n, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^n\right)\right), \quad i \in \mathcal{A}, \\
\text{and } I(j_{\mathcal{M}}(X_{\mathcal{M}}^n) \wedge G^n) < \epsilon_n \Bigg\}\Bigg) \geq 1 - \epsilon - \epsilon'.
\end{aligned}$$

This shows the existence of a particular realization $j_{\mathcal{M}}$ of $J_{\mathcal{M}}$ such that G^n is ϵ_n-SC from $\left(X_i^n, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^n\right)\right)$ for each $i \in \mathcal{A}$.

It now remains to prove (3.23). Fix $i \in \mathcal{M}$ and note that for each $j_i \in \mathcal{J}_i$, with $\|j_i\|$ denoting the cardinality of the (image) set $j_i(\mathcal{X}_i^n)$,

$$
\begin{aligned}
&I\Bigg(j_i\left(X_i^n\right) \wedge G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\Bigg) \\
&\leq I\left(j_i\left(X_i^n\right) \wedge G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\right) + \log\|j_i\| - H\left(j_i\left(X_i^n\right)\right) \\
&= D\Bigg(j_i(X_i^n), \left(G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\right) \| U_{j_i\left(\mathcal{X}_i^n\right)} \times \Bigg(G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\Bigg)\Bigg), \quad (3.24)
\end{aligned}
$$

where the right side above denotes the (Kullback-Leibler) divergence between the joint pmf of $j_i(X_i^n),\Big(G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\Big)$ and the product of the uniform pmf on $j_i(\mathcal{X}_i^n)$ and the pmf of $\Big(G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\Big)$. Using [20, Lemma 1], the right side of (3.24) is bounded above further by

$$
s_{var} \log \frac{\|j_i\|}{s_{var}}, \tag{3.25}
$$

where $s_{var} = s_{var}\Big(j_i(X_i^n); G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\Big)$ is the variational distance between the pmfs in the divergence above. Therefore, to prove (3.23), it suffices to show that

$$
\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : \, s_{var}\left(j_i(X_i^n); G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\right) \geq \frac{\epsilon_n}{m}\right\}\right) \leq \frac{\epsilon'}{m}, \quad i \in \mathcal{M}, \tag{3.26}
$$

on account of the fact that $\log\|j_i(X_i^n)\| = O(n)$, and the exponential decay to 0 of ϵ_n. Defining

$$
\tilde{\mathcal{J}}_i = \Bigg\{j_{\mathcal{M}\backslash\{i\}} \in \mathcal{J}_{\mathcal{M}\backslash\{i\}} : \, X_{\mathcal{M}}^n \text{ is } \epsilon_n\text{-recoverable from } \left(X_i^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right), Z_i^n\right)\Bigg\},
$$

we have by (3.20) that $\mathrm{P}\left(J_{\mathcal{M}\backslash\{i\}} \in \tilde{\mathcal{J}}_i\right) \geq 1 - \epsilon$. It follows that

$$
\begin{aligned}
&\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : s_{var}\left(j_i\left(X_i^n\right); G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\right) \geq \frac{\epsilon_n}{m}\right\}\right) \\
&\leq \epsilon + \sum_{j_{\mathcal{M}\backslash\{i\}} \in \tilde{\mathcal{J}}_i} \mathrm{P}\left(J_{\mathcal{M}\backslash\{i\}} = j_{\mathcal{M}\backslash\{i\}}\right) \times \\
&\qquad \mathrm{P}\left(\left\{j_i \in \mathcal{J}_i : s_{var}\left(j_i(X_i^n); G^n, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^n\right)\right) \geq \frac{\epsilon_n}{m}\right\}\right),
\end{aligned}
$$

since J_i is independent of $J_{\mathcal{M}\backslash\{i\}}$. Thus, (3.26), and hence (3.23), will follow upon showing that

$$\mathrm{P}\left(\left\{j_i \in \mathcal{J}_i : s_{var}\left(j_i(X_i^n); G^n, j_{\mathcal{M}\backslash\{i\}}\left(X^n_{\mathcal{M}\backslash\{i\}}\right)\right) \geq \frac{\epsilon_n}{m}\right\}\right) \leq \frac{\epsilon'}{m} - \epsilon, \quad j_{\mathcal{M}\backslash\{i\}} \in \tilde{\mathcal{J}}_i, \tag{3.27}$$

for all n sufficiently large. Fix $j_{\mathcal{M}\backslash\{i\}} \in \tilde{\mathcal{J}}_i$. We take recourse to Lemma 2.7 and set $U = X^n_{\mathcal{M}}, U' = X_i^n, V = G^n, h = j_{\mathcal{M}\backslash\{i\}}$, and

$$\mathcal{U}_0 = \left\{x^n_{\mathcal{M}} \in \mathcal{X}^n_{\mathcal{M}} : x^n_{\mathcal{M}} = \psi_i\left(x_i^n, j_{\mathcal{M}\backslash\{i\}}\left(x^n_{\mathcal{M}\backslash\{i\}}\right), g^n\left(x^n_{\mathcal{M}}\right)\mathbf{1}\left(i \in \mathcal{A}^c\right)\right)\right\}$$

for some mapping ψ_i. By the definition of $\tilde{\mathcal{J}}_i$,

$$\mathrm{P}\left(U \in \mathcal{U}_0\right) \geq 1 - \epsilon_n,$$

so that condition (2.22)(i) preceding Lemma 2.7 is met. Condition (2.22)(ii), too, is met since conditioned on the events in (2.22)(ii), only those $x^n_{\mathcal{M}} \in \mathcal{U}_0$ can occur that are determined uniquely by their i^{th} components x_i^n.

Upon choosing

$$d = \exp\left[n\left(H(X_{\mathcal{M}}|G) - \frac{\delta}{6}\right)\right],$$

in (2.23), the hypotheses of Lemma 2.7 are satisfied with $\lambda = \sqrt{\epsilon_n}$ for an appropriate exponentially vanishing ϵ_n. Then, by Lemma 2.7, with

$$r = \lceil \exp[nR_i] \rceil, \quad r' = \left\lceil \exp\left[n\left(\sum_{l \in \mathcal{M}\backslash\{i\}} R_l + \frac{\delta}{6}\right)\right]\right\rceil,$$

and with J_i in the role of ϕ, we get from (2.24) and (3.21) that

$$\mathrm{P}\left(\left\{j_i \in \mathcal{J}_i : \; s_{var}\left(j_i(X_i^n); G^n, j_{\mathcal{M}\backslash\{i\}}\left(X^n_{\mathcal{M}\backslash\{i\}}\right)\right) \geq 14\sqrt{\epsilon_n}\right\}\right)$$

decays to 0 doubly exponentially in n, which proves (3.27). This completes the proof of Theorem 3.4. □

Proof of Proposition 3.5. (i) Since the rv $(K_g^{(n)}, G^n)$, with nearly independent components, constitutes an ASK for $\mathcal{M}$ with side information $Z_{\mathcal{M}}$ as in (3.15), it is clear that

$$H(G) + C^g\left(\mathcal{M}; Z_{\mathcal{M}}\right) \leq C\left(\mathcal{M}; Z_{\mathcal{M}}\right). \tag{3.28}$$

In order to prove the reverse of (3.28), we show that $C\left(\mathcal{M}; Z_{\mathcal{M}}\right) - H(G)$ is an achievable ASK rate for K_g that additionally satisfies (3.19). First, note that in the proof of Theorem 3.4, the assertions (3.20) and (3.23) mean that for all sufficiently large n, there exists a public communication $F_{\mathcal{M}}$, say, such that $I(F_{\mathcal{M}} \wedge G^n) < \epsilon_n$ and $X_{\mathcal{M}}^n$ is ϵ_n-recoverable from $(X_i^n, F_{\mathcal{M}}, Z_i^n)$ for every $i \in \mathcal{M}$, with $\lim_n \epsilon_n = 0$. Fix $0 < \tau < \delta$, where δ is as in the proof of Theorem 3.4. Apply Lemma 2.7, choosing

$$\begin{aligned} &U = U' = X_{\mathcal{M}}^n, \quad \mathcal{U}_0 = \mathcal{X}_{\mathcal{M}}^n, \quad V = G^n, \quad h = F_{\mathcal{M}}, \\ &d = \exp\left[n\left(H\left(X_{\mathcal{M}}|G\right) - \frac{\tau}{6}\right)\right], \end{aligned} \tag{3.29}$$

whereby the hypothesis (2.23) of Lemma 2.7 is satisfied for all n sufficiently large. Fixing

$$r' = \left\lceil \exp\left[n\left(R_{CO}\left(\mathcal{M}; Z_{\mathcal{M}}\right) + \frac{\tau}{2}\right)\right]\right\rceil,$$

by Lemma 2.7 a randomly chosen ϕ of rate

$$\begin{aligned} \frac{1}{n}\log r &= H(X_{\mathcal{M}}|G) - R_{CO}\left(\mathcal{M}; Z_{\mathcal{M}}\right) - \tau \\ &= C\left(\mathcal{M}; Z_{\mathcal{M}}\right) - H(G) - \tau \end{aligned}$$

will yield an ASK $K_g = K_g^{(n)} = \phi\left(X_{\mathcal{M}}^n\right)$ which is nearly independent of $(F_{\mathcal{M}}, G^n)$ (and, in particular, satisfies (3.19)) with positive probability, for all n sufficiently large.

(ii) The proof can be completed as that of part (i) upon showing that for a securely computable g, for all $\tau > 0$ and n sufficiently large, there exists a public communication $F'_{\mathcal{M}}$ that meets the following requirements: its rate does not exceed $R_{CO}(\mathcal{A}) + \tau$; $I(F'_{\mathcal{M}} \wedge G^n) < \epsilon_n$; and $X_{\mathcal{M}}^n$ is ϵ_n-recoverable from $(X_i^n, F'_{\mathcal{M}})$ for every $i \in \mathcal{A}$. To that end, for $R_{\mathcal{M}} = (R_1, ..., R_m) \in \mathcal{R}(\mathcal{M}; Z_{\mathcal{M}})$ as in the proof of Theorem 3.4, consider $R'_{\mathcal{M}} = (R'_1, ..., R'_m) \in \mathcal{R}(\mathcal{A})$ that satisfies $R'_i \leq R_i$ for all $i \in \mathcal{M}$ and

$$\sum_{i=1}^{m} R'_i \leq R_{CO}(\mathcal{A}) + \tau,$$

noting that $\mathcal{R}\left(\mathcal{M}; Z_{\mathcal{M}}\right) \subseteq \mathcal{R}(\mathcal{A})$. Further, for $J_{\mathcal{M}}$ and $\mathcal{J}_{\mathcal{M}}$ as in that proof, define a (map-valued) rv J'_i that is uniformly distributed on the family $\mathcal{J}'_i$ of all mappings from $\{1, \ldots, \lceil \exp(nR_i) \rceil\}$ to $\{1, \ldots, \lceil \exp(nR'_i) \rceil\}$, $i \in \mathcal{M}$. The rvs $J_1, ..., J_m$, $J'_1, ..., J'_m, X_{\mathcal{M}}^n$ are taken to be mutually independent. Define $\mathcal{J}_{\mathcal{M}}^0$ as the set of mappings $j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}}$ for which there exists a $j'_{\mathcal{M}} \in \mathcal{J}'_{\mathcal{M}}$ such that $X_{\mathcal{M}}^n$ is ϵ_n-recoverable from $(X_i^n, j'_{\mathcal{M}}\left(j_{\mathcal{M}}\left(X_{\mathcal{M}}^n\right)\right))$ for every $i \in \mathcal{A}$. By the general source network coding theorem [17, Lemma 3.1.13 and Theorem 3.1.14], applied to the random mapping $J'_{\mathcal{M}}\left(J_{\mathcal{M}}\right)$, it follows that for all sufficiently large n,

$$\mathrm{P}\left(J_{\mathcal{M}} \in J_{\mathcal{M}}^0\right) \geq 1 - \epsilon.$$

This, together with (3.20) and (3.23) in the proof of Theorem 3.4, imply that for a securely computable g there exist $j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}}$ and $j'_{\mathcal{M}} \in \mathcal{J}'_{\mathcal{M}}$ for which the public communication $F'_{\mathcal{M}} \triangleq j'_{\mathcal{M}}(j_{\mathcal{M}})$ satisfies the aforementioned requirements. Finally, apply Lemma 2.7 with $U, U', \mathcal{U}_0, V$ and d as in (3.29) but with $h = F'_{\mathcal{M}}$ and

$$r' = \left\lceil \exp\left[n\left(R_{CO}\left(\mathcal{A}\right) + \frac{\tau}{2}\right)\right]\right\rceil.$$

As in the proof above of part (i), an SK $K'_g = K'^{(n)}_g$ of rate

$$\frac{1}{n}\log r = H(X_{\mathcal{M}}|G) - R_{CO}\left(\mathcal{A}\right) - \tau = C\left(\mathcal{A}\right) - H(G) - \tau$$

which is nearly independent of $(F'_{\mathcal{M}}, G^n)$ (and, hence, satisfies (3.19)) exists for all n sufficiently large. □

3.5 Discussion

We obtain simple necessary and sufficient conditions for secure computability expressed in terms of function entropy and ASK capacity. The latter is the largest rate of an SK for a new model in which side information is provided for use in only the recovery stage of SK generation. This model could be of independent interest. In particular, a function is securely computable if its entropy is less than the ASK capacity of an associated secrecy model. The difference is shown to correspond to the maximum achievable rate of an ASK which is independent of the securely computed function and, together with it, forms an ASK of optimum rate. Also, a function that is securely computed by $\mathcal{A}$ can be augmented to form an SK for $\mathcal{A}$ of maximum rate.

Our results extend trivially to functions defined on a block of symbols of *fixed* length in an obvious manner by considering larger alphabets composed of supersymbols of such length. However, they do not cover sequences of functions of symbols of increasing length (in n), e.g., a running average (in n).

In our proof of Theorem 3.4, g was securely computed from omniscience at all the terminals in $\mathcal{A} \subseteq \mathcal{M}$ that was attained using noninteractive public communication. However, omniscience is not necessary for the secure computation of g, and it is possible to make do with communication of rate less than $R_{CO}(\mathcal{A})$ using an interactive protocol. A related unresolved question is: What is the minimum rate of public communication for secure computation?

A natural generalization of the conditions for secure computability of g by $\mathcal{A} \subseteq \mathcal{M}$ given here entails a characterization of conditions for the secure computability of multiple functions $g_1, ..., g_k$ by subsets $\mathcal{A}_1, ..., \mathcal{A}_k$ of $\mathcal{M}$, respectively. This unsolved problem, in general, will not permit omniscience for any $\mathcal{A}_i, i = 1, ..., k$. For instance with $m = 2$, $\mathcal{A}_1 = \{1\}$, $\mathcal{A}_2 = \{2\}$, and X_1 and X_2 being independent, the functions $g_i(x_i) = x_i$, $i = 1, 2$, are securely computable trivially, but not through omniscience since, in this example, public communication is forbidden for the secure computation of g_1, g_2. The next chapter addresses a version of the mentioned generalization.

Yet another direction involves a model in which the terminals in $\mathcal{M}$ securely compute $G = g\left(X_{\mathcal{M}}\right)$, and the eavesdropper has additional access to correlated side information that may not be available to the terminals in $\mathcal{M}$. Specifically, the eavesdropper observes n i.i.d. repetitions Z^n of a $\mathcal{Z}$-valued rv Z that has a given joint pmf with $X_{\mathcal{M}}$, in addition to the public communication $\mathbf{F}$ of the terminals in

$\mathcal{M}$. The secrecy condition (2.2) is replaced by

$$I\left(G^n \wedge \mathbf{F} \mid Z^n\right) \leq \epsilon_n, \tag{3.30}$$

noting that G need not be independent of Z. Having computed g securely, the terminals in $\mathcal{M}$ can extract a rv $K = K\left(G^n\right)$, of rate $H(G \mid Z)$, that is (nearly) independent of Z^n. Together with (3.30), this means that K is similarly independent of $(\mathbf{F}, Z^n)$. Since K constitutes a *wiretap secret key* (WSK), its rate $H(G \mid Z)$ necessarily cannot exceed the corresponding WSK capacity [48, 1, 20]. A single-letter characterization of WSK capacity remains unresolved in general (cf. [31]). The sufficiency of the previous necessary condition is unclear even when WSK capacity is known. In the special circumstance in which the terminals in $\mathcal{M}$, too, have access to Z^n, a single-letter characterization of WSK capacity is known [20]. In this case, our proof technique shows that the aforementioned necessary condition is also sufficient.

CHAPTER 4

Secure Computation: Multiple Functions

4.1 Synopsis

This chapter generalizes and extends the results of the previous chapter to the computation of multiple given functions of the observations at the terminals while maintaining the privacy of a specified function. Specifically, multiple terminals observe correlated data and seek to compute functions of the data using interactive public communication. At the same time, it is required that the value of a private function of the data remain concealed from an eavesdropper observing this communication. In general, the private function and the functions computed by the terminals can be all different. We show that a class of functions are securely computable if and only if the conditional entropy of data given the value of private function is greater than the least rate of interactive communication required for a related multiterminal source-coding task. A single-letter formula is provided for this rate in special cases.

The problem of secure computing for multiple functions is formulated in the next section, followed by our results in Section 4.3. The proofs are given in Sections 4.4 and 4.5. The final section discusses alternative forms of the necessary conditions. The results of this chapter were reported in [62, 64].

4.2 Formulation: Secure computation of multiple functions

We consider a multiterminal source model for function computation using public communication, with a confidentiality requirement. Terminals $1, \ldots, m$ observe, respectively, the sequences $X_1^n, \ldots, X_m^n$ of length n. For $0 \leq i \leq m$, let $g_i : \mathcal{X}_\mathcal{M} \to \mathcal{Y}_i$ be given mappings, where the sets $\mathcal{Y}_i$ are finite. Further, for $0 \leq i \leq m$ and $n \geq 1$, the (single-letter) mapping $g_i^n : \mathcal{X}_\mathcal{M}^n \to \mathcal{Y}_i^n$ is defined by

$$g_i^n(x_\mathcal{M}^n) = (g_i(x_{11}, \ldots, x_{m1}), \ldots, g_i(x_{1n}, \ldots, x_{mn})),$$
$$x_\mathcal{M}^n = (x_1^n, \ldots, x_m^n) \in \mathcal{X}_\mathcal{M}^n.$$

For convenience, we shall denote the rv $g_i^n\left(X_\mathcal{M}^n\right)$ by $G_i^n, n \geq 1$, and, in particular, $G_i^1 = g_i\left(X_\mathcal{M}\right)$ simply by G_i.

Each terminal $i \in \mathcal{M}$ wishes to compute the function $g_i^n(x_\mathcal{M}^n)$, without revealing $g_0^n(x_\mathcal{M}^n)$, $x_\mathcal{M}^n \in \mathcal{X}_\mathcal{M}^n$. To this end, the terminals are allowed to communicate over a noiseless public channel, possibly interactively in several rounds. An interactive communication protocol is as in Definition 2.1 but, for simplicity, in this chapter we do not allow local randomization, i.e., $U_\mathcal{M} = \emptyset$. The rate of the interactive communication $\mathbf{F}$ is $\frac{1}{n} \log \|\mathbf{F}\|$.

Definition 4.1. For $\epsilon_n > 0$, $n \geq 1$, we say that functions[1] $g_\mathcal{M} = (g_0, g_1, ..., g_m)$, with private function g_0, are ϵ_n-*securely computable* (ϵ_n- SC) from observations of length n, and public communication $\mathbf{F} = \mathbf{F}^{(n)}$, if

(i) G_i^n is ϵ_n- recoverable from $(X_i^n, \mathbf{F})$ for every $i \in \mathcal{M}$, and

[1] The abuse of notation $g_\mathcal{M} = (g_0, g_1, ..., g_m)$ simplifies our presentation.

(ii) $\mathbf{F}$ satisfies the secrecy condition

$$\frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right) \leq \epsilon_n.$$

Remark. The definition of secrecy here corresponds to the notion of weak secrecy. When our results have a single-letter form, our achievability schemes for secure computing attain strong secrecy (see Definition 2.3 and remarks (ii), (iii) following Theorem 2.1).

By definition, for ϵ_n-SC functions $g_{\mathcal{M}}$, the private function G_0 is effectively concealed from an eavesdropper with access to the public communication $\mathbf{F}$.

Definition 4.2. For private function g_0, we say that functions $g_{\mathcal{M}}$ are *securely computable* if $g_{\mathcal{M}}$ are ϵ_n- SC from observations $X_{\mathcal{M}}^n$ of length n and public communication $\mathbf{F} = \mathbf{F}^{(n)}$, such that $\lim\limits_{n} \epsilon_n = 0$.

Figure 4.1 shows the setup for secure computing.

In this dissertation, we give necessary and sufficient conditions for the secure computability of certain classes of functions $g_{\mathcal{M}} = (g_0, g_1, ..., g_m)$. The formulation in Chapter 3, in which the terminals in a given subset $\mathcal{A}$ of $\mathcal{M}$ are required to compute (only) g_0 securely, is a special case with

$$g_i = \begin{cases} g_0, & i \in \mathcal{A}, \\ \\ \text{constant}, & \text{otherwise}. \end{cases} \tag{4.1}$$

Upon rearranging the terms in Theorem 3.4, we see that conditions

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq R^* \tag{4.2}$$

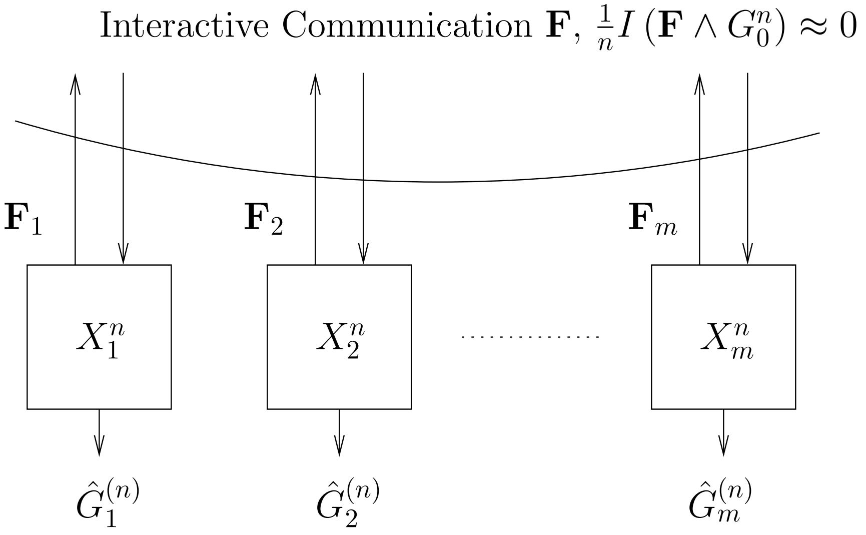

Figure 4.1: Secure computation of $g_1, ..., g_m$ with private function g_0

and

$$H\left(X_{\mathcal{M}} \mid G_0\right) > R^* \tag{4.3}$$

constitute, respectively, necessary and sufficient conditions for the functions above to be securely computable, with R^* being the minimum rate of interactive communication $\mathbf{F}$ that enables all the terminals in $\mathcal{M}$ to attain omniscience, using $\mathbf{F}$ and with decoder side information G_0^n given to the terminals in $\mathcal{M} \setminus \mathcal{A}$. In fact, it was shown that when condition (4.3) holds, it is possible to recover $X_{\mathcal{M}}^n$ using communication that is independent of G_0^n.

The guiding heuristic in this chapter is the following general principle, which is also consistent with the results of the previous chapter:

Conditions (4.2) and (4.3) constitute, respectively, the necessary and sufficient conditions for functions $g_{\mathcal{M}} = (g_0, g_1, ..., g_m)$ *to be securely computable, where* R^* *is the infimum of the rates of interactive communication* $\mathbf{F}'$ *such that, for each*

$1 \leq i \leq m$, *the following hold simultaneously:*

(P1) G_i^n *is* ϵ_n*-recoverable from* $(X_i^n, \mathbf{F}')$*, and*

(P2) $X_{\mathcal{M}}^n$ *is* ϵ_n*-recoverable from* $(X_i^n, G_0^n, \mathbf{F}')$*, i.e., terminals attain omniscience, with* G_0^n *as side information that is used only for decoding (but is not used for the communication* $\mathbf{F}'$*),*

where $\epsilon_n \to 0$ *as* $n \to \infty$.

Thus, (P1) and (P2) require any terminal computing g_0 to become omniscient, an observation that was also made for the special case in Chapter 3. The first condition (P1) above is straightforward and ensures the computability of the functions $g_1, ..., g_m$, by the terminals $1, ..., m$, respectively. The omniscience condition (P2) facilitates the decomposition of total entropy into mutually independent components that include the random values of the private function G_0^n and the communication $\mathbf{F}'$. For the specific case in (4.1), R^* above has a single-letter formula. In general, a single-letter expression for R^* is not known.

Our results, described in section 4.3, are obtained by simple adaptations of this principle. However, unlike in the previous chapter, our conditions, in general, are not of a single-letter form. Nevertheless, they provide a structural characterization of secure computability. As an application, our results provide simple conditions for secure computability in the following illustrative example.

Example 4.1. **Secure Computing for Binary Symmetric Sources.** We consider the case of $m = 2$ terminals that observe BSS (see Example 3.1) with underlying rvs X_1, X_2 with joint pmf given by

$$\mathrm{P}\left(X_1 = 0, X_2 = 0\right) = \mathrm{P}\left(X_1 = 1, X_2 = 1\right) = \frac{1-\delta}{2},$$
$$\mathrm{P}\left(X_1 = 0, X_2 = 1\right) = \mathrm{P}\left(X_1 = 1, X_2 = 0\right) = \frac{\delta}{2},$$

where $0 < \delta < 1/2$ (cf. Example 3.1). The results of this dissertation will allow us to provide conditions for the secure computability of the four choices of g_0, g_1, g_2 below; it will follow by Theorem 4.1 that functions g_0, g_1, g_2 are securely computable if

$$h(\delta) < \tau,$$

and conversely, if the functions above are securely computable, then

$$h(\delta) \leq \tau,$$

where $h(\tau) = -\tau \log \tau - (1-\tau) \log(1-\tau)$, and the constant $\tau = \tau(\delta)$ depends on the choice of the function. These characterizations are summarized in the next table. Denote the AND and the OR of two random bits X_1 and X_2 by $X_1.X_2$ and $X_1 \oplus X_2$, respectively.

g_0	g_1	g_2	τ
$X_1 \oplus X_2$	$X_1 \oplus X_2$	$X_1 \oplus X_2$	$1/2$
$X_1 \oplus X_2$	$X_1 \oplus X_2$	ϕ	1
$X_1 \oplus X_2,\ X_1.X_2$	$X_1 \oplus X_2,\ X_1.X_2$	$X_1.X_2$	$2\delta/3$
$X_1 \oplus X_2$	$X_1 \oplus X_2$	$X_1.X_2$	$2/3$

The results for the first two settings follow from Examples 3.1, 3.2. The third and fourth results are new. In these settings, terminal 1 is required to recover the private function; our results below show that the conditions for the secure computability in these cases remain unchanged even if this terminal is required to attain omniscience. Note that since $h(\delta) < 1$ for all $0 < \delta < 1/2$, there exists a communication protocol for securely computing the functions in the second setting. By contrast, a secure computing protocol for the functions in the third setting does not exist for any $0 < \delta < 1/2$, since $h(\delta) > 2\delta/3$. □

4.3 Characterization of securely computable functions

In this section, we characterize securely computable functions for three settings. Our necessary and sufficient conditions entail the comparison of $H\left(X_{\mathcal{M}}|G_0\right)$ with a rate R^*; the specific choice of R^* depends on the functions $g_{\mathcal{M}}$. Below we consider three different classes of functions $g_{\mathcal{M}}$. Although the first class is a special case of the second, the two are handled separately as the more restrictive case is amenable to simpler analysis. Furthermore, for $m = 2$, the obtained necessary and sufficient conditions for secure computability take a single-letter form in the first case (see Corollary 4.4).

(1) In the first class we consider, values of all the functions $g_1, ..., g_m$ must be kept secret. In addition, at least one of the terminals must compute all the functions $g_1, ..., g_m$. This case arises in distributed function computation over a network where all the computed values are collated at a single sink node, and we are interested in securing the collated function values. Alternatively, denoting the function computed at the sink node by the private function g_0, the computed functions $g_1, ..., g_m$ can be restricted to be functions of g_0. Specifically, for $0 < m_0 < m$, and for private function g_0, let

$$g_i = \begin{cases} g_0, & i \in [1, m_0], \\ g_i(g_0), & i \in [m_0 + 1, m]. \end{cases} \tag{4.4}$$

(2) The next case is a relaxation of the previous model in that the restriction $g_i = g_i(g_0)$ for $i \in [m_0 + 1, m]$ is dropped. For this general case, our analysis below implies roughly that requiring the terminals $[1, m_0]$ that compute the private function g_0 to recover the entire data $X^n_{\mathcal{M}}$ does not change the conditions for secure computability, which is a key observation of this dissertation.

(3) The last class of problems we study is an instance of *secure multiterminal source coding*, which arises in the data download problems in sensor networks where each node is interested in downloading the data observed by a subset of nodes. Specifically, we consider the situation where each terminal wishes to recover some subset $X^n_{\mathcal{M}_i}$ of the sources where $\mathcal{M}_i \subseteq \mathcal{M} \setminus \{i\}$, i.e.,

$$g_i(X_{\mathcal{M}}) = X_{\mathcal{M}_i}, \quad i \in \mathcal{M}. \tag{4.5}$$

This last case appears at first blush to be disconnected from the previous two cases. However, our characterizations of secure computability below have the same form for all cases above. Moreover, the same heuristic principle, highlighted in (P1) and (P2), leads to a characterization of secure computability in all three cases.

The necessary and sufficient conditions for secure computability are stated in terms of quantities $R^*_i(g_{\mathcal{M}})$, $i = 1, 2, 3$, which are defined next. The subscript

i corresponds to case (i) above. In particular, the quantity R^* corresponds to the minimum rate of communication needed for an appropriate modification of the source-coding task in (P1), (P2). Below we give specific expressions for R_i^*, $i = 1, 2, 3$, along with their operational roles (for a complete description of this role see the sufficiency proof in Section 4.4).

Denote by $\mathcal{R}_1^*\left(g_{\mathcal{M}}\right)$ the closure of the (nonempty) set of pairs[2]

$$\left(R_{\mathbf{F}}^{(1)}, \frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right)\right),$$

for all $n \geq 1$ and interactive communication $\mathbf{F}$, where

$$R_{\mathbf{F}}^{(1)} = \frac{1}{n} H(\mathbf{F}) + \frac{1}{n} \sum_{i=m_0+1}^{m} H\left(G_i^n | X_i^n, \mathbf{F}\right) + \inf R_{\mathcal{M}}, \tag{4.6}$$

with the infimum taken over the rates $R_1, ..., R_m$ satisfying the following constraints:

(1a) $\forall \mathcal{L} \subsetneq \mathcal{M}$, $[1, m_0] \nsubseteq \mathcal{L}$,

$$R_{\mathcal{L}} \geq \frac{1}{n} H\left(X_{\mathcal{L}}^n | X_{\mathcal{M}\setminus\mathcal{L}}^n, \mathbf{F}\right);$$

(1b) $\forall \mathcal{L} \subsetneq \mathcal{M}$, $[1, m_0] \subseteq \mathcal{L}$,

$$R_{\mathcal{L}} \geq \frac{1}{n} H\left(X_{\mathcal{L}}^n | X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right).$$

The quantity $\inf_{n,\mathbf{F}} R_{\mathbf{F}}^{(1)}$ corresponds to the solution of a multiterminal source coding problem. Specifically, it is the infimum of the rates of interactive communication that satisfy (P1) and (P2) above (see [18, Theorem 13.15], [20]).

[2] The first term accounts for the rate of the communication and the second term tracks the information about G_0^n leaked by $\mathbf{F}$ (see (4.11)) below

Next, let $\mathcal{R}_2^*\left(g_{\mathcal{M}}\right)$ denote the closure of the set of pairs

$$\left(R_{\mathbf{F}}^{(2)}, \frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right)\right),$$

for all $n \geq 1$ and interactive communication $\mathbf{F}$, where

$$R_{\mathbf{F}}^{(2)} = \frac{1}{n} H(\mathbf{F}) + \inf\left[R'_{[m_0+1,m]} + R_{\mathcal{M}}\right], \tag{4.7}$$

with the infimum taken over the rates $R_1, \ldots, R_m$ and $R'_{m_0+1}, \ldots, R'_m$ satisfying the following constraints:

(2a) $\forall \mathcal{L} \subsetneq \mathcal{M}$, $[1, m_0] \nsubseteq \mathcal{L}$,

$$R_{\mathcal{L}} \geq \frac{1}{n} H\left(X_{\mathcal{L}}^n | X_{\mathcal{M}\setminus\mathcal{L}}^n, \mathbf{F}\right);$$

(2b) for $m_0 < j \leq m$,

$$R'_j \geq \frac{1}{n} H\left(G_j^n | X_j^n, \mathbf{F}\right);$$

(2c) $\forall \mathcal{L} \subseteq \mathcal{M}, [1, m_0] \subseteq \mathcal{L}$, and $\mathcal{L}' \subseteq [m_0+1, m]$ with either $\mathcal{L} \neq \mathcal{M}$ or $\mathcal{L}' \neq [m_0+1, m]$,

$$R'_{\mathcal{L}'} + R_{\mathcal{L}} \geq \frac{1}{n} H\left(G_{\mathcal{L}'}^n, X_{\mathcal{L}}^n | G_{[m_0+1,m]\setminus\mathcal{L}'}^n, X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right).$$

The quantity $\inf_{n,\mathbf{F}} R_{\mathbf{F}}^{(2)}$ corresponds to the solution of a multiterminal source coding problem, and is the infimum of the rates of interactive communication $\mathbf{F}'$ that satisfy (P1) and (P2) above, and additionally satisfies:

(P3) *$X_{\mathcal{M}}^n$ is ϵ_n-recoverable from $\left(G_j^n, G_0^n, \mathbf{F}'\right)$, $m_0 < j \leq m$.*

This modification corresponds to the introduction of $m - m_0$ dummy terminals, with the jth dummy terminal observing G_j^n, $m_0 < j \leq m$ (see section 4.6); the dummy terminals can be realized by a terminal i in $[1, ..., m_0]$ that recovers $X_{\mathcal{M}}^n$ from $(X_i^n, \mathbf{F})$. The conditions (P2) and (P3) above correspond to the omniscience at the terminals in the extended model, with G_0^n provided as side information only for decoding.

Finally, denote by $\mathcal{R}_3^*(g_{\mathcal{M}})$ the closure of the set of pairs

$$\left(R_{\mathbf{F}}^{(3)}, \frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right)\right),$$

for all interactive communication $\mathbf{F}$, where

$$R_{\mathbf{F}}^{(3)} = \frac{1}{n} H(\mathbf{F}) + \inf R_{\mathcal{M}}, \tag{4.8}$$

with rates $R_1, ..., R_m$ satisfying the following constraints:

(3a) For $1 \leq i \leq m$, $\forall\, \mathcal{L} \subseteq \mathcal{M}_i \subseteq \mathcal{M} \setminus \{i\}$,

$$R_{\mathcal{L}} \geq \frac{1}{n} H\left(X_{\mathcal{L}}^n | X_{\mathcal{M}_i \setminus \mathcal{L}}^n, X_i^n, \mathbf{F}\right);$$

(3b) $\forall \mathcal{L} \subsetneq \mathcal{M}$,

$$R_{\mathcal{L}} \geq \frac{1}{n} H\left(X_{\mathcal{L}}^n | X_{\mathcal{M} \setminus \mathcal{L}}^n, G_0^n, \mathbf{F}\right).$$

As before, the quantity $\inf_{n,\mathbf{F}} R_{\mathbf{F}}^{(3)}$ corresponds to the infimum of the rates of interactive communication that satisfy (P1) and (P2) above.

Our main result below characterizes securely computable functions for the three settings above.

Theorem 4.1. *For $i = 1, 2, 3$, with functions $g_0, g_1, ..., g_m$ as in the case (i) above, the functions $g_{\mathcal{M}}$ are securely computable if the following condition holds:*

$$H\left(X_{\mathcal{M}}|G_0\right) > R_i^*\left(g_{\mathcal{M}}\right). \tag{4.9}$$

Conversely, if the functions above are securely computable, then

$$H\left(X_{\mathcal{M}}|G_0\right) \geq R_i^*\left(g_{\mathcal{M}}\right), \tag{4.10}$$

where

$$R_i^*\left(g_{\mathcal{M}}\right) = \inf_{(x,0)\in\mathcal{R}_i^*(g_{\mathcal{M}})} x, \quad i = 1, 2, 3. \tag{4.11}$$

Remark. Although the first setting above is a special case of the second, it is unclear if for $g_{\mathcal{M}}$ in (4.4) the quantities $R_1^*(g_{\mathcal{M}})$ and $R_2^*(g_{\mathcal{M}})$ are identical (also, see Section 4.6). In general, the multi-letter characterizations of secure computability of $g_{\mathcal{M}}$ above can have different forms. For case (1) with $m = 2$, Corollary 4.4 below provides a single-letter formula for $R_1^*(g_{\mathcal{M}})$. However, a similar single-letter formula for $R_2^*(g_{\mathcal{M}})$ is not known.

Theorem 4.1 affords the following heuristic interpretation. The quantity $H\left(X_{\mathcal{M}}|G_0\right)$ represents the maximum rate of randomness in $X_{\mathcal{M}}^n$ that is (nearly) independent of G_0^n. On the other hand, $R_i^*\left(g_{\mathcal{M}}\right)$ is an appropriate rate of communication for the computation of $g_{\mathcal{M}}$; we show that latter being less than $H\left(X_{\mathcal{M}}|G_0\right)$ guarantees the secure computability of $g_{\mathcal{M}}$.

Although the characterization in Theorem 4.1 is not of a single-letter form, the following result provides a sufficient condition for obtaining such forms. Denote by $R_{\text{constant}}^{(i)}$, $i = 1, 2, 3$, the quantity $R_{\mathbf{F}}^{(i)}$ for $\mathbf{F} = \text{constant}$.

Lemma 4.2. *For case (i), $i = 1, 2, 3$, if for all $n \geq 1$ and interactive communication* $\mathbf{F}$

$$R_{\mathbf{F}}^{(i)} \geq R_{constant}^{(i)}, \tag{4.12}$$

then $R_i^* \left(g_{\mathcal{M}}\right) = R_{constant}^{(i)} = \inf_{n,\mathbf{F}} R_{\mathbf{F}}^{(i)}$.

The proof is a simple consequence of the definition of $R_i^* \left(g_{\mathcal{M}}\right)$ in (4.11). Note that $R_{\text{constant}}^{(i)}$ has a single-letter form.

Remark. As mentioned before, the quantity $\inf_{n,\mathbf{F}} R_{\mathbf{F}}^{(i)}$ is the infimum of the rates of interactive communication that satisfies (P1), (P2) for $i = 1, 3$, and satisfies (P1)-(P3) for $i = 2$. Thus, when the conditions of Lemma 4.2 hold, we have from Theorem 4.1 that $g_{\mathcal{M}}$ are securely computable if

$$H \left(X_{\mathcal{M}} | G_0\right) > R_{\text{constant}}^{(i)},$$

and if $g_{\mathcal{M}}$ are securely computable then

$$H \left(X_{\mathcal{M}} | G_0\right) \geq R_{\text{constant}}^{(i)},$$

where $R_{\text{constant}}^{(i)}$ is the minimum rate of communication that satisfies (P1), (P2) for $i = 1, 3$, and satisfies (P1)-(P3) for $i = 2$.

As a consequence of Lemma 4.2, we obtain below a single-letter characterization of securely computable functions, with $m = 2$, in a special case; the following lemma is instrumental to our proof.

Lemma 4.3. *Let $m = 2$. For an interactive communication $\mathbf{F}$, we have*

$$H(\mathbf{F}) \geq H \left(\mathbf{F} | X_1^n\right) + H \left(\mathbf{F} | X_2^n\right).$$

Proof. Lemma 4.3 is a special case of [21, Lemma B.1] (also, see [46]). We provide a proof here for completion.

$$\begin{aligned}
H(\mathbf{F}) &= H(F_1) + H(\mathbf{F} \mid F_1) \\
&\geq H(F_1 \mid Y^n) + H(\mathbf{F} \mid F_1) \\
&\geq H(F_1 \mid Y^n) + H(F_2 \mid F_1) + H(\mathbf{F} \mid F_1, F_2) \\
&\geq H(F_1 \mid Y^n) + H(F_2 \mid F_1, X^n) + H(\mathbf{F} \mid F_1, F_2) \\
&= H(F_1, F_2 \mid Y^n) + H(F_1, F_2 \mid X^n) + H(\mathbf{F} \mid F_1, F_2),
\end{aligned}$$

where the last step uses $H(F_1 \mid X^n) = H(F_2 \mid F_1, Y^n) = 0$. The proof is completed by an iterative application of these steps. □

We next consider case (1) for two terminals.

Corollary 4.4. *For $m = 2$, for functions g_0, g_1, g_2 with $g_1 = g_0$ and $g_2 = g_2(g_0)$, we have*

$$R_1^*(g_{\mathcal{M}}) = H(X_2|X_1) + H(G_2|X_2) + H(X_1|X_2, G_0). \tag{4.13}$$

Proof: The constraints (1a) and (1b) satisfied by rates R_1, R_2 in the definition of $R_{\mathbf{F}}^{(1)}$ are

$$\begin{aligned}
R_2 &\geq \frac{1}{n} H(X_2^n | X_1^n, \mathbf{F}), \\
R_1 &\geq \frac{1}{n} H(X_1^n | X_2^n, G_0^n, \mathbf{F}),
\end{aligned}$$

which further yields

$$R_{\mathbf{F}}^{(1)} = \frac{1}{n}\left[H(\mathbf{F}) + H(G_2^n | X_2^n, \mathbf{F}) + H(X_2^n | X_1^n, \mathbf{F}) + H(X_1^n | X_2^n, G_0^n, \mathbf{F})\right]. \tag{4.14}$$

Thus, $R^{(1)}_{\text{constant}}$ equals the term on the right side of (4.13). Upon manipulating the expression for $R^{(1)}_{\mathbf{F}}$ above, we get

$$R^{(1)}_{\mathbf{F}} = \frac{1}{n}\left[H(\mathbf{F}) - H\left(\mathbf{F}|X_1^n\right) - H\left(\mathbf{F}|X_2^n, G_0^n\right) - I\left(G_2^n \wedge \mathbf{F}|X_2^n\right)\right] + R^{(1)}_{constant}. \quad (4.15)$$

Further, since $H\left(G_2|G_0\right) = 0$, it holds that

$$I\left(G_2^n \wedge \mathbf{F}|X_2^n\right) \leq I\left(G_0^n \wedge \mathbf{F}|X_2^n\right),$$

which along with (4.15) yields

$$\begin{aligned} R^{(1)}_{\mathbf{F}} &\geq \frac{1}{n}\Bigg[H(\mathbf{F}) - H\left(\mathbf{F}|X_1^n\right) - H\left(\mathbf{F}|X_2^n\right)\Bigg] + R^{(1)}_{constant} \\ &\geq R^{(1)}_{constant}, \end{aligned}$$

where the last inequality follows from Lemma 4.3. The result then follows from Lemma 4.2. □

We next derive simple conditions for secure computability for the BSS in Example 4.1

Example 4.2. Consider the setup of Example 4.1, with $g_0 = g_1 = X_1 \oplus X_2, X_1.X_2$ and $g_2 = X_1.X_2$. By Corollary 4.4 and the observation $H\left(G_2|X_2\right) = h(\delta)/2$, we get $R_1^*\left(g_{\mathcal{M}}\right) = 3h(\delta)/2$. Since $H\left(X_1, X_2 \mid G_0\right) = H\left(X_1, X_2 \mid X_1 \oplus X_2\right) - H\left(X_1.X_2 \mid X_1 \oplus X_2\right) = \delta$, the characterization of secure computability claimed in Example 4.1 follows from Theorem 4.1. □

Example 4.3. In the setup of Example 4.1, consider $g_0 = g_1 = X_1 \oplus X_2$ and $g_2 = X_1.X_2$. This choice of g_0, g_1, g_2 is an instance of case (2) above. For an interactive communication $\mathbf{F}$, the constraints (2a), (2b), (2c) in the definition of $R^{(2)}_{\mathbf{F}}$, upon simplification, reduce to

$$\begin{aligned}
R_1 &\geq \frac{1}{n} H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right), \\
R_2 &\geq \frac{1}{n} H\left(X_2^n | X_1^n, \mathbf{F}\right), \\
R_1 + R_2 &\geq \frac{1}{n} H\left(X_1^n, X_2^n | G_0^n, G_2^n, \mathbf{F}\right), \\
R_2' &\geq \frac{1}{n} H\left(G_2^n | X_2^n, \mathbf{F}\right).
\end{aligned}$$

Therefore, $\inf\left[R_1 + R_2 + R_2'\right]$ with R_1, R_2, R_2' satisfying (2a), (2b), (2c), is given by

$$\frac{1}{n}\Bigg[H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right) + \max\left\{H\left(X_2^n | G_0^n, G_2^n, \mathbf{F}\right), H\left(X_2^n | X_1^n, \mathbf{F}\right)\right\} \\ + H\left(G_2^n | X_2^n, \mathbf{F}\right)\Bigg],$$

which further gives

$$R_{\mathbf{F}}^{(2)} = \frac{1}{n}\Bigg[H(\mathbf{F}) + H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right) + \max\left\{H\left(X_2^n | G_0^n, G_2^n, \mathbf{F}\right), H\left(X_2^n | X_1^n, \mathbf{F}\right)\right\} \\ + H\left(G_2^n | X_2^n, \mathbf{F}\right)\Bigg]. \qquad (4.16)$$

It follows from $H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right) = 0$ that

$$\begin{aligned}
R_{constant}^{(2)} &= H\left(G_2 | X_2\right) + \max\left\{H\left(X_2 | G_0, G_2\right), H\left(X_2 | X_1\right)\right\} \\
&= \frac{h(\delta)}{2} + \max\left\{\delta, h(\delta)\right\} = \frac{3}{2} h(\delta), \qquad (4.17)
\end{aligned}$$

as $h(\delta) > \delta$ for $0 < \delta < 1/2$.

Next, note from (4.16) that for any interactive communication $\mathbf{F}$

$$\begin{aligned} R_{\mathbf{F}}^{(2)} &\geq \frac{1}{n}\left[H(\mathbf{F}) + H\left(X_2^n|X_1^n, \mathbf{F}\right) + H\left(G_2^n|X_2^n, \mathbf{F}\right)\right] \\ &= \frac{1}{n}\left[H(\mathbf{F}) + H\left(X_2^n|X_1^n\right) - H\left(\mathbf{F}|X_1^n\right) + H\left(G_2^n, \mathbf{F}|X_2^n\right) - H\left(\mathbf{F}|X_2^n\right)\right] \\ &\geq \frac{1}{n}\left[H(\mathbf{F}) - H\left(\mathbf{F}|X_1^n\right) - H\left(\mathbf{F}|X_2^n\right)\right] + H\left(G_2|X_2\right) + H\left(X_2|X_1\right) \\ &\geq H\left(G_2|X_2\right) + H\left(X_2|X_1\right) = \frac{3}{2}h(\delta), \end{aligned} \tag{4.18}$$

where the last inequality above follows from Lemma 4.3. The characterization in Example 4.1 follows from (4.17), (4.18), and $H\left(X_1, X_2|G_0\right) = 1$, using Lemma 4.2 and Theorem 4.1. □

4.4 Proof of sufficiency

Sufficiency of (4.9) for $i = 1$: We propose a two step protocol for securely computing $g_0, g_1, ..., g_m$. In the first step, for sufficient large N, the terminals $[1, m_0]$ (g_0-seeking terminals) attain omniscience, using an interactive communication $\mathbf{F}'' = \mathbf{F}''\left(X_{\mathcal{M}}^N\right)$ that satisfies

$$\frac{1}{N} I\left(G_0^N \wedge \mathbf{F}''\right) \leq \epsilon, \tag{4.19}$$

where $\epsilon > 0$ is sufficiently small. Next, upon attaining omniscience, one of the terminals in $[1, m_0]$ computes the following for $m_0 < j \leq m$:

(i) Slepian-Wolf codewords $\hat{F}_j = \hat{F}_j\left(G_j^N\right)$ of appropriate rates R_j' for a recovery of G_j^N by a decoder with the knowledge of X_j^N and previous communication $\mathbf{F}''$, and

(ii) the rvs $K_j = K_j\left(X_j^N\right)$ of rates R'_j that satisfy:

$$\left|\frac{1}{N} H\left(K_j\right) - R'_j\right| \leq \epsilon, \tag{4.20}$$

$$\frac{1}{N} I\left(K_j \wedge G_0^N, \mathbf{F}'', \left\{K_l \oplus \hat{F}_l\right\}_{m_0 < l \leq j-1}\right) \leq \epsilon. \tag{4.21}$$

Note that $K_j \oplus \hat{F}_j$ denotes the encrypted version of the Slepian-Wolf code $\hat{F}_j$, encrypted with a one-time pad using the SK K_j. Thus, terminal j, with the knowledge of K_j, can recover $\hat{F}_j$ from $K_j \oplus \hat{F}_j$, and hence can recover G_j^N. The operation $K_j \oplus \hat{F}_j$ is valid since the SK K_j has size greater than $\|\hat{F}_j\|$. Furthermore, we have from (4.19) and (4.21) that

$$\begin{aligned}
&\frac{1}{N} I\left(G_0^N \wedge \mathbf{F}'', \left\{K_j \oplus \hat{F}_j\right\}_{m_0 < j \leq m}\right) \\
&\leq \frac{1}{N} I\left(G_0^N \wedge \left\{K_j \oplus \hat{F}_j\right\}_{m_0 < j \leq m} \mid \mathbf{F}''\right) + \epsilon \\
&\leq \sum_{j=m_0+1}^{m} \frac{1}{N}\left[\log \|K_j \oplus \hat{F}_j\| H\left(K_j \oplus \hat{F}_j \mid \mathbf{F}'', \left\{K_i \oplus \hat{F}_i\right\}_{m_0 < i \leq j-1}, G_0^N\right)\right] + \epsilon \\
&\leq \sum_{j=m_0+1}^{m} \frac{1}{N}\left[H\left(K_j\right) - H\left(K_j \oplus \hat{F}_j \mid \mathbf{F}'', \left\{K_i \oplus \hat{F}_i\right\}_{m_0 < i \leq j-1}, G_0^N\right)\right] + 2\epsilon \\
&= \sum_{j=m_0+1}^{m} \frac{1}{N}\left[H\left(K_j\right) - H\left(K_j \mid \mathbf{F}'', \left\{K_i \oplus \hat{F}_i\right\}_{m_0 < i \leq j-1}, G_0^N\right)\right] + 2\epsilon \qquad (4.22) \\
&\leq 3m\epsilon,
\end{aligned}$$

where the third inequality above uses (4.20) and the last inequality follows from (4.21). The equality in (4.22) follows from the fact that $\hat{F}_j = \hat{F}_j\left(G_j^N\right)$ is a function of G_0^N, since G_j is a function of G_0. We note that this is the only place in the proof where the functional relation between G_j and G_0 is used.

Thus, the communication $\left(\mathbf{F}'', K_j \oplus \hat{F}_j, m_0 < j \leq m\right)$ constitutes the required secure computing protocol for $g_{\mathcal{M}}$. It remains to show the existence of $\mathbf{F}''$ and K_j, $m_0 < j \leq m$ that satisfy (4.19)-(4.21).

Specifically, when (4.9) holds for $i = 1$, we have from the definition of $R_1^* \left(g_{\mathcal{M}}\right)$ in (4.11) that for all $0 < \epsilon \leq \epsilon_0$ (ϵ_0 to be specified later), there exists $n \geq 1$ and interactive communication $\mathbf{F} = \mathbf{F}\left(X_{\mathcal{M}}^n\right)$ such that

$$\frac{1}{n} I \left(G_0^n \wedge \mathbf{F}\right) < \epsilon, \tag{4.23}$$

and

$$R_{\mathbf{F}}^{(1)} \leq R_1^* \left(g_{\mathcal{M}}\right) + \frac{\epsilon}{2},$$

where $R_{\mathbf{F}}^{(1)}$ is as in (4.6). This further implies that there exist $R_1, ..., R_m$ satisfying (1a) and (1b) (for $\mathbf{F}$) such that

$$\frac{1}{n} H(\mathbf{F}) + \frac{1}{n} \sum_{i=m_0+1}^{m} H\left(G_j^n \mid X_j^n, \mathbf{F}\right) + R_{\mathcal{M}} \leq R_1^* \left(g_{\mathcal{M}}\right) + \epsilon. \tag{4.24}$$

Choosing

$$\epsilon_0 < H\left(X_{\mathcal{M}} \mid G_0\right) - R_1^* \left(g_{\mathcal{M}}\right) - \delta,$$

for some $\delta < H\left(X_{\mathcal{M}} \mid G_0\right) - R_1^* \left(g_{\mathcal{M}}\right)$, we get from (4.23) and (4.24) upon simplification:

$$\frac{1}{n} \sum_{i=m_0+1}^{m} H\left(G_j^n \mid X_j^n, \mathbf{F}\right) + R_{\mathcal{M}} + \delta < \frac{1}{n} H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right). \tag{4.25}$$

Next, for $k \geq 1$, denote by $\mathbf{F}^k = \left(\mathbf{F}_1, ..., \mathbf{F}_k\right)$ the i.i.d. rvs $\mathbf{F}_i = \mathbf{F}\left(X_{\mathcal{M},n(i-1)+1}, ..., X_{\mathcal{M},ni}\right)$, $1 \leq i \leq k$. Further, let $N = nk$. In Appendix D, we follow the approach in the proof of Theorem 3.4 and use (4.25) to show that for sufficiently large k there exists an interactive communication $\mathbf{F}' = \mathbf{f}'\left(X_{\mathcal{M}}^{nk}\right)$ of overall rate $R_{\mathcal{M}} + \delta/2$ that satisfies the following:

$$X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from } \left(X_i^N, \mathbf{F}^k, \mathbf{F}'\right) \text{ for } 1 \leq i \leq m_0,$$
$$\text{and from } \left(X_i^N, \mathbf{F}^k, G_0^N, \mathbf{F}'\right) \text{ for } m_0 < i \leq m, \tag{4.26}$$

and further,

$$\frac{1}{N} I\left(G_0^N, \mathbf{F}^k \wedge \mathbf{F}'\right) < \epsilon. \tag{4.27}$$

The proposed communication $\mathbf{F}''$ comprises $\mathbf{F}', \mathbf{F}^k$, and condition (4.19) follows from (4.23) and (4.27). Finally, we show the existence of $\hat{F}_j$ and K_j, $m_0 < j \leq m$, as above. From the Slepian-Wolf theorem [60], there exist rvs $\hat{F}_j = \hat{F}_j\left(G_j^N\right)$ of rates

$$R'_j \leq \frac{1}{N} H\left(G_j^N \mid X_j^N, \mathbf{F}^k\right) + \frac{\delta}{2m}, \tag{4.28}$$

such that G_j^N is ϵ-recoverable from $\left(X_j^N, \mathbf{F}^k, \hat{F}_j\right)$, $m_0 < j \leq m$, for k sufficiently large. Suppose the rvs $K_{m_0+1}, K_{m_0+2}, ..., K_j$ of rates $R'_{m_0+1}, R'_{m_0+2}, ..., R'_j$, respectively, satisfy (4.20) and (4.21) for some $j \leq m-1$. Denote by $\mathbf{F}'(j)$ the communication $\left(\mathbf{F}', K_i \oplus \hat{F}_i, m_0 < i \leq j\right)$ of rate $R^{(j)}$ that satisfies

$$R^{(j)} \leq R_{\mathcal{M}} + \frac{1}{N} \sum_{i=m_0+1}^{j} H\left(G_i^N \mid X_i^N, \mathbf{F}^k\right) + \delta \tag{4.29}$$

We have from (4.25)-(4.29) that

$$R'_{j+1} < \frac{1}{N} H\left(X_{\mathcal{M}}^N \mid G_0^N, \mathbf{F}^k\right) - R^{(j)}. \tag{4.30}$$

Heuristically, since $X_{\mathcal{M}}^N$ is recoverable from $\left(X_{j+1}^N, \mathbf{F}^k, \mathbf{F}'\right)$, (4.30) gives

$$
\begin{aligned}
&\frac{1}{N} H\left(X_{j+1}^N \mid G_0^N, \mathbf{F}^k, \mathbf{F}'(j)\right) \\
&\approx \frac{1}{N} H\left(X_{\mathcal{M}}^N \mid G_0^N, \mathbf{F}^k\right) - \frac{1}{N} H\left(\mathbf{F}'(j) \mid G_0^N, \mathbf{F}^k\right) \\
&\geq \frac{1}{N} H\left(X_{\mathcal{M}}^N \mid G_0^N, \mathbf{F}^k\right) - R^{(j)} \\
&> R'_{j+1}.
\end{aligned}
$$

Thus, a randomly chosen mapping $K_{j+1} = K_{j+1}\left(X_{j+1}^N\right)$ of rate R'_{j+1} is almost jointly-independent of $G_0^N, \mathbf{F}^k, \mathbf{F}'(j)$ (see [16]). This argument is made rigorous using a version of the balanced coloring lemma. Specifically, in Lemma 2.7, set $U = X_{\mathcal{M}}^N$, $U' = X_{j+1}^N$, $V = G_0^N, \mathbf{F}^k$, $h = \mathbf{F}'(j)$, and

$$
\mathcal{U}_0 = \left\{ x_{\mathcal{M}}^N \in \mathcal{X}_{\mathcal{M}}^N : x_{\mathcal{M}}^N = \psi_{j+1}\left(x_{j+1}^N, f'\left(x_{\mathcal{M}}^N\right), \mathbf{F}^k, g_0^n\left(x_{\mathcal{M}}^N\right)\right) \right\},
$$

for some mapping ψ_{j+1}, where $f'\left(X_{\mathcal{M}}^N\right) = \mathbf{F}'$ is as in (4.26). By the definition of $\mathbf{F}'$,

$$
\mathrm{P}\left(U \in \mathcal{U}_0\right) \geq 1 - \epsilon,
$$

so that condition (2.22)(i) preceding Lemma 2.7 is met. Condition (2.22)(ii), too, is met from the definition of $\mathcal{U}_0, h$ and V.

Upon choosing

$$
d = \exp\left[k\left(H\left(X_{\mathcal{M}}^n | G_0^n, \mathbf{F}\right) - \frac{n\delta}{2m}\right)\right],
$$

in (2.23), the hypotheses of Lemma 2.7 are satisfied for appropriately chosen λ, and for sufficiently large k. Then, by Lemma 2.7, with

$$
r = \left\lceil \exp\left(NR'_{j+1}\right) \right\rceil, \quad r' = \left\lceil \exp\left(NR^{(j)}\right) \right\rceil,
$$

and with K_{j+1} in the role of ϕ, it follows from (2.25) that there exists rv $K_{j+1} = K_{j+1}\left(X_{j+1}^N\right)$ that satisfies (4.20) and (4.21), for k sufficiently large. The proof is completed upon repeating this argument for $m_0 < j < m$. □

Sufficiency of (4.9) for $i = 2$: The secure computing protocol for this case also consists of two stages. In the first stage, as before, the terminals $[1, m_0]$ (g_0-seeking terminals) attain omniscience, using an interactive communication $\mathbf{F}'' = \mathbf{F}''\left(X_{\mathcal{M}}^N\right)$. The second stage, too, is similar to the previous case and involves one of the omniscience-attaining terminals in $[1, m_0]$ transmitting communication $\hat{F}_j = \hat{F}_j\left(G_j^N\right)$ to the terminals j, for $m_0 < j \leq m$. However, the encryption-based scheme of the previous case is not applicable here; in particular, (4.22) no longer holds. Instead, the communication $\hat{F}_j$ now consists of the Slepian-Wolf codewords for G_j^N given X_j^N, and previous communication $\mathbf{F}''$. We show below that if (4.9) holds, then there exist communication $\mathbf{F}''$ and $\hat{F}_j$, $m_0 < j \leq m$, of appropriate rate such that the following holds:

$$\frac{1}{N} I\left(G_0^N \wedge \mathbf{F}'', \hat{F}_{m_0+1}, ..., \hat{F}_m\right) < \epsilon,$$

for sufficiently large N.

Specifically, when (4.9) holds for $i = 2$, using similar manipulations as in the previous case we get that for all $0 < \epsilon < \epsilon_0$, there exist interactive communication $\mathbf{F} = \mathbf{F}\left(X_{\mathcal{M}}^n\right)$, and rates $R_1, ..., R_m, R'_{m_0+1}, ..., R'_m$ satisfying (2a)-(2c) (for $\mathbf{F}$) such that

$$\frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right) < \frac{\epsilon}{2},$$

and

$$R_{\mathcal{M}} + R'_{[m_0+1,m]} + \delta < \frac{1}{n} H\left(X^n_{\mathcal{M}} \mid G^n_0, \mathbf{F}\right), \tag{4.31}$$

with $\delta < H\left(X_{\mathcal{M}} \mid G_0\right) - R^*_2\left(g_{\mathcal{M}}\right) - \epsilon_0$; (4.31) replaces (4.25) in the previous case.

Next, for $N = nk$ consider $2m - m_0$ correlated sources X^N_j, $1 \leq j \leq m$, and G^N_j, $m_0 < j \leq m$. Since $R_1, ..., R_m, R'_{m_0+1}, ..., R'_m$ satisfy (2a)-(2c), random mappings $F'_j = F'_j\left(X^N_j\right)$ of rates R_j, $1 \leq j \leq m$, and $F'_{j+m-m_0} = F'_{j+m-m_0}\left(G^N_j\right)$ of rates R'_j, $m_0 < j \leq m$ satisfy the following with high probability, for k sufficiently large (see [18, Lemma 13.13 and Theorem 13.14]):

(i) for $1 \leq i \leq m$, $X^{nk}_{\mathcal{M}}$ is ϵ-recoverable from $\left(F'_1, ..., F'_m, \mathbf{F}^k, X^{nk}_i\right)$;

(ii) for $m_0 < j \leq m$, G^{nk}_j is ϵ-recoverable from $\left(F'_{j+m-m_0}, \mathbf{F}^k, X^{nk}_j\right)$;

(iii) for $m_0 < j \leq m$, $X^{nk}_{\mathcal{M}}$ is ϵ-recoverable from $\left(\mathbf{F}', \mathbf{F}^k, X^{nk}_j, G^{nk}_0\right)$ and from $\left(\mathbf{F}', \mathbf{F}^k, G^{nk}_j, G^{nk}_0\right)$,

where $\mathbf{F}^k = \left(\mathbf{F}_1, ..., \mathbf{F}_k\right)$ are i.i.d. rvs $\mathbf{F}_i = \mathbf{F}\left(X_{\mathcal{M},n(i-1)+1}, ..., X_{\mathcal{M},ni}\right)$, $1 \leq i \leq k$. It follows from (4.31) in a manner similar to the proof in Appendix D that there exist communication F'_j, $1 \leq j \leq 2m - m_0$ as above such that

$$\frac{1}{nk} I\left(G^{nk}_0 \wedge \mathbf{F}', \mathbf{F}^k\right) < \epsilon,$$

for sufficiently large k.

The first stage of the protocol entails transmission of $\mathbf{F}^k$, followed by the transmission of $F'_1, ..., F'_m$, i.e., $\mathbf{F}'' = \left(\mathbf{F}^k, F'_1, ..., F'_m\right)$. The second stage of communication

$\hat{F}_j$ is given by F'_{j+m-m_0}, for $m_0 < j \leq m$. □

Sufficiency of (4.9) for $i = 3$: Using the definition of $R_3^*(g_{\mathcal{M}})$ and the manipulations above, the sufficiency condition (4.9) implies that for all $0 < \epsilon < \epsilon_0$, there exist interactive communication $\mathbf{F} = \mathbf{F}(X_{\mathcal{M}}^n)$, and rates $R_1, ..., R_m$ satisfying (3a), (3b) (for $\mathbf{F}$) such that

$$\frac{1}{n} I(G_0^n \wedge \mathbf{F}) < \frac{\epsilon}{2},$$

and

$$R_{\mathcal{M}} + \delta < \frac{1}{n} H(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}), \tag{4.32}$$

for $\delta < H(X_{\mathcal{M}} \mid G_0) - R_3^*(g_{\mathcal{M}}) - \epsilon_0$. Denoting by $\mathbf{F}^k = (\mathbf{F}_1, ..., \mathbf{F}_k)$ the i.i.d. rvs $\mathbf{F}_i = \mathbf{F}\left(X_{n(i-1)+1}^{ni}\right)$, $1 \leq i \leq k$, it follows from (3a) and (3b) that for $N = nk$ the random mappings $F'_i = F'_i\left(X_i^{nk}\right)$ of rates R_i, $1 \leq i \leq m$, satisfy the following with high probability, for k sufficiently large (see [18, Lemma 13.13 and Theorem 13.14]):

(i) for $i \in \mathcal{M}$, $X_{\mathcal{M}_i}^{nk}$ is ϵ-recoverable from $\left(\mathbf{F}', \mathbf{F}^k, X_i^{nk}\right)$;

(ii) for $i \in \mathcal{M}$, $X_{\mathcal{M}}^{nk}$ is ϵ-recoverable from $\left(\mathbf{F}', \mathbf{F}^k, X_i^{nk}, G_0^{nk}\right)$.

From (4.32), proceeding along the lines of the arguments in Appendix D, it follows that there exist F'_i, $i \in \mathcal{M}$, as above such that

$$\frac{1}{nk} I\left(G_0^{nk} \wedge \mathbf{F}', \mathbf{F}^k\right) < \epsilon,$$

for sufficiently large k. The interactive communication $\left(\mathbf{F}', \mathbf{F}^k\right)$ constitutes the protocol for securely computing $g_{\mathcal{M}}$, where $g_i(X_{\mathcal{M}}) = X_{\mathcal{M}_i}, i \in \mathcal{M}$. □

4.5 Proof of necessity

Necessity of (4.10) for $i = 1$: If functions $g_{\mathcal{M}}$ are securely computable then there exists an interactive communication $\mathbf{F}$ such that G_i^n is ϵ_n-recoverable from $(X_i^n, \mathbf{F})$, $i \in \mathcal{M}$, and

$$\frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right) < \epsilon_n, \tag{4.33}$$

where $\epsilon_n \to 0$ as $n \to \infty$. It follows from the Fano's inequality that[3]

$$\frac{1}{n} H\left(G_i^n \mid X_i^n, \mathbf{F}\right) < c_1 \epsilon_n, \qquad i \in \mathcal{M}. \tag{4.34}$$

Using an approach similar to that in [20], we have from (4.33):

$$\begin{aligned}
\frac{1}{n} H\left(X_{\mathcal{M}}^n\right) &= \frac{1}{n} H\left(G_0^n, \mathbf{F}\right) + \frac{1}{n} H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right) \\
&\geq \frac{1}{n} H\left(G_0^n\right) + \frac{1}{n} H\left(\mathbf{F}\right) + \frac{1}{n} H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right) - \epsilon_n, \qquad (4.35) \\
&= \frac{1}{n} H\left(G_0^n\right) + \frac{1}{n} H\left(\mathbf{F}\right) + \frac{1}{n} \sum_{i=1}^{m} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) - \epsilon_n. \qquad (4.36)
\end{aligned}$$

Next, for $\mathcal{L} \subsetneq \mathcal{M}$, with $[1, m_0] \nsubseteq \mathcal{L}$, we have

$$\begin{aligned}
&\frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M} \setminus \mathcal{L}}^n, \mathbf{F}\right) \\
&= \frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M} \setminus \mathcal{L}}^n, G_0^n, \mathbf{F}\right) + \frac{1}{n} H\left(G_0^n \mid X_{\mathcal{M} \setminus \mathcal{L}}^n, \mathbf{F}\right) \\
&\leq \frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M} \setminus \mathcal{L}}^n, G_0^n, \mathbf{F}\right) + c_1 \epsilon_n,
\end{aligned}$$

where the last step follows from (4.34) and the assumption that $g_i = g_0$ for $i \in$

[3] The constants c_1, c_2, c_3, c_4 depend only on $\log \|\mathcal{X}_{\mathcal{M}}\|$, m, m_0 (and not on n).

$[1, m_0]$. Continuing with the inequality above, we get

$$\begin{aligned}
&\frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, \mathbf{F}\right) \\
&\leq \frac{1}{n} \sum_{i \in \mathcal{L}} \left[H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1 \epsilon_n\right], \qquad (4.37)
\end{aligned}$$

Letting

$$R_i = \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1 \epsilon_n, \quad i \in \mathcal{M},$$

by (4.37) $R_1, ..., R_m$ satisfy (1a) and (1b) for $\mathbf{F}$, whereby it follows from (4.34) and (4.36) that

$$\begin{aligned}
&H\left(X_{\mathcal{M}} \mid G_0\right) \\
&\geq \frac{1}{n} H(\mathbf{F}) + \frac{1}{n} \sum_{i=m_0+1}^{m} H\left(G_i^n \mid X_i^n, \mathbf{F}\right) + R_{\mathcal{M}} - c_2 \epsilon_n \\
&\geq R_{\mathbf{F}}^{(1)} - c_2 \epsilon_n,
\end{aligned}$$

where $\mathbf{F}$ satisfies (4.33). Taking the limit $n \to \infty$, and using the definition of $R_1^*\left(g_{\mathcal{M}}\right)$ we get $H\left(X_{\mathcal{M}} \mid G_0\right) \geq R_1^*\left(g_{\mathcal{M}}\right)$. □

Necessity of (4.10) for $i = 2$: If $g_{\mathcal{M}}$ are securely computable, the approach above implies that there exists an interactive communication $\mathbf{F}$ satisfying (4.33) and (4.34) such that, with

$$R_i = \begin{cases} \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1 \epsilon_n, & 1 \leq i \leq m_0, \\ \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_{[m_0+1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1 \epsilon_n, & m_0 < i \leq m, \end{cases}$$

$$R_j' = c_1 \epsilon_n, \quad m_0 < j \leq m,$$

we have by (4.35),

$$\begin{aligned}
H\left(X_{\mathcal{M}} \mid G_0\right) &\geq \frac{1}{n}H(\mathbf{F}) + \frac{1}{n}H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right) - \epsilon_n \\
&\geq \frac{1}{n}H(\mathbf{F}) + \frac{1}{n}\sum_{i=1}^{m_0} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) \\
&\quad + \frac{1}{n}\sum_{i=m_0+1}^{m} H\left(X_i^n \mid X_{[1,i-1]}^n, G_{[m_0+1,i-1]}^n, G_0^n, \mathbf{F}\right) - \epsilon_n \\
&\geq \frac{1}{n}H(\mathbf{F}) + R_{\mathcal{M}} + R'_{[m_0+1,m]} - c_3\epsilon_n. \qquad (4.38)
\end{aligned}$$

Furthermore, (4.34) and the assumption $g_i = g_0$, $1 \leq i \leq m_0$, yield for $[1, m_0] \nsubseteq \mathcal{L} \subsetneq \mathcal{M}$ that

$$\begin{aligned}
\frac{1}{n}H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\backslash\mathcal{L}}^n, \mathbf{F}\right) &\leq \frac{1}{n}H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\backslash\mathcal{L}}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n \\
&\leq \sum_{i\in\mathcal{L},\, i\leq m_0} \left[\frac{1}{n}H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n\right] + \\
&\quad \sum_{i\in\mathcal{L},\, i> m_0} \left[\frac{1}{n}H\left(X_i^n \mid X_{[1,i-1]}^n, G_{[m_0+1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n\right] \\
&= R_{\mathcal{L}}, \qquad (4.39)
\end{aligned}$$

and similarly, for $[1, m_0] \subseteq \mathcal{L} \subseteq \mathcal{M}$, $\mathcal{L}' \subseteq [m_0+1, m]$, with either $\mathcal{L} \neq \mathcal{M}$ or $\mathcal{L}' \neq [m_0+1, m]$ that

$$\begin{aligned}
\frac{1}{n}H\left(G_{\mathcal{L}'}^n, X_{\mathcal{L}}^n | G_{[m_0+1,m]\backslash\mathcal{L}'}^n, X_{\mathcal{M}\backslash\mathcal{L}}^n, G_0^n, \mathbf{F}\right) &= \frac{1}{n}H\left(X_{\mathcal{L}}^n | G_{[m_0+1,m]\backslash\mathcal{L}'}^n, X_{\mathcal{M}\backslash\mathcal{L}}^n, G_0^n, \mathbf{F}\right) \\
&\leq \frac{1}{n}H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\backslash\mathcal{L}}^n, G_0^n, \mathbf{F}\right) \\
&\leq R_{\mathcal{L}} + R'_{\mathcal{L}'}, \qquad (4.40)
\end{aligned}$$

Therefore, (4.39), (4.34) and (4.40) imply that $R_1, ..., R_m$, $R'_{m_0}, ..., R'_m$ satisfy (2a)-(2c) for $\mathbf{F}$, which along with (4.38) yields

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq R_{\mathbf{F}}^{(2)} - c_3\epsilon_n,$$

where $R_{\mathbf{F}}^{(2)}$ is as in (4.7), and $\mathbf{F}$ satisfies (4.33), which completes the proof of necessity (4.10) for $i = 2$ upon taking the limit $n \to \infty$. □

Necessity of (4.10) for $i = 3$: If the functions $g_{\mathcal{M}}$ in (4.5) are securely computable then, as above, there exists an interactive communication $\mathbf{F}$ that satisfies (4.33) and (4.34). Defining

$$R_i = \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n, \quad i \in \mathcal{M},$$

similar manipulations as above yield

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq \frac{1}{n} H(\mathbf{F}) + R_{\mathcal{M}} - c_4\epsilon_n. \tag{4.41}$$

Further, from (4.34) we get that $R_1, ..., R_m$ satisfy (3a) and (3b) for $\mathbf{F}$. It follows from (4.41) that

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq R_{\mathbf{F}}^{(3)} - c_4\epsilon_n,$$

where $R_{\mathbf{F}}^{(2)}$ is as in (4.8), and $\mathbf{F}$ satisfies (4.33), which completes the proof of necessity (4.10) for $i = 3$ as above. □

4.6 Discussion: Alternative necessary conditions for secure computability

The necessary condition (4.10) for secure computing given in section 4.3 is in terms of quantities $R_{\mathbf{F}}^{(i)}$, $i = 1, 2, 3$, defined in (4.6), (4.7), (4.8), respectively. As remarked before, for $i = 1, 3$, the quantity $\inf_{\mathbf{F}} R_{\mathbf{F}}^{(i)}$ is the infimum over the rates of interactive communication that satisfy conditions (P1) and (P2). However, this is not true for

$i = 2$. Furthermore, although $i = 1$ is special case of $i = 2$, it is not clear if the necessary condition (4.10) for $i = 2$ reduces to that for $i = 1$ upon imposing the restriction in (4.4). In this section, we shed some light on this baffling observation.

First, consider the functions $g_{\mathcal{M}}$ in (4.1). For this choice of functions, denoting by R_0^* the minimum rate of interactive communication that satisfies (P1) and (P2), the results in [68] imply that (4.2) constitutes a necessary condition for secure computability, with $R^* = R_0^*$.

Next, consider an augmented model obtained by introducing a new terminal $m+1$ that observes rv $X_{m+1} = \tilde{g}\left(X_{\mathcal{M}}\right)$ and seeks to compute $g_{m+1} = \emptyset$. Further, the terminal does not communicate, i.e., observation X_{m+1}^n is available only for decoding. Clearly, secure computability in the original model implies secure computability in the new model. It follows from the approach of [68] that for the new model also, (4.2) constitutes a necessary condition for secure computability, with R^* now being the minimum rate of interactive communication that satisfies (P1) and (P2) when terminal $m+1$ does not communicate; this R^* is given by

$$\max\{H\left(X_{\mathcal{M}} \mid \tilde{g}(X_{\mathcal{M}}), G_0\right), R_0^*\}.$$

Note that the new necessary condition (4.2) is

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq R_0^* = \max\{H\left(X_{\mathcal{M}} \mid \tilde{g}(X_{\mathcal{M}}), G_0\right), R_0^*\},$$

which is, surprisingly, same as the original condition

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq R_0^*.$$

Our necessary condition (4.10) for $i = 2$ is based on a similar augmentation that entails introduction of $m-m_0$ new terminals observing $g_{m_0+1}\left(X_{\mathcal{M}}\right),...,g_m\left(X_{\mathcal{M}}\right)$ (to be used only for decoding). Now, however, this modification may result in a different necessary condition.

CHAPTER 5

Common Randomness and Minimal Communication for Optimum Rate Secret Keys

5.1 Synopsis

We focus on the generation of a secret key of maximum rate by a pair of terminals observing correlated data and with the means to communicate over a noiseless public communication channel. Our main result establishes a structural equivalence between the generation of a maximum rate secret key and the generation of a common randomness that renders the observations of the two terminals conditionally independent. The minimum rate of such common randomness, termed interactive common information, is related to Wyner's notion of common information, and serves to characterize the minimum rate of interactive public communication required to generate an optimum rate secret key. This characterization yields a single-letter expression for the aforementioned communication rate when the number of rounds of interaction are bounded. An application of our results shows that interaction does not reduce this rate for binary symmetric sources. Further, we provide an example for which interaction does reduce the minimum rate of communication.

The definition of interactive common information and the heuristics underlying our approach are given in Section 5.2. Our main results are provided in Section 5.3, followed by illustrative examples in the subsequent section. Section 5.5 shows an invariance property of interactive common information. A discussion of our results and possible extensions is given in the final section. The results of this chapter were reported in [61, 63].

5.2 Interactive common information

The results of this chapter are for the case of two terminals; for convenience, we simplify our notations for the interactive communication **F**. Terminals 1 and 2 communicate interactively, with, say, terminal 1 transmitting first. Each terminal then communicates alternately for r rounds. Specifically, an *r-interactive communication* $\mathbf{f} = (f_1, f_2, ..., f_r)$ is a sequence of finite-valued mappings with

$$f_{2i+1} : \mathcal{X}_1^n \times \mathcal{F}^{2i} \rightarrow \mathcal{F}_{2i+1}, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor,$$

$$f_{2i} : \mathcal{X}_2^n \times \mathcal{F}^{2i-1} \rightarrow \mathcal{F}_{2i}, \quad 1 \leq i \leq \lfloor r/2 \rfloor,$$

where $\{\mathcal{F}_i\}_{i=1}^r$ are finite sets and $\mathcal{F}_0 = \emptyset$. As in Chapter 2, this set-up subsumes protocols where terminal 2 initiates the communication upon choosing $f_1 =$ constant. Let $\mathbf{F} = \mathbf{f}\left(X_1^n, X_2^n\right)$ describe collectively the corresponding rv. The rate of this communication is given by

$$\frac{1}{n} \log \|\mathbf{F}\|.$$

We assume that the communication from each terminal is a (deterministic) function of its knowledge. In particular, randomization is not allowed. This is not a limiting assumption; see Section 5.6.1.

We define a new notion of CI, termed interactive CI, which will be a key concept for this chapter. Recall from Chapter 2 that Wyner's CI is defined as the minimum rate of rvs $L = L^{(n)}(X_1^n, X_2^n)$ such that for all $\epsilon > 0$, the following holds for n sufficiently large:

$$\frac{1}{n} I\left(X_1^n \wedge X_2^n \mid L\right) \leq \epsilon. \tag{5.1}$$

Interactive CI is defined by restricting the rvs L to be CR.

Definition 5.1. An achievable *r-interactive* CI rate is defined in a manner analogous to the achievable CI rate, but with the restriction that the rvs L in (5.1) be ϵ-CR, i.e., $L = (J, \mathbf{F})$, where $\mathbf{F}$ is an r-interactive communication and J is ϵ-recoverable from $\mathbf{F}$. The infimum of all achievable r-interactive CI rates, denoted $CI_i^r(X_1; X_2)$, is called the r-interactive CI of the rvs X_1 and X_2. By definition, the nonnegative sequence $\{CI_i^r(X_1; X_2)\}_{r=1}^{\infty}$ is nonincreasing in r and is bounded below by $CI_W(X_1 \wedge X_2)$. Define

$$CI_i(X_1 \wedge X_2) = \lim_{r \to \infty} CI_i^r(X_1; X_2).$$

Then $CI_i(X_1 \wedge X_2) \geq CI_W(X_1 \wedge X_2) \geq 0$. Note that $CI_i^r(X_1; X_2)$ may not be symmetric in X_1 and X_2 since the communication is initiated at terminal 1. However, since

$$CI_i^{r+1}(X_1; X_2) \leq CI_i^r(X_2; X_1) \leq CI_i^{r-1}(X_1; X_2),$$

clearly,

$$CI_i(X_1 \wedge X_2) = \lim_{r\to\infty} CI_i^r(X_1; X_2) = \lim_{r\to\infty} CI_i^r(X_2; X_1) = CI_i(X_2 \wedge X_1). \qquad (5.2)$$

Further, for all $0 < \epsilon < 1$, $J = X_1^n$ is ϵ-recoverable from X_2^n and a communication (of a Slepian-Wolf codeword) $F = F\left(X_1^n\right)$, and $L = (J, F)$ satisfies (5.1). Hence, $CI_i(X_1 \wedge X_2) \leq H(X_1)$; similarly, $CI_i(X_1 \wedge X_2) \leq H(X_2)$. To summarize, we have

$$0 \leq CI_W(X_1 \wedge X_2) \leq CI_i(X_1 \wedge X_2) \leq \min\{H(X_1); H(X_2)\}, \qquad (5.3)$$

where the first and the last inequalities can be strict. In Section 5.4.1 we show that the second inequality is strict for BSS X_1, X_2.

The r-interactive CI plays a pivotal role in optimum rate SK generation. Loosely speaking, our main result asserts the following. *A CR that satisfies (5.1) can be used to generate an optimum rate SK and conversely, an optimum rate SK yields a CR satisfying (5.1). In fact, such a CR of rate R can be recovered from an interactive communication of rate $R - C$, where C is the SK capacity for X_1 and X_2. Therefore, to find the minimum rate of interactive communication needed to generate an optimum rate SK, it is sufficient to characterize $CI_i(X_1 \wedge X_2)$.*

5.3 Formulation and main results

Definition 5.2. A number $R' \geq 0$ is an achievable r-interactive communication rate for CI_i^r if, for all $0 < \epsilon < 1$, there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate $(1/n)\log\|\mathbf{F}\| \leq R' + \epsilon$, and an ϵ-CR J recoverable from $\mathbf{F}$, with $L = (J, \mathbf{F})$ satisfying (5.1). Let R_{CI}^r denote the infimum of all achievable

r-interactive communication rates for CI_i^r. Similarly, $R'' \geq 0$ is an achievable r-interactive communication rate for SK capacity if, for all $0 < \epsilon < 1$, there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate $(1/n) \log \|\mathbf{F}\| \leq R'' + \epsilon$, and an ϵ-SK K, recoverable from $\mathbf{F}$, of rate $(1/n)H(K) \geq I(X_1 \wedge X_2) - \epsilon$; R_{SK}^r denotes the infimum of all achievable r-interactive communication rates for SK capacity. Note that by their definitions, both R_{CI}^r and R_{SK}^r are nonincreasing with increasing r, and are bounded below by zero. Define

$$R_{CI} = \lim_{r \to \infty} R_{CI}^r, \quad R_{SK} = \lim_{r \to \infty} R_{SK}^r.$$

Although $R_{CI}^r(X_1; X_2)$ and $R_{SK}^r(X_1; X_2)$ are not equal to $R_{CI}^r(X_2; X_1)$ and $R_{SK}^r(X_2; X_1)$, respectively, the quantities R_{CI} and R_{SK} are symmetric in X_1 and X_2 using an argument similar to the one leading to (5.2).

Theorem 5.1. *For every $r \geq 1$,*

$$R_{SK}^r = R_{CI}^r = CI_i^r(X_1; X_2) - I(X_1 \wedge X_2). \tag{5.4}$$

Corollary 5.2. *It holds that*

$$R_{SK} = R_{CI} = CI_i(X_1 \wedge X_2) - I(X_1 \wedge X_2). \tag{5.5}$$

Remark. The relation (5.5) can be interpreted as follows. Any CR J recoverable from (interactive communication) $\mathbf{F}$, with $L = (J, \mathbf{F})$ satisfying (5.1), can be decomposed into two mutually independent parts: An SK K of maximum rate and the interactive communication $\mathbf{F}$. It follows upon rewriting (5.5) as $CI_i(X_1 \wedge X_2) = I(X_1 \wedge X_2) + R_{CI}$ that the communication $\mathbf{F}$ is (approximately) of rate R_{CI}. Furthermore, R_{CI} is the same as R_{SK}.

A computable characterization of the operational term $CI_i(X_1 \wedge X_2)$ is not known. However, the next result gives a single-letter characterization of $CI_i^r(X_1; X_2)$.

Theorem 5.3. *Given rvs X_1, X_2 and $r \geq 1$, we have*

$$CI_i^r(X_1; X_2) = \min_{U_1, \ldots, U_r} I(X_1, X_2 \wedge U_1, ..., U_r), \tag{5.6}$$

where the minimum is taken over rvs $U_1, ..., U_r$ taking values in finite sets $\mathcal{U}_1, ..., \mathcal{U}_r$, respectively, that satisfy the following conditions

$$\begin{aligned}
&(P1)\ \ U_{2i+1} \mathrel{-\!\!\circ\!\!-} X_1, U^{2i} \mathrel{-\!\!\circ\!\!-} X_2, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\
&\qquad U_{2i} \mathrel{-\!\!\circ\!\!-} X_2, U^{2i-1} \mathrel{-\!\!\circ\!\!-} X_1, \quad 1 \leq i \leq \lfloor r/2 \rfloor, \\
&(P2)\ \ X_1 \mathrel{-\!\!\circ\!\!-} U^r \mathrel{-\!\!\circ\!\!-} X_2, \\
&(P3)\ \ |\mathcal{U}_{2i+1}| \leq |\mathcal{X}_1| \prod_{j=1}^{2i} |\mathcal{U}_j| + 1, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\
&\qquad |\mathcal{U}_{2i}| \leq |\mathcal{X}_2| \prod_{j=1}^{2i-1} |\mathcal{U}_j| + 1, \quad 1 \leq i \leq \lfloor r/2 \rfloor,
\end{aligned}$$

with $\mathcal{U}_0 = \emptyset$ and $U_0 =$ constant.

Remark. Note that (5.6) has the same form as the expression for $CI_W(X_1 \wedge X_2)$ in (2.17) with W replaced by $(U_1, ..., U_r)$ satisfying the conditions above.

Before presenting the proof of our main Theorems 5.1 and 5.3, we give pertinent technical results that will constitute central tools for the proofs.

Lemma 5.4. *For an r-interactive communication $\mathbf{F}$, define*

$$\mathbf{F}_i = \mathbf{F}\left(X_{1(n(i-1)+1)}^{ni}, X_{(2n(i-1)+1)}^{ni}\right), \quad 1 \leq i \leq k.$$

Then, for all $k \geq k_0(n, \epsilon, |\mathcal{X}_1|, |\mathcal{X}_2|)$ there exists an r-interactive communication $\mathbf{F}' = \mathbf{F}'\left(X_1^{nk}, X_2^{nk}\right)$ of rate

$$\frac{1}{nk} \log \|\mathbf{F}'\| \leq \frac{1}{n} \left[H\left(\mathbf{F}|X_1^n\right) + H\left(\mathbf{F}|X_2^n\right)\right] + \epsilon, \tag{5.7}$$

such that $\mathbf{F}^k$ is an ϵ-CR recoverable from $\mathbf{F}'$.

Proof. From the Slepian-Wolf theorem [60], there exist mappings $f_1, ..., f_r$ of $F_1^k, ..., F_r^k$, respectively, of rates

$$\begin{aligned}\frac{1}{k} \log \|f_{2i+1}\| &\leq H(F_{2i+1} \mid X_2^n, F_1, ..., F_{2i}) + \frac{n\epsilon}{2r}, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\ \frac{1}{k} \log \|f_{2i}\| &\leq H(F_{2i} \mid X_1^n, F_1, ..., F_{2i-1}) + \frac{n\epsilon}{2r}, \quad 1 \leq i \leq \lfloor r/2 \rfloor,\end{aligned}$$

such that

$$\begin{aligned}&F_{2i+1}^k \text{ is } \frac{\epsilon}{2r}\text{-recoverable from } \left(f_{2i+1}(F_{2i+1}^k), X_2^N, F_1^k, ..., F_{2i}^k\right), \ 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\ &F_{2i}^k \text{ is } \frac{\epsilon}{2r}\text{-recoverable from } \left(f_{2i}(F_{2i}^k), X_1^N, F_1^k, ..., F_{2i-1}^k\right), \ 1 \leq i \leq \lfloor r/2 \rfloor,\end{aligned}$$

for all k sufficiently large. Thus, the communication $\mathbf{F}'$ given by $F_i' = f_i\left(F_i^k\right)$, $1 \leq i \leq r$ constitutes the required communication of rate

$$\frac{1}{nk} \log \|\mathbf{F}'\| \leq \frac{1}{n} \left[H\left(\mathbf{F}|X_1^n\right) + H\left(\mathbf{F}|X_2^n\right)\right] + \epsilon.$$

□

Remark. Lemma 5.4 says that, in essence, for an optimum rate communication $\mathbf{F}$,

$$\frac{1}{n} \log \|\mathbf{F}\| \approx \frac{1}{n} \left[H\left(\mathbf{F}|X_1^n\right) + H\left(\mathbf{F}|X_2^n\right)\right].$$

Lemma 5.5. *(A General Decomposition) For a CR J recoverable from an interactive communication $\mathbf{F}$ we have*

$$\begin{aligned} nI(X_1 \wedge X_2) = I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}\right) + H(J, \mathbf{F}) - H\left(\mathbf{F} \mid X_1^n\right) \\ - H\left(\mathbf{F} \mid X_2^n\right) - H\left(J \mid X_1^n, \mathbf{F}\right) - H\left(J \mid X_2^n, \mathbf{F}\right). \end{aligned} \tag{5.8}$$

Proof. For $T = T\left(X_1^n, X_2^n\right)$ we have,

$$\begin{aligned} & nI(X_1 \wedge X_2) \\ & = H\left(X_1^n, X_2^n\right) - H\left(X_1^n \mid X_2^n\right) - H\left(X_2^n \mid X_1^n\right) \\ & = H\left(X_1^n, X_2^n \mid T\right) - H\left(X_1^n \mid X_2^n, T\right) - H\left(X_2^n \mid X_1^n, T\right) \\ & \qquad + H(T) - H\left(T \mid X_1^n\right) - H\left(T \mid X_2^n\right) \\ & = I\left(X_1^n \wedge X_2^n \mid T\right) + H(T) - H\left(T \mid X_1^n\right) - H\left(T \mid X_2^n\right). \end{aligned}$$

Lemma 5.5 follows upon choosing $T = J, \mathbf{F}$. □

Note that a simplification of (5.8) gives

$$I(X_1 \wedge X_2) \leq \frac{1}{n}\Big[I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}\right) + H(J, \mathbf{F}) - H\left(\mathbf{F} \mid X_1^n\right) - H\left(\mathbf{F} \mid X_2^n\right)\Big]. \tag{5.9}$$

If J is an ϵ-CR recoverable from $\mathbf{F}$, Fano's inequality implies

$$\frac{1}{n}\big[H(J \mid X_1^n, \mathbf{F}) + H(J \mid X_2^n, \mathbf{F})\big] \leq 2\epsilon \log|\mathcal{X}_1||\mathcal{X}_2| + 2h(\epsilon) = \delta(\epsilon), \text{ say,} \tag{5.10}$$

where $\delta(\epsilon) \to 0$ as $\epsilon \to 0$. Combining (5.8) and (5.10) we get

$$I(X_1 \wedge X_2) \geq \frac{1}{n}\Big[I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}\right) + H(J, \mathbf{F}) - H\left(\mathbf{F} \mid X_1^n\right) - H\left(\mathbf{F} \mid X_2^n\right)\Big] - \delta(\epsilon), \tag{5.11}$$

and further, by Lemma 4.3,

$$I(X_1 \wedge X_2) \geq \frac{1}{n}\left[I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}\right) + H(J, \mathbf{F}) - H(\mathbf{F})\right] - \delta(\epsilon). \tag{5.12}$$

Proof of Theorem 5.1.

In this section we give a proof for (5.4). The proof of (5.5) then follows upon taking limit $r \to \infty$ on both sides of (5.4). The proof of (5.4) follows from claims 1-3 below. In particular, the proofs of claims 1-3 establish a structural equivalence between a maximum rate SK and an SK of rate $\approx \frac{1}{n}H(J \mid \mathbf{F})$ extracted from a CR J recoverable from $\mathbf{F}$ such that $L = (J, \mathbf{F})$ satisfies (5.1).

Claim 1: $R_{CI}^r \geq CI_i^r(X_1; X_2) - I(X_1 \wedge X_2)$.

Proof. By the definition of R_{CI}^r, for every $0 < \epsilon < 1$ there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate

$$\frac{1}{n}\log\|\mathbf{F}\| \leq R_{CI}^r + \epsilon, \tag{5.13}$$

and J, an ϵ-CR recoverable from $\mathbf{F}$, such that $L = (J, \mathbf{F})$ satisfies (5.1). It follows upon rearranging the terms in (5.12) that

$$\frac{1}{n}H(J, \mathbf{F}) \leq I(X_1 \wedge X_2) + \frac{1}{n}H(\mathbf{F}) + \delta(\epsilon),$$

which with (5.13) gives

$$\frac{1}{n}H(J, \mathbf{F}) \leq I(X_1 \wedge X_2) + R_{CI}^r + \epsilon + \delta(\epsilon). \tag{5.14}$$

Since $(J, \mathbf{F})$ satisfies

$$\frac{1}{n}I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}\right) \leq \epsilon \leq \epsilon + \delta(\epsilon),$$

the inequality (5.14), along with the fact that $(\epsilon + \delta(\epsilon)) \to 0$ as $\epsilon \to 0$, implies that $I(X_1 \wedge X_2) + R^r_{CI}$ is an achievable r-interactive CI rate; hence, $CI^r_i(X_1; X_2) \leq I(X_1 \wedge X_2) + R^r_{CI}$.

Claim 2: $R^r_{SK} \geq R^r_{CI}$.

Proof. Using the definition of R^r_{SK}, for $0 < \epsilon < 1$ there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate $\frac{1}{n}\log\|\mathbf{F}\| \leq R^r_{SK} + \epsilon$, and an ϵ-SK K recoverable from $\mathbf{F}$ of rate

$$\frac{1}{n}H(K) \geq I(X_1 \wedge X_2) - \epsilon. \tag{5.15}$$

By choosing $J = K$ in (5.12) and rearranging the terms we get,

$$\frac{1}{n}I\left(X_1^n \wedge X_2^n \mid K, \mathbf{F}\right) \leq I(X_1 \wedge X_2) - \frac{1}{n}H(K \mid \mathbf{F}) + \delta(\epsilon).$$

Next, from $(1/n)I(K \wedge \mathbf{F}) < \epsilon$, we have

$$\begin{aligned}\frac{1}{n}I\left(X_1^n \wedge X_2^n \mid K, \mathbf{F}\right) &\leq I(X_1 \wedge X_2) - \frac{1}{n}H(K) + \epsilon + \delta(\epsilon)\\ &\leq 2\epsilon + \delta(\epsilon),\end{aligned}$$

where the last inequality follows from (5.15). Since $(2\epsilon + \delta(\epsilon)) \to 0$ as $\epsilon \to 0$, R^r_{SK} is an achievable r-interactive communication rate for CI^r_i, and thus, $R^r_{SK} \geq R^r_{CI}$.

Claim 3: $R^r_{SK} \leq CI^r_i(X_1; X_2) - I(X_1 \wedge X_2)$.

Proof. For $0 < \epsilon < 1$, let J be an ϵ-CR recoverable from an r-interactive communication $\mathbf{F}$, with

$$\frac{1}{n}H(J, \mathbf{F}) \leq CI^r_i(X_1; X_2) + \epsilon, \tag{5.16}$$

such that $L = (J, \mathbf{F})$ satisfies (5.1), and so, by (5.9),

$$\begin{aligned}\frac{1}{n}\left[H(\mathbf{F} \mid X_1^n) + H(\mathbf{F} \mid X_2^n)\right] &\leq \frac{1}{n}H(J, \mathbf{F}) - I(X_1 \wedge X_2) + \epsilon \\ &\leq CI_i^r(X_1; X_2) - I(X_1 \wedge X_2) + 2\epsilon. \qquad (5.17)\end{aligned}$$

To prove the assertion in claim 3, we show that for some $N \geq 1$ there exists $\Delta(\epsilon)$-SK $K = K(X_1^N, X_2^N)$ of rate

$$\frac{1}{n}\log \|K\| \geq I(X_1 \wedge X_2) - \Delta(\epsilon)$$

recoverable from an r-interactive communication $\mathbf{F}'' = \mathbf{F}''(X_1^N, X_2^N)$ of rate

$$\frac{1}{N}\log \|\mathbf{F}''\| \leq \frac{1}{n}\left[H(\mathbf{F} \mid X_1^n) + H(\mathbf{F} \mid X_2^n)\right] + \Delta(\epsilon) - 2\epsilon, \qquad (5.18)$$

where $\Delta(\epsilon) \to 0$ as $\epsilon \to 0$. Then (5.18), along with (5.17), would yield

$$\frac{1}{N}\log \|\mathbf{F}''\| \leq CI_i^r(X_1; X_2) - I(X_1 \wedge X_2) + \Delta(\epsilon), \qquad (5.19)$$

so that $CI_i^r(X_1; X_2) - I(X_1 \wedge X_2)$ is an achievable r-interactive communication rate for SK capacity, thereby establishing the claim.

It remains to find K and $\mathbf{F}''$ as above. To that end, let J be recovered as $J_1 = J_1(X_1^n, \mathbf{F})$ and $J_2 = J_2(X_2^n, \mathbf{F})$ by terminals 1 and 2, respectively, i.e.,

$$\mathrm{P}\left(J = J_1 = J_2\right) \geq 1 - \epsilon.$$

Further, for $k \geq 1$, let

$$J_{1i} = J_1\left(X_{1(n(i-1)+1)}^{ni}, \mathbf{F}_i\right), \qquad J_{2i} = J_2\left(X_{(2n(i-1)+1)}^{ni}, \mathbf{F}_i\right), \quad 1 \leq i \leq k,$$

where $\mathbf{F}_i = \mathbf{F}\left(X_{1(n(i-1)+1)}^{ni}, X_{(2n(i-1)+1)}^{ni}\right)$. For odd r, we find an r-interactive communication $\mathbf{F}''$ such that $\left(J_1^k, \mathbf{F}^k\right)$ is a ϵ-CR recoverable from $\mathbf{F}''$, for all k sufficiently large; the the SK K will be chosen to be a function of $\left(J_1^k, \mathbf{F}^k\right)$ of appropriate rate. The proof for even r is similar and is obtained by interchanging the roles of J_1 and J_2. In particular, by Lemma 5.4, for all k sufficiently large there exists an r-interactive communication $\mathbf{F}'$ such that $\mathbf{F}^k$ is ϵ-CR recoverable from $\mathbf{F}'$ of rate given by (5.7). Next, from Fano's inequality

$$\frac{1}{n}\max\{H(J \mid J_1); H(J_1 \mid J_2)\} \leq \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon). \tag{5.20}$$

By the Slepian-Wolf theorem [60] there exists a mapping f of J_1^k of rate

$$\frac{1}{k}\log \|f\| \leq H(J_1 \mid J_2) + n\epsilon, \tag{5.21}$$

such that

$$J_1^k \text{ is } \epsilon\text{-recoverable from } \left(f\left(J_1^k\right), J_2^k\right), \tag{5.22}$$

for all k sufficiently large. It follows from (5.20), (5.21) that

$$\frac{1}{nk}\log \|f\| \leq \epsilon + \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon). \tag{5.23}$$

For $N = nk$, we define the r-interactive communication $\mathbf{F}'' = \mathbf{F}''\left(X_1^N, X_2^N\right)$ as

$$F_i'' = F_i', \quad 1 \leq i \leq r-1,$$

$$F_k'' = F_r', f(J_1^k), \quad i = r,$$

Thus, $\left(J_1^k, \mathbf{F}^k\right)$ is 2ϵ-CR recoverable from $\mathbf{F}''$, where, by (5.7) and (5.23), the rate of communication $\mathbf{F}''$ is bounded by

$$\begin{aligned}&\frac{1}{nk}\log \|\mathbf{F}''\| \\ &\leq \frac{1}{n}\left[H\left(\mathbf{F}|X_1^n\right) + H\left(\mathbf{F}|X_2^n\right)\right] + 2\epsilon + \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon).\end{aligned} \tag{5.24}$$

Finally, to construct the SK $K = K\left(J_1^k, \mathbf{F}^k\right)$, using the corollary of Balanced Coloring Lemma in [20, Lemma B.3], with

$$U = (J_1, \mathbf{F}),\ V = \phi,\ n = k,\ g = \mathbf{F}',$$

we get from (5.24) that there exists a function K of $J_1^k, \mathbf{F}^k$ such that

$$\begin{aligned}\frac{1}{k}\log\|K\| &\geq H(U) - \frac{1}{k}\log\|\mathbf{F}''\| \\ &\geq H(J_1,\mathbf{F}) - H(\mathbf{F} \mid X_1^n) - H(\mathbf{F} \mid X_2^n) - n(2\epsilon + \epsilon\log|\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon)),\end{aligned} \tag{5.25}$$

and

$$I(K \wedge \mathbf{F}') \leq \exp(-ck),$$

where $c > 0$, for all sufficiently large k. We get from (5.25) and (5.9) that the rate of K is bounded below as follows:

$$\frac{1}{nk}\log\|K\| \geq I(X_1 \wedge X_2) - \frac{1}{n}I\left(X_1^n \wedge X_2^n \mid J_1, \mathbf{F}\right) - 2\epsilon - \epsilon\log|\mathcal{X}_1||\mathcal{X}_2| - h(\epsilon). \tag{5.26}$$

Observe that

$$\begin{aligned}I(X_1^n \wedge X_2^n \mid J, \mathbf{F}) &= I(J_1, X_1^n \wedge X_2^n \mid J, \mathbf{F}) \\ &\geq I(X_1^n \wedge X_2^n \mid J, J_1, \mathbf{F}) \\ &\geq I(X_1^n \wedge X_2^n \mid J_1, \mathbf{F}) - H(J \mid J_1),\end{aligned}$$

which along with (5.20), and the fact that $L = (J, \mathbf{F})$ satisfies (5.1), yields

$$\frac{1}{n}I(X_1^n \wedge X_2^n \mid J_1, \mathbf{F}) \leq \epsilon + \epsilon\log|\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon). \tag{5.27}$$

Upon combining (5.26) and (5.27) we get,

$$\frac{1}{nk}\log\|K\| \geq I(X_1 \wedge X_2) - 3\epsilon - 2\epsilon\log|\mathcal{X}_1||\mathcal{X}_2| - 2h(\epsilon).$$

Thus, for $\Delta(\epsilon) = 4\epsilon + 2\epsilon\log|\mathcal{X}_1||\mathcal{X}_2| + 2h(\epsilon)$, K is a $\Delta(\epsilon)$-SK of rate $(1/nk)\log\|K\| \geq I(X_1 \wedge X_2) - \Delta(\epsilon)$, recoverable from r-interactive communication $\mathbf{F}''$, which with (5.24), completes the proof. □

Proof of Theorem 5.3.

Achievability. Consider rvs $U_1, ..., U_r$ satisfying conditions (P1)-(P3) in the statement of Theorem 5.3. It suffices to show for every $0 < \epsilon < 1$, for some $n \geq 1$, there exists an r-interactive communication $\mathbf{F}$, and ϵ-CR J recoverable from $\mathbf{F}$, such that

$$I(X_1, X_2 \wedge U^r) - \epsilon \leq \frac{1}{n}H(J, \mathbf{F}) \leq I(X_1, X_2 \wedge U^r) + \epsilon, \tag{5.28}$$

and

$$\frac{1}{n}H(\mathbf{F}) \leq I(X_1, X_2 \wedge U^r) - I(X_1 \wedge X_2) + \epsilon, \tag{5.29}$$

since from (5.12), (5.28) and (5.29), we have

$$\begin{aligned}\frac{1}{n}I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}\right) &\leq \frac{1}{n}H(\mathbf{F}) - \frac{1}{n}H(J, \mathbf{F}) + I(X_1 \wedge X_2) + \delta(\epsilon)\\ &\leq 2\epsilon + \delta(\epsilon).\end{aligned}$$

We show below that

$$\begin{aligned}I(X_1, X_2 \wedge U^r) - I(X_1 \wedge X_2) = &\sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X_1 \wedge U_{2i+1} \mid X_2, U^{2i}) +\\ &\sum_{i=1}^{\lfloor r/2 \rfloor} I(X_2 \wedge U_{2i} \mid X_1, U^{2i-1}).\end{aligned} \tag{5.30}$$

Thus, the proof will be completed upon showing that there exists an ϵ-CR J, recoverable from $\mathbf{F}$ of rate

$$\frac{1}{n}H(\mathbf{F}) \leq \sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X_1 \wedge U_{2i+1} \mid X_2, U^{2i}) \quad + \sum_{i=1}^{\lfloor r/2 \rfloor} I(X_2 \wedge U_{2i} \mid X_1, U^{2i-1}) + \epsilon, \tag{5.31}$$

such that $(J, \mathbf{F})$ satisfies (5.28). For $r = 2$, such a construction was given by Ahlswede-Csiszár [2, Theorem 4.4]. (In their construction, $\mathbf{F}$ was additionally a function of J.) The extension of their construction to a general r is straightforward, and is relegated to Appendix E.

It remains to prove (5.30). Note that

$$\begin{aligned} I(X_1, X_2 \wedge U^r) - &\sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X_1 \wedge U_{2i+1} \mid X_2, U^{2i}) - \sum_{i=1}^{\lfloor r/2 \rfloor} I(X_2 \wedge U_{2i} \mid X_1, U^{2i-1}) \\ = &\sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X_2 \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=1}^{\lfloor r/2 \rfloor} I(X_1 \wedge U_{2i} \mid U^{2i-1}). \end{aligned} \tag{5.32}$$

Further, from conditions (P1)-(P3) it follows that

$$\begin{aligned} &\sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X_2 \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=1}^{\lfloor r/2 \rfloor} I(X_1 \wedge U_{2i} \mid U^{2i-1}) - I(X_2 \wedge X_1) \\ &= \sum_{i=1}^{\lfloor (r-1)/2 \rfloor} I(X_2 \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=2}^{\lfloor r/2 \rfloor} I(X_1 \wedge U_{2i} \mid U^{2i-1}) \\ &\qquad + I(X_1 \wedge U_2 \mid U_1) + I(X_2 \wedge U_1) - I(X_2 \wedge X_1) \\ &= \sum_{i=1}^{\lfloor (r-1)/2 \rfloor} I(X_2 \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=2}^{\lfloor r/2 \rfloor} I(X_1 \wedge U_{2i} \mid U^{2i-1}) \\ &\qquad + I(X_1 \wedge U_2 \mid U_1) - I(X_1 \wedge X_2 \mid U_1) \\ &= \sum_{i=1}^{\lfloor (r-1)/2 \rfloor} I(X_2 \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=2}^{\lfloor r/2 \rfloor} I(X_1 \wedge U_{2i} \mid U^{2i-1}) - I(X_1 \wedge X_2 \mid U_1, U_2) \\ &= \ldots = -I(X_1 \wedge X_2 \mid U^r) = 0. \end{aligned} \tag{5.33}$$

Combining (5.32) and (5.33) we get (5.30).

Converse. Let $R \geq 0$ be an achievable r-interactive CI rate. Then, for all $0 < \epsilon < 1$, for some $n \geq 1$, there exists an r-interactive communication $\mathbf{F}$, and ϵ-CR J recoverable from $\mathbf{F}$, such that $(1/n)H(J, \mathbf{F}) \leq R + \epsilon$ and $L = (J, \mathbf{F})$ satisfies (5.1). Let J be recovered as $J_1 = J_1(X_1^n, \mathbf{F})$ and $J_2 = J_2(X_2^n, \mathbf{F})$ by terminals 1 and 2, respectively, i.e., $\mathrm{P}\left(J = J_1 = J_2\right) \geq 1 - \epsilon$. Further, let rv T be distributed uniformly over the set $\{1, ..., n\}$. Define rvs U^r as follows:

$$
\begin{aligned}
&U_1 = F_1, X_1^{T-1}, X_{2(T+1)}^n, T,\\
&U_i = F_i, \qquad 2 \leq i < r,\\
&U_r = \begin{cases} (F_r, J_1), & r \text{ odd},\\ \\ (F_r, J_2), & r \text{ even}. \end{cases}
\end{aligned}
$$

We complete the proof for odd r; the proof for even r can be completed similarly. It was shown by Kaspi [39, equations (3.10)-(3.13)] that

$$
\begin{aligned}
&U_{2i+1} -\!\!\circ\!\!- X_{1T}, U^{2i} -\!\!\circ\!\!- X_{2T}, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor,\\
&U_{2i} -\!\!\circ\!\!- X_{2T}, U^{2i-1} -\!\!\circ\!\!- X_{1T}, \quad 1 \leq i \leq \lfloor r/2 \rfloor.
\end{aligned}
$$

Next, note from (5.27) that

$$
\begin{aligned}
\epsilon + \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon) &\geq \frac{1}{n} I(X_1^n \wedge X_2^n \mid J_1, \mathbf{F})\\
&\geq \frac{1}{n} \sum_{i=1}^n I(X_{1i} \wedge X_2^n \mid X_1^{i-1}, J_1, \mathbf{F})\\
&\geq \frac{1}{n} \sum_{i=1}^n I(X_{1i} \wedge X_{2i} \mid X_1^{i-1}, X_{2(i+1)}^n, J_1, \mathbf{F})\\
&= I(X_{1T} \wedge X_{2T} \mid U^r). \qquad (5.34)
\end{aligned}
$$

Similarly, it holds that

$$\epsilon + \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon) \geq I(X_{1T} \wedge X_{2(T+1)}^n \mid X_1^{T-1}, J_1, \mathbf{F}, T). \tag{5.35}$$

The entropy rate of $(J, \mathbf{F})$ is now bounded as

$$\begin{aligned}
&\frac{1}{n} H(J, \mathbf{F}) \\
&\geq \frac{1}{n} H(J_1, \mathbf{F}) - \frac{1}{n} H(J_1 \mid J) \\
&\geq \frac{1}{n} H(J_1, \mathbf{F}) - \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| - h(\epsilon) \\
&= \frac{1}{n} I(X_1^n, X_2^n \wedge J_1, \mathbf{F}) - \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| - h(\epsilon) \\
&= H(X_{1T}, X_{2T}) - \frac{1}{n} H(X_1^n \mid J_1, \mathbf{F}) - \frac{1}{n} H(X_2^n \mid X_1^n, J_1, \mathbf{F}) - \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| - h(\epsilon) \\
&= H(X_{1T}, X_{2T}) - H(X_{1T} \mid X_1^{T-1}, J_1, \mathbf{F}, T) \\
&\qquad - H(X_{2T} \mid X_1^{T-1}, X_{2(T+1)}^n, X_{1T}, X_{1(T+1)}^n, J_1, \mathbf{F}, T) - \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| - h(\epsilon) \\
&\geq I(X_{1T}, X_{2T} \wedge U^r) - \epsilon - 2\epsilon \log |\mathcal{X}_1||\mathcal{X}_2| - 2h(\epsilon),
\end{aligned}$$

where the second inequality follows from Fano's inequality, and the last inequality follows from (5.35). Consequently,

$$R \geq \frac{1}{n} H(J, \mathbf{F}) - \epsilon \geq I(X_{1T}, X_{2T} \wedge U^r) - 2(\epsilon + \epsilon \log |\mathcal{X}_1||\mathcal{X}_2| + h(\epsilon)). \tag{5.36}$$

We now replace the rvs $U_1, ..., U_r$ with those taking values in finites sets $\mathcal{U}_1, ..., \mathcal{U}_r$, respectively, with $\mathcal{U}_1, ..., \mathcal{U}_r$ satisfying the cardinality bounds in condition (iii). Similar bounds were derived in the context of interactive function computation in [42]. For $1 \leq l \leq r$, assume that rvs $\mathcal{U}_1, ..., \mathcal{U}_{l-1}$ satisfy the cardinality bounds. We consider odd l; the steps for even l are similar. If the rv U_l does not satisfy the cardinality bound, from the Support Lemma [18, Lemma 15.4], we can replace it with another rv $\tilde{U}_l$ that takes less than or equal to $|\mathcal{X}_1| \prod_{i=1}^{l-1} |\mathcal{U}_i| + 1$ values, while keeping the following quantities unchanged:

$$\mathrm{P}_{X_{1T}U^{l-1}}, \; I(X_{1T} \wedge X_{2T} \mid U^r), \text{ and } I(X_{1T}, X_{2T} \wedge U^r).$$

Note that we have only altered P_{U_l} in the joint pmf $\mathrm{P}_{X_{1T}X_{2T}U^r} = \mathrm{P}_{U_l}\mathrm{P}_{X_{1T}U^{l-1}|U_l}\mathrm{P}_{X_{2T}|X_{1T}U^{l-1}}$. Hence, the Markov relations in (P1) remain unaltered. Furthermore, $\mathrm{P}_{X_{1T}X_{2T}} = \mathrm{P}_{X_1X_2}$. Finally, since the set of pmfs on a finite alphabet is compact, and the choice of ϵ above was arbitrary, it follows upon taking $\epsilon \to 0$ in (5.34) and (5.36) that there exists U_1^r satisfying (P1)-(P3) such that

$$R \geq I(X_1, X_2 \wedge U^r),$$

which completes the proof. □

5.4 Can interaction reduce the communication rate?

It is well known that the SK capacity can be attained by using a simple one-way communication from terminal 1 to terminal 2 (or from $\mathcal{X}_2$ to $\mathcal{X}_1$). Here we derive the minimum rate R_{NI} of such noninteractive communication using the expression for $CI_i^r(X_1; X_2)$ in (5.6). Since this expression has a *double Markov structure*, it can be simplified by the following observation (see [18, Problem 16.25]): If rvs U, X_1, X_2 satisfy

$$U -\!\!\circ\!\!- X_1 -\!\!\circ\!\!- X_2, \quad X_1 -\!\!\circ\!\!- U -\!\!\circ\!\!- X_2, \tag{5.37}$$

then there exist functions $f = f(U)$ and $g = g(X_1)$ such that

(i) $\mathrm{P}\left(f(U) = g(X_1)\right) = 1$;

(ii) $X_1 \mathrel{-\!\!\circ\!\!-} g(X_1) \mathrel{-\!\!\circ\!\!-} X_2$.

In particular, for rvs U, X_1, X_2 that satisfy (5.37), it follows from (i) above that

$$I(X_1, X_2 \wedge U) = I(X_1 \wedge U) \geq I(g(X_1) \wedge f(U)) = H(g(X_1)).$$

Turning to (5.6), for rvs U^r with r odd, the observations above applied to the rvs X_1 and X_2 conditioned on each realization $U^{r-1} = u^{r-1}$ implies that there exists a function $g = g\left(X_1, U^{r-1}\right)$ such that

$$X_1 \mathrel{-\!\!\circ\!\!-} g\left(X_1, U^{r-1}\right), U^{r-1} \mathrel{-\!\!\circ\!\!-} X_2, \tag{5.38}$$

and

$$I\left(X_1, X_2 \wedge U^r\right) \geq I\left(X_1, X_2 \wedge U^{r-1}\right) + H\left(g\left(X_1, U^{r-1}\right) \mid U^{r-1}\right),$$

where rv U^{r-1} satisfies (P1), (P3). Similar observations hold for even r. Thus, for the minimization in (5.6), conditioned on arbitrarily chosen rvs U^{r-1} satisfying (P1), (P3), the rv U_r is selected as a *sufficient statistic for X_2 given the observation* X_1 (sufficient statistic for X_1 given the observation X_2) when r is odd (r is even). Specifically, for $r = 1$, we have

$$CI_i^1(X_1; X_2) = \min_{X_1 \mathrel{-\!\!\circ\!\!-} g_1(X_1) \mathrel{-\!\!\circ\!\!-} X_2} H\left(g_1(X_1)\right), \tag{5.39}$$

and

$$CI_i^1(X_2; X_1) = \min_{X_2 \mathrel{-\!\!\circ\!\!-} g_2(X_2) \mathrel{-\!\!\circ\!\!-} X_1} H\left(g_2(X_2)\right). \tag{5.40}$$

The answer to the optimization problems in (5.39) and (5.40) can be given explicitly. In fact, we specify next a *minimal sufficient statistic* for X_2 on the basis of X_1. Define an equivalence relation on $\mathcal{X}_1$ as follows:

$$x \sim x' \Leftrightarrow \mathrm{P}_{X_2|X_1}\left(y \mid x\right) = \mathrm{P}_{X_2|X_1}\left(y \mid x'\right), \quad y \in \mathcal{X}_2. \tag{5.41}$$

Let g_1^* be the function corresponding to the equivalence classes of $\sim$. We claim that g_1^* is a minimal sufficient statistic for X_2 on the basis of X_1. This expression for the minimal sufficient statistic was also given in [38, Lemma 3.5(4)]. Specifically, $X_1 \mathrel{-\!\!\circ\!\!-} g_1^*(X_1) \mathrel{-\!\!\circ\!\!-} X_2$ since with $g_1^*(X_1) = c$, say, we have

$$\begin{aligned}
\mathrm{P}_{X_2|g_1^*(X_1)}\left(y \mid c\right) &= \sum_{x \in \mathcal{X}_1} \mathrm{P}_{X_2,X_1|g_1^*(X_1)}\left(y, x \mid c\right) \\
&= \sum_{x : g_1^*(x) = c} \mathrm{P}_{X_1|g_1^*(X_1)}\left(x \mid c\right) \mathrm{P}_{X_2|X_1,g_1^*(X_1)}\left(y \mid x, c\right) \\
&= \mathrm{P}_{X_2|X_1,g_1^*(X_1)}\left(y \mid x, c\right), \quad \forall\, x \text{ with } g_1^*(x) = c.
\end{aligned}$$

Also, if $g_1(X_1)$ satisfies $X_1 \mathrel{-\!\!\circ\!\!-} g_1(X_1) \mathrel{-\!\!\circ\!\!-} X_2$ then g_1^* is a function of g_1. To see this, let $g_1(x) = g_1(x') = c$ for some $x, x' \in \mathcal{X}_1$. Then,

$$\mathrm{P}_{X_2|g_1(X_1)}\left(y \mid c\right) = \mathrm{P}_{X_2|X_1}\left(y \mid x\right) = \mathrm{P}_{X_2|X_1}\left(y \mid x'\right), \quad y \in \mathcal{X}_2,$$

so that $g_1^*(x) = g_1^*(x')$. Since g_1^* is a minimal sufficient statistic for X_2 on the basis of X_1, it follows from (5.39) that

$$CI_i^1(X_1; X_2) = H\left(g_1^*(X_1)\right),$$

and similarly, with $g_2^*(X_2)$ defined analogously,

$$CI_i^1(X_2; X_1) = H\left(g_2^*(X_2)\right).$$

Therefore, from (5.4), the minimum rate R_{NI} of a noninteractive communication for generating a maximum rate SK is given by

$$R_{NI} = \min\left\{H\left(g_1^*(X_1)\right); H\left(g_1^*(X_1)\right)\right\} - I(X_1 \wedge X_2). \tag{5.42}$$

From the expression for R_{NI}, it is clear that the rate of noninteractive communication can be reduced by replacing X_1 and X_2 with their respective minimal sufficient statistics $g_1^*(X_1)$ and $g_2^*(X_2)$. Can the rate of communication required for generating an optimum rate SK be reduced by resorting to complex interactive communication protocols? To answer this question we must compare the expression for R_{NI} with R_{SK}. Specifically, from Theorem 5.1 and the Corollary following it, interaction reduces the rate of communication iff, for some $r > 1$,

$$CI_i^r(X_1; X_2) < \min\left\{H\left(g_1^*(X_1)\right); H\left(g_1^*(X_1)\right)\right\}, \tag{5.43}$$

where g_1^* and g_2^* are as in (5.42); interaction does not help iff

$$CI_i(X_1 \wedge X_2) = \min\left\{H\left(g_1^*(X_1)\right); H\left(g_1^*(X_1)\right)\right\}.$$

Note that instead of comparing with $CI_i^r(X_1; X_2)$ in (5.43), we can also compare with $CI_i^r(X_2; X_1)$.

We shall explore this question here, and give an example where the answer is in the affirmative. In fact, we first show that interaction does not help in the case of BSS. Then we give an example where interaction does help.

5.4.1 Binary symmetric sources

For BSS X_1 and X_2, we note a property of rvs U^r that satisfy the conditions (P1)-(P3) in Theorem 5.3.

Lemma 5.6. *Let X_1 and X_2 be $\{0,1\}$ valued rvs with $I(X_1 \wedge X_2) \neq 0$. Then, for rvs $U_1, ..., U_r$ that satisfy the conditions (P1)-(P3) in Theorem 5.3, for every realization $u_1, ..., u_r$ of $U_1, ..., U_r$, one of the following holds:*

$$H(X_1 \mid U^r = u^r) = 0, \quad \textit{or} \;\; H(X_2 \mid U^r = u^r) = 0. \tag{5.44}$$

Proof. Given a sequence u^r, assume that

$$H(X_1 \mid U^r = u^r) > 0 \;\; \text{and} \;\; H(X_2 \mid U^r = u^r) > 0,$$

which is equivalent to

$$\mathrm{P}_{X_1|U^r}(1 \mid u^r)\, \mathrm{P}_{X_1|U^r}(0 \mid u^r) > 0 \;\; \text{and} \;\; \mathrm{P}_{X_2|U^r}(1 \mid u^r)\, \mathrm{P}_{X_2|U^r}(0 \mid u^r) > 0. \tag{5.45}$$

We consider the case when r is even; the case of odd r is handled similarly. From the Markov conditions $X_1 -\!\!\circ\!\!- U^r -\!\!\circ\!\!- X_2$ and $X_1 -\!\!\circ\!\!- X_2, U^{r-1} -\!\!\circ\!\!- U_r$, we have

$$\begin{aligned}\mathrm{P}_{X_1,X_2|U^r}(x, y \mid u^r) &= \mathrm{P}_{X_1|U^r}(x \mid u^r)\, \mathrm{P}_{X_2|U^r}(y \mid u^r) \\ &= \mathrm{P}_{X_1|X_2,U^{r-1}}\left(x \mid y, u^{r-1}\right) \mathrm{P}_{X_2|U^r}(y \mid u^r), \quad x, y \in \{0,1\}.\end{aligned}$$

Since $\mathrm{P}_{X_2|U^r}(y \mid u^r) > 0$ from (5.45), we have

$$\mathrm{P}_{X_1|U^r}(x \mid u^r) = \mathrm{P}_{X_1|X_2,U^{r-1}}\left(x \mid y, u^{r-1}\right), \quad x, y \in \{0,1\},$$

which further implies

$$\mathrm{P}_{X_1|X_2,U^{r-1}}\left(x \mid 1, u^{r-1}\right) = \mathrm{P}_{X_1|X_2,U^{r-1}}\left(x \mid 0, u^{r-1}\right),$$

$$x \in \{0,1\}.$$

Hence, $I(X_1 \wedge X_2 \mid U^{r-1} = u^{r-1}) = 0$. Noting from (5.45) that

$$\mathrm{P}_{X_1|U^{r-1}}\left(1 \mid u^{r-1}\right)\mathrm{P}_{X_1|U^{r-1}}\left(0 \mid u^{r-1}\right) > 0,$$

we can do the same analysis as above, again for $r-1$. Upon repeating this process r times we get $I(X_1 \wedge X_2) = 0$, which is a contradiction. Therefore, either $H(X_1 \mid U^r = u^r) = 0$ or $H(X_2 \mid U^r = u^r) = 0$ holds. □

Note that

$$CI_i^r(X_1; X_2) = H(X_1, X_2) - \max_{U^r} H(X_1, X_2 \mid U^r),$$

where the max is taken over rvs U^r as in Theorem 5.3. If $H(X_1 \mid U^i = u^i) = 0$, it follows that

$$\begin{aligned} &I\left(X_1 \wedge X_2 \mid U^i = u^i\right) = 0, \text{ and} \\ &H(X_1, X_2 \mid U^i = u^i, U_{i+1}^r) = H(X_2 \mid U^i = u^i, U_{i+1}^r) \\ &\qquad \le H(X_2 \mid U^i = u^i). \end{aligned} \tag{5.46}$$

Similarly, $H(X_2 \mid U^i = u^i) = 0$ implies

$$\begin{aligned} &I\left(X_1 \wedge X_2 \mid U^i = u^i\right) = 0, \text{ and} \\ &H(X_1, X_2 \mid U^i = u^i, U_{i+1}^r) \le H(X_1 \mid U^i = u^i). \end{aligned} \tag{5.47}$$

For a sequence u^r with $\mathrm{P}_{U^r}(u^r) > 0$, let $\tau(u^r)$ be the minimum value of i such that

$$H(X_1 \mid U^i = u^i) = 0 \text{ or } H(X_2 \mid U^i = u^i) = 0;$$

if X_1 and X_2 are independent, $\tau(u^r) = 0$. Note that τ is a stopping-time adapted to $U_1, ..., U_r$. Then, from (5.46), (5.47), $CI_i^r(X_1; X_2)$ remains unchanged if we restrict the support of U^r to sequences u^r with $u_i = \phi$ for all $i > \tau(u^r)$. Furthermore, the Markov condition (P1) implies that if for a sequence u^r, $\tau = \tau(u^r)$ is odd then

$$\mathrm{P}_{X_2|X_1,U^\tau}(y \mid x, u^\tau) = \mathrm{P}_{X_2|X_1,U^{\tau-1}}\left(y \mid x, u^{\tau-1}\right),$$

and so if

$$\mathrm{P}_{X_1|U^\tau}(1 \mid u^\tau)\,\mathrm{P}_{X_1|U^\tau}(0 \mid u^\tau) > 0,$$

it holds from the definition of τ that

$$\mathrm{P}_{X_2|U^\tau}(1 \mid u^\tau)\,\mathrm{P}_{X_2|U^\tau}(0 \mid u^\tau) > 0,$$

which is a contradiction. Therefore, we have $H(X_1 \mid U^\tau = u^\tau) = 0$. Similarly, $H(X_2 \mid U^\tau = u^\tau) = 0$ holds for even τ. To summarize,

$$CI_i^r(X_1; X_2) = \min_{U^\tau} I\left(X_1, X_2 \wedge U^\tau\right), \tag{5.48}$$

where U^r are rvs satisfying (P1)-(P3), and τ is the stopping-time defined above.

We show next that for BSS, interaction can never reduce the rate of communication for optimum rate SK generation. In fact, *we conjecture that for any BSS X_1, X_2, $R_{NI} = R_{SK}$.*

Theorem 5.7. *Let X_1 and X_2 be $\{0,1\}$-valued rvs, with*

$$\begin{aligned}
&\mathrm{P}\left(X_1 = 0, X_2 = 0\right) = \mathrm{P}\left(X_1 = 1, X_2 = 1\right) = \frac{1}{2}(1-\delta), \\
&\mathrm{P}\left(X_1 = 0, X_2 = 1\right) = \mathrm{P}\left(X_1 = 1, X_2 = 0\right) = \frac{1}{2}\delta, \quad 0 < \delta < \frac{1}{2}.
\end{aligned} \tag{5.49}$$

Then,

$$CI_i(X_1 \wedge X_2) = \min\{H(X_1); H(X_2)\},$$

i.e., interaction does not help to reduce the communication required for optimum rate SK generation.

Remark. As a consequence of Theorem 5.7, for sources with joint distribution as in (5.49), the second inequality in (5.3) can be strict. Specifically, it was noted by Wyner (see the discussion following equation (1.19) in [77]) that for BSS, $CI_W(X_1 \wedge X_2) < 1$. From Theorem 5.7, we have

$$CI_i(X_1 \wedge X_2) = \min\{H(X_1); H(X_2)\} = 1.$$

Thus, for BSS X_1 and X_2, $CI_W(X_1 \wedge X_2) < CI_i(X_1 \wedge X_2)$.

Proof. Denote by $\mathcal{U}_0^r$ the following set of stopped sequences in $\mathcal{U}^r$:

For $i \leq r$, for a sequence $u^r \in \mathcal{U}^r$ the stopped sequence $u^i \in \mathcal{U}_0^r$ if:

$$\begin{aligned}
&H\left(X_1 \mid U^j = u^j\right) > 0, H\left(X_2 \mid U^j = u^j\right) > 0, \quad \forall\, j < i, \text{ and} \\
&H\left(X_1 \mid U^i = u^i\right) = 0 \text{ or } H\left(X_2 \mid U^i = u^i\right) = 0.
\end{aligned}$$

For $i \in \{0,1\}$, define the following subsets of $\mathcal{U}_0^r$:

$$\mathcal{U}_i^1 = \left\{u^\tau \in \mathcal{U}_0^r : \tau \text{ is odd}, \mathrm{P}_{X_1|U^\tau}\left(i \mid u^\tau\right) = 1\right\},$$

$$\mathcal{U}_i^2 = \left\{u^\tau \in \mathcal{U}_0^r : \tau \text{ is even}, \mathrm{P}_{X_2|U^\tau}\left(i \mid u^\tau\right) = 1\right\}.$$

By their definition the sets $\mathcal{U}_0^1$,$\mathcal{U}_1^1$,$\mathcal{U}_0^2$, and $\mathcal{U}_1^2$ are disjoint, whereby we have

$$\begin{aligned}\mathrm{P}_{U^\tau}\left(\mathcal{U}_0^r\right) &= \mathrm{P}_{U^r}\left(\mathcal{U}_0^1\bigcup\mathcal{U}_1^1\bigcup\mathcal{U}_0^2\bigcup\mathcal{U}_1^2\right)\\ &= \sum_{i=0}^{1}\left[\mathrm{P}_{U^r}\left(\mathcal{U}_i^1\right)+\mathrm{P}_{U^r}\left(\mathcal{U}_i^2\right)\right] = 1. \end{aligned} \tag{5.50}$$

For $u^\tau \in \mathcal{U}_0^r$, denote by $p(u^\tau)$ the probability $\mathrm{P}_{U^\tau}\left(u^\tau\right)$. Further, for $u^\tau \in \mathcal{U}_0^1\bigcup\mathcal{U}_1^1$, denote by $W^{u^\tau} : \mathcal{X}_1 \to \mathcal{X}_2$ the stochastic matrix corresponding to $\mathrm{P}_{X_2|X_1,U^\tau}\left(\cdot \mid \cdot, u^\tau\right)$, and for $u^\tau \in \mathcal{U}_0^2\bigcup\mathcal{U}_1^2$, denote by $T^{u^\tau} : \mathcal{X}_2 \to \mathcal{X}_1$ the stochastic matrix corresponding to $\mathrm{P}_{X_1|X_2,U^\tau}\left(\cdot \mid \cdot, u^\tau\right)$. With this notation, the following holds:

$$\begin{aligned}\frac{1}{2}(1-\delta) &= \mathrm{P}_{X_1,X_2}\left(i,i\right)\\ &= \sum_{u^\tau\in\mathcal{U}_i^1} p(u^\tau)W^{u^\tau}(i\mid i) + \sum_{u^\tau\in\mathcal{U}_i^2} p(u^\tau)T^{u^\tau}(i\mid i), \quad i\in\{0,1\},\end{aligned} \tag{5.51}$$

since the sets $\mathcal{U}_0^1$,$\mathcal{U}_1^1$,$\mathcal{U}_0^2$,$\mathcal{U}_1^2$ are disjoint. Upon adding (5.51) for $i = 0, 1$, we get

$$\sum_{i=0}^{1}\left[\sum_{u^\tau\in\mathcal{U}_i^1} p(u^\tau)W^{u^\tau}(i\mid i) + \sum_{u^\tau\in\mathcal{U}_i^2} p(u^\tau)T^{u^\tau}(i\mid i)\right] = (1-\delta).$$

Furthermore, from (5.50) we get

$$1 = \sum_{i=0}^{1}\sum_{u^\tau\in\mathcal{U}_i^1} p(u^\tau) + \sum_{u^\tau\in\mathcal{U}_i^2} p(u^\tau).$$

Therefore, since the function $g(z) = -z\log z$ is concave for $0 < z < 1$, the Jensen's inequality yields

$$g(1-\delta) \geq \sum_{i=0}^{1}\sum_{u^\tau\in\mathcal{U}_i^1} p(u^\tau)g\left(W^{u^\tau}(i\mid i)\right) + \sum_{u^\tau\in\mathcal{U}_i^2} p(u^\tau)g\left(T^{u^\tau}(i\mid i)\right) \tag{5.52}$$

Similarly, using for $i \neq j, i, j \in \{0,1\}$,

$$\begin{aligned}\frac{1}{2}\delta &= \mathrm{P}_{X_1,X_2}\left(i,j\right)\\ &= \sum_{u^\tau\in\mathcal{U}_i^1} p(u^\tau)\left(1-W^{u^\tau}(i\mid i)\right) + \sum_{u^\tau\in\mathcal{U}_j^2} p(u^\tau)\left(1-T^{u^\tau}(j\mid j)\right),\end{aligned}$$

we get

$$g(\delta) \geq \sum_{i=0}^{1} \sum_{u^\tau \in \mathcal{U}_i^1} p(u^\tau) g\left(1 - W^{u^\tau}(i \mid i)\right) + \sum_{u^\tau \in \mathcal{U}_i^2} p(u^\tau) g\left(1 - T^{u^\tau}(i \mid i)\right) \qquad (5.53)$$

On adding (5.52) and (5.53) we get

$$\begin{aligned} h(\delta) &= g(\delta) + g(1-\delta) \\ &\geq \sum_{i=0}^{1} \sum_{u^\tau \in \mathcal{U}_i^1} p(u^\tau) h\left(W^{u^\tau}(i \mid i)\right) + \sum_{u^\tau \in \mathcal{U}_i^2} p(u^\tau) h\left(T^{u^\tau}(i \mid i)\right). \end{aligned}$$

Note that the right-side above equals $H(X_1, X_2 \mid U^\tau)$, which yields

$$h(\delta) = \max\{H(X_1 \mid X_2); H(X_2 \mid X_1)\} \geq H(X_1, X_2 \mid U^\tau).$$

Since rvs U^r above were arbitrary, we have from (5.48),

$$\begin{aligned} CI_i^r(X_1; X_2) &\geq H(X_1, X_2) - \max\{H(X_1 \mid X_2); H(X_2 \mid X_1)\} \\ &= \min\{H(X_1); H(X_2)\}. \end{aligned}$$

Combining this with (5.3), we obtain

$$CI_i^r(X_1; X_2) = \min\{H(X_1); H(X_2)\}.$$

□

5.4.2 An example where interaction does help

Consider rvs X_1 and X_2 with $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1, 2\}$, and with joint pmf:

$$\begin{bmatrix} a & a & a \\ b & a & a \\ a & c & a \end{bmatrix},$$

where a, b, c are nonnegative, $7a + b + c = 1$, and $c \neq a$, which holds iff $b \neq 1 - 8a$. Assume that

$$2a > b > a. \tag{5.54}$$

From (5.43), to show that interaction helps, it suffices to find rvs $U_1, ..., U_r$ satisfying (P1)-(P3) such that

$$I\left(X_1, X_2 \wedge U_1, ..., U_r\right) < \min\left\{H\left(g_1^*(X_1)\right); H\left(g_2^*(X_2)\right)\right\}, \tag{5.55}$$

where g_1^* and g_2^* are as in (5.42). From (5.41), $g_1^*(x) = g_1^*(x')$ iff

$$\frac{\mathrm{P}_{X_2,X_1}\left(y, x\right)}{\mathrm{P}_{X_2,X_1}\left(y, x'\right)} = \frac{\mathrm{P}_{X_1}\left(x\right)}{\mathrm{P}_{X_1}\left(x'\right)}, \quad y \in \mathcal{X}_2, \tag{5.56}$$

i.e., the ratio $\frac{\mathrm{P}_{X_2,X_1}(y,x)}{\mathrm{P}_{X_2,X_1}(y,x')}$ does not depend on y. Therefore, for the pmf above, $g_1^*(X_1)$ and $g_2^*(X_2)$ are equivalent to X_1 and X_2, respectively. Thus,

$$\min\left\{H\left(g_1^*(X_1)\right); H\left(g_2^*(X_2)\right)\right\} = \min\{H(X_1); H(X_2)\},$$

where $H(X_1) = H(X_2)$ for the given pmf.

Next, let $U_1 = f_1(X_1)$, $U_2 = f_2(X_2, f_1(X_1))$, where f_1 and f_2 are given below:

$$f_1(x) = \begin{cases} 1, & x = 2, \\ 2, & x = 0, 1, \end{cases}$$

$$f_2(y, 1) = 0, \forall\, y \in \{0, 1, 2\}, \text{ and } \; f_2(y, 2) = \begin{cases} 1, & y = 0, \\ 2, & y = 1, 2. \end{cases}$$

Clearly, U_1 and U_2 satisfy (P1) and (P2). For (P3), note that if $(U_1, U_2) = (1, 0)$, then $X_1 = 2$, and if $(U_1, U_2) = (2, 1)$, then $X_2 = 0$. Finally, if $(U_1, U_2) = (2, 2)$, then

$X_1 \in \{0,1\}$ and $X_2 \in \{1,2\}$, implying

$$\begin{aligned}\mathrm{P}_{X_1,X_2|U_1,U_2}\left(x,y \mid 2,2\right) &= \frac{\mathrm{P}_{X_1,X_2}\left(x,y\right)}{4a} \\ &= \frac{1}{4}, \quad \forall\ (x,y) \in \{0,1\} \times \{1,2\}.\end{aligned}$$

Therefore, $I(X_1 \wedge X_2 \mid U_1, U_2) = 0$, and so U_1, U_2 satisfy (P3). We show that (5.55) holds for this choice of U_1, U_2. Specifically, $I\left(X_1, X_2 \wedge U_1, U_2\right) = H\left(U_1, U_2\right)$, and the following holds:

$$\begin{aligned}H(X_2) - H\left(U_1, U_2\right) &= H(X_1) - H\left(U_1, U_2\right) \\ &= H\left(X_1|U_1\right) - H\left(U_2|U_1\right) \\ &= \mathrm{P}\left(f_1(X_1) = 2\right)\Big[H\left(X_1|f_1(X_1)=2\right) - H\left(f_2(2,X_2)|f_1(X_1)=2\right)\Big] \\ &= (5a+b)\left[h\left(\mathrm{P}_{X_1|f_1(X_1)}\left(0|2\right)\right) - h\left(\mathrm{P}_{X_2|f_1(X_1)}\left(0|2\right)\right)\right] \\ &= (5a+b)\left[h\left(\frac{3a}{5a+b}\right) - h\left(\frac{a+b}{5a+b}\right)\right].\end{aligned}$$

Then, from (5.54),

$$\frac{a+b}{5a+b} < \frac{3a}{5a+b} < \frac{1}{2},$$

which implies (5.55) for U_1, U_2.

5.5 Interactive common information and sufficient statistics

In Chapter 2, we saw that several CI quantities remain unchanged if X_1 and X_2 are replaced, respectively, with their sufficient statistics (for the other). This is true even for the interactive CI.

Theorem 5.8. *For rvs X_1 and X_2, let functions g_1 of X_1 and g_2 of X_2 be such that $X_1 -\!\!\circ\!\!- g_1(X_1) -\!\!\circ\!\!- X_2$ and $X_1 -\!\!\circ\!\!- g_2(X_2) -\!\!\circ\!\!- X_2$. Then the following relations hold:*

$$CI_i^r(X_1; X_2) = CI_i^r\left(g_1(X_1); g_2(X_2)\right), \quad r \geq 1,$$

$$CI_i(X_1 \wedge X_2) = CI_i\left(g_1(X_1) \wedge g_2(X_2)\right).$$

Remark. (i) Theorem 5.8 implies that the minimum rate of communication for generating a maximum rate SK remains unchanged if X_1 and X_2 are replaced by $g_1(X_1)$ and $g_2(X_2)$ as above, respectively.
(ii) Note that $g_1(X_1)$ and $g_2(X_2)$ above are, respectively, functions of $g_1^*(X_1)$ and $g_2^*(X_2)$ defined through (5.41).

Proof. First note that

$$I(X_1 \wedge X_2) = I\left(g_1(X_1) \wedge X_2\right) = I\left(g_1(X_1) \wedge g_2(X_2)\right). \tag{5.57}$$

Proof. . From (5.57), any protocol that generates an optimum rate SK for the sources $g_1(X_1)$ and $g_2(X_2)$ also generates an optimum rate SK for the sources X_1 and X_2. Thus, the minimum communication rate for prior protocols is bounded below by the minimum communication rate for the latter protocols, so that by Theorem 5.1,

$$CI_i^r\left(g_1(X_1); g_2(X_2)\right) - I\left(g_1(X_1) \wedge g_2(X_2)\right) \geq CI_i^r\left(X_1; X_2\right) - I\left(X_1 \wedge X_2\right),$$

which, by (5.57), is

$$CI_i^r\left(g_1(X_1); g_2(X_2)\right) \geq CI_i^r\left(X_1; X_2\right). \tag{5.58}$$

In fact, (5.58) holds with equality: We claim that any choice of rvs U^r that satisfy (P1)-(P3) also satisfy the following Markov relations:

$$\begin{aligned}
&U_{2i+1} -\!\!\circ\!\!- g_1(X_1), U^{2i} -\!\!\circ\!\!- g_2(X_2), \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\
&U_{2i} -\!\!\circ\!\!- g_2(X_2), U^{2i-1} -\!\!\circ\!\!- g_1(X_1), \quad 1 \leq i \leq \lfloor r/2 \rfloor, \\
&g_1(X_1) -\!\!\circ\!\!- U^r -\!\!\circ\!\!- g_2(X_2). \qquad (5.59)
\end{aligned}$$

It follows that

$$CI_i^r \left(g_1(X_1); g_2(X_2) \right) \leq I \left(g_1(X_1), g_2(X_2) \wedge U^r \right) \leq I \left(X_1, X_2 \wedge U^r \right), \qquad (5.60)$$

and consequently,

$$CI_i^r \left(g_1(X_1); g_2(X_2) \right) \leq CI_i^r \left(X_1; X_2 \right).$$

Thus, by (5.58),

$$CI_i^r \left(g_1(X_1); g_2(X_2) \right) = CI_i^r \left(X_1; X_2 \right). \qquad (5.61)$$

Taking the limit $r \to \infty$ we get

$$CI_i \left(g_1(X_1) \wedge g_2(X_2) \right) = CI_i \left(X_1 \wedge X_2 \right).$$

It remains to establish (5.59); instead, using induction we establish the following stronger Markov relations: For $1 \leq i \leq r$,

$$\begin{aligned}
&U_i -\!\!\circ\!\!- g_1(X_1), U^{i-1} -\!\!\circ\!\!- X_2, \quad i \text{ odd}, \\
&U_i -\!\!\circ\!\!- g_2(X_2), U^{i-1} -\!\!\circ\!\!- X_1, \quad i \text{ even}, \\
&X_1 -\!\!\circ\!\!- g_1(X_1), U^i -\!\!\circ\!\!- X_2 \text{ and } X_1 -\!\!\circ\!\!- g_2(X_2), U^i -\!\!\circ\!\!- X_2. \qquad (5.62)
\end{aligned}$$

Clearly, (5.62) implies the first two Markov relations in (5.59). The last Markov chain in (5.59) follows upon observing

$$0 = I\left(X_1 \wedge X_2 \mid U^r\right) \geq I\left(g_1(X_1) \wedge g_2(X_2) \mid U^r\right).$$

To see that (5.62) holds for $i = 1$ note that

$$I\left(X_1 \wedge X_2 \mid g_1(X_1), U_1\right) \leq I\left(X_1 \wedge X_2 \mid g_1(X_1)\right) + I\left(U_1 \wedge X_2 \mid g_1(X_1), X_1\right) = 0,$$

and

$$I\left(X_1 \wedge X_2 \mid g_2(X_2), U_1\right) \leq I\left(X_1 \wedge X_2 \mid g_2(X_2)\right) + I\left(U_1 \wedge X_2, g_2(X_2) \mid X_1\right) = 0.$$

Next, assume that (5.62) holds for an even i. Then, from (P1) we get:

$$\begin{aligned}
&I\left(X_2 \wedge U_{i+1} \mid X_1, U^i\right) = 0 \\
\Leftrightarrow &I\left(X_2 \wedge U_{i+1} \mid X_1, g_1(X_1), U^i\right) = 0 \\
\Leftrightarrow &I\left(X_2 \wedge X_1, U_{i+1} \mid g_1(X_1), U^i\right) = I\left(X_2 \wedge X_1 \mid g_1(X_1), U^i\right) = 0,
\end{aligned}$$

where the last equality follows from (5.62). From the last inequality above we have

$$U_{i+1} \mathrel{-\!\!\circ\!\!-} g_1(X_1), U^i \mathrel{-\!\!\circ\!\!-} X_2 \text{ and } X_1 \mathrel{-\!\!\circ\!\!-} g_1(X_1), U^{i+1} \mathrel{-\!\!\circ\!\!-} X_2.$$

Furthermore, it also follows from (5.62) that

$$\begin{aligned}
I\left(X_1 \wedge X_2 \mid g_2(X_2), U^{i+1}\right) &\leq I\left(X_1, U_{i+1} \wedge X_2 \mid g_2(X_2), U^i\right) \\
&= I\left(U_{i+1} \wedge X_2 \mid g_2(X_2), X_1, U^i\right) \\
&\leq I\left(U_{i+1} \wedge X_2 \mid X_1, U^i\right) = 0,
\end{aligned}$$

where the last equality follows from (P1). Thus, we have

$$X_1 \mathrel{-\!\!\circ\!\!-} g_2(X_2), U^{i+1} \mathrel{-\!\!\circ\!\!-} X_2,$$

establishing the validity of (5.62) for $i+1$. The proof of (5.59) can be completed by induction by using a similar argument for odd i. □

5.6 Discussion

5.6.1 Local randomization

Although independent local randomization was not allowed in our formulation, our main result characterizing R_{SK} holds even when such randomization is available. Consider a model where terminals 1 and 2, in additional to their respective observations X_1^n and X_2^n, have access to finite-valued[1] rvs T_1 and T_2, respectively. The rvs T_1, T_2, and (X_1^n, X_2^n) are mutually independent. The SK capacity is defined as before, with X_1^n and X_2^n now replaced by (X_1^n, T_1) and (X_2^n, T_2), respectively. It is known [48, 2] that even with randomization the SK capacity equals $I(X_1 \wedge X_2)$. For this model, denote the minimum rate of r-interactive communication required to generate an SK of rate $I(X_1 \wedge X_2)$ by $\tilde{R}_{SK}^r$.

Lemma 5.9. *For $r \geq 1$,*

$$\tilde{R}_{SK}^r = R_{SK}^r.$$

To see this, we define quantities $\tilde{R}_{CI}^r$ and $\tilde{CI}_i^r$ analogously to R_{SK}^r and CI_i^r, with X_1^n and X_2^n replaced by (X_1^n, T_1) and (X_2^n, T_2), respectively. Note that this substitution is made even in condition (5.1), i.e., the CR J and the communication

[1] The cardinalities of the range spaces of T_1 and T_2 are allowed to be at most exponential in n.

$\mathbf{F}$ now are required to satisfy:

$$\frac{1}{n} I\left(X_1^n, T_1 \wedge X_2^n, T_2 \mid J, \mathbf{F}\right) \leq \epsilon. \tag{5.63}$$

We observe that (5.8) still holds, with (X_1^n, T_1) and (X_2^n, T_2) replacing, respectively, X_1^n and X_2^n on the right-side. Therefore, the proof of Theorem 5.1 is valid, and we get:

$$\tilde{R}_{CI}^r = \tilde{R}_{SK}^r = \tilde{CI}_i^r - I(X_1 \wedge X_2). \tag{5.64}$$

By its definition $\tilde{R}_{CI}^r \leq R_{CI}^r$, since $L = (J, \mathbf{F}) = L(X_1^n, X_2^n)$ satisfying (5.1) will meet (5.63) as well. We claim that $\tilde{R}_{CI}^r \geq R_{CI}^r$, which by (5.64) and Theorem 5.1 implies Lemma 5.9. Indeed, consider CR J recoverable from $\mathbf{F}$ such that $(J, \mathbf{F})$ attain $\tilde{R}_{CI}^r$. Then, the condition (5.63) gives

$$\frac{1}{n} I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}, T_1, T_2\right) \approx 0.$$

So, there exist t_1, t_2 such that conditioned on $T_1 = t_1, T_2 = t_2$ the CR J is still recoverable from $\mathbf{F}$, and

$$\frac{1}{n} I\left(X_1^n \wedge X_2^n \mid J, \mathbf{F}, T_1 = t_1, T_2 = t_2\right) \approx 0.$$

Thus, with $T_1 = t_1, T_2 = t_2$ fixed, $(J, \mathbf{F})$ constitutes a feasible choice in the definition of R_{CI}^r. Since the number of values taken by $\mathbf{F}$ can only decrease upon fixing $T_1 = t_1, T_2 = t_2$, we get $\tilde{R}_{CI}^r \geq R_{CI}^r$. Therefore, the availability of local randomization does not decrease the rate of communication required for generating an optimum rate SK.

5.6.2 Less-than-optimum rate secret keys

SK generation is linked intrinsically to the efficient generation of CR. For $\rho \geq 0$, a rate $R \geq 0$ is an achievable CR rate for ρ if for every $0 < \epsilon < 1$ there exists, for some $n \geq 1$, an ϵ-CR L with

$$\frac{1}{n} H(L) \geq R - \epsilon,$$

recoverable from an r-interactive communication $\mathbf{F}$, for arbitrary r, of rate

$$\frac{1}{n} H(\mathbf{F}) \leq \rho + \epsilon;$$

the maximum achievable CR rate for ρ is denoted by $CR(\rho)$. Similarly, denote by $C(\rho)$ the maximum rate of an SK that can be generated using a communication as above. It can be shown in a straightforward manner that

$$C(\rho) = CR(\rho) - \rho. \tag{5.65}$$

The graph of CR as a function of ρ is plotted in Fig. 5.1. $CR(\rho)$ is an increasing and a concave function of ρ, as seen from a simple time-sharing argument. Since R_{SK} is the minimum rate of communication required to generate a maximum rate SK, $CR(\rho) - \rho = I(X_1 \wedge X_2)$ for $\rho \geq R_{SK}$. Thus, our results characterize the graph of $CR(\rho)$ for all $\rho \geq R_{SK}$. The quantity R_{SK} is the minimum value of ρ for which the slope of $CR(\rho)$ is 1; $CR\left(R_{SK}\right)$ is equal to the interactive CI $CI_i(X_1 \wedge X_2)$. Furthermore, from the proof of Theorem 5.1, a CR L that satisfies (5.1) must yield an optimum rate SK. Thus, any CR recoverable from a communication of rate less than R_{SK} cannot satisfy (5.1). A characterization of $CR(\rho)$ for $\rho < R_{SK}$ is central

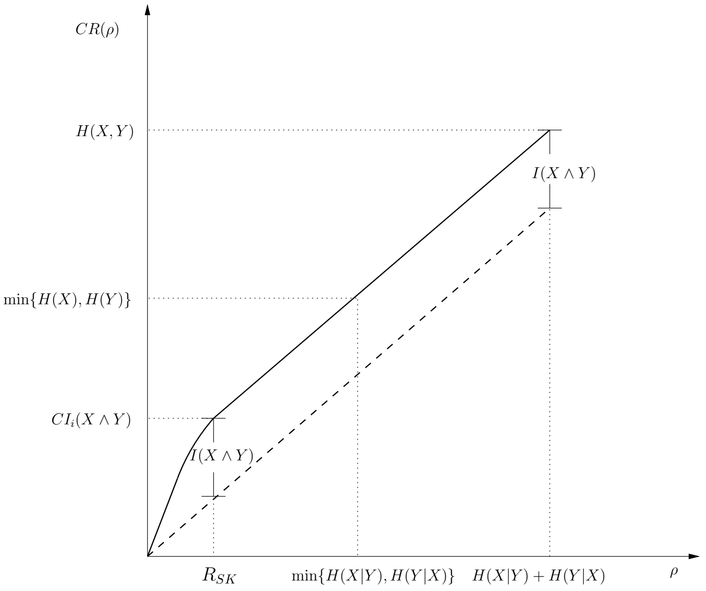

Figure 5.1: Minimum rate of communication R_{SK} for optimum rate SK generation

to the characterization of $C(\rho)$, and this, along with a single-letter characterization of R_{SK}, is an open problem.

We close this chapter by remarking that an extension of the results of this chapter to the case of multiple terminals is not known. In particular, an appropriate notion of interactive CI for multiple terminals, with connections to the minimal communication for optimum rate SK generation, is unavailable. The identification of such a notion of multiterminal interactive CI will lead to a characterization of the CR underlying optimum rate SK generation for multiple terminals and shed light on the structure of such SKs; it remains an interesting open problem.

CHAPTER 6

Querying Common Randomness

6.1 Synopsis

A set of m terminals, observing correlated signals, communicate interactively to generate common randomness for a given subset of them. Knowing only the communication, how many direct queries of the value of the common randomness will resolve it? A general upper bound, valid for arbitrary signal alphabets, is developed for the number of such queries by using a query strategy that applies to all common randomness and associated communication. When the underlying signals are i.i.d. repetitions of m correlated random variables, the number of queries can be exponential in signal length. For this case, the mentioned upper bound is tight and leads to a single-letter formula for the largest query exponent, which coincides with the secret key capacity of a corresponding multiterminal source model. In fact, the upper bound constitutes a strong converse for the optimum query exponent, and implies also a new strong converse for secret key capacity.

The problem formulation and our main result characterizing the optimum query exponent are given in the next section. Simple and essential technical tools are presented in Section 6.3. Achievability is proved in Section 6.4. The less complex converse proof for the case $\mathcal{A} = \{1, ..., m\}$ is given in Section 6.5. However, this proof does not extend to an arbitrary $\mathcal{A} \subseteq \{1, ..., m\}$, for which a different converse is provided in Section 6.6. Section 6.7 contains the strong converse result for SK capacity. A converse for the optimum query exponent for rvs with arbitrary alphabets is proved in Section 6.8, with jointly Gaussian rvs as a special case. The results of this chapter were reported in [66, 65].

6.2 Main result: How many queries will resolve common randomness?

Consider the multiterminal source model of Section 2.2. The terminals communicate interactively with each other over a public communication network, as described in Definition 2.1, in order to generation a CR (see Definition 2.2). For simplicity, we do not allow independent local randomization $U_{\mathcal{M}}$. However, our results do not change if such randomization is allowed; see Section 6.9.

A querier observing the communication $\mathbf{F}$ wants to resolve the value of this CR L by asking questions of the form "Is $L = l$?" with yes-no answers. While queries of this form have been termed "guessing" [47, 4, 5, 34], we use the terminology "query" since our approach covers a broader class of query strategies; see Section 6.9.

Definition 6.1. For rvs U, V with values in the sets $\mathcal{U}, \mathcal{V}$, a *query strategy* q for U given $V = v$ is a bijection $q(\cdot|v) : \mathcal{U} \rightarrow \{1, ..., |\mathcal{U}|\}$, where the querier, upon observing $V = v$, asks the question "Is $U = u$?" in the $q(u|v)^{\text{th}}$ query.

Thus, a query strategy q for resolving a CR L on the basis of an observed communication $\mathbf{F} = \mathbf{i}$ is an ordering of the possible values of L. The terminals seek to generate a CR L for $\mathcal{A}$ using communication $\mathbf{F}$ so as to make the task of the querier observing $\mathbf{F}$ as onerous as possible. For instance, if L were to be independent of $\mathbf{F}$, then the querier necessarily must search exhaustively over the set of possible values of L, which can be exponentially large (in n).

Definition 6.2. Given $0 < \epsilon < 1$, a *query exponent* $E > 0$ is ϵ-achievable if for every $0 < \epsilon' < 1$, there exists an ϵ-CR $L = L^{(n)}\left(X_{\mathcal{M}}^n\right)$ for $\mathcal{A} \subseteq \mathcal{M}$ from communication $\mathbf{F} = \mathbf{F}\left(X_{\mathcal{M}}^n\right)$ such that for every query strategy q for L given $\mathbf{F}$,

$$\mathrm{P}\big(q(L \mid \mathbf{F}) \geq \exp(nE)\big) > 1 - \epsilon', \tag{6.1}$$

for all $n \geq N(\epsilon, \epsilon')$. The ϵ-optimum query exponent, denoted $E^*(\epsilon)$, is the supremum of all ϵ-achievable query exponents; $E^*(\epsilon)$ is nondecreasing in ϵ. The *optimum query exponent* E^* is the infimum of $E^*(\epsilon)$ for $0 < \epsilon < 1$, i.e.,

$$E^* = \lim_{\epsilon \to 0} E^*(\epsilon).$$

Remark. Clearly, $0 \leq E^* \leq \log |\mathcal{X}_{\mathcal{M}}|$.

Condition (6.1) forces any query strategy adopted by the querier to have an exponential complexity (in n) with large probability; E^* is the largest value of the exponent that can be inflicted on the querier.

Our main result is a single-letter characterization of the optimum query exponent E^*.

Theorem 6.1. *The optimum query exponent E^* equals*

$$E^* = E^*(\epsilon) = H\left(X_{\mathcal{M}}\right) - \max_{\lambda \in \Lambda(\mathcal{A})} \sum_{B \in \mathcal{B}} \lambda_B H\left(X_B \mid X_{B^c}\right), \qquad 0 < \epsilon < 1. \tag{6.2}$$

Remarkably, the value of E^* coincides with the SK capacity of a multiterminal source model; see (2.12). In fact, the achievability proof of Theorem 6.1 is straightforward and employs, in effect, an SK in forming an appropriate CR L. We show that for such a CR L, any query strategy is tantamount to an exhaustive search over the set of values of the SK, a feature that is apparent for a "perfect" SK with $I(K \wedge \mathbf{F}) = 0$. The difficult step in the proof of Theorem 6.1 is the converse part which involves an appropriate query strategy, for arbitrary L and $\mathbf{F}$, that limits the incurred query exponents. Our *strong converse* yields a uniform upper bound for $E^*(\epsilon)$, $0 < \epsilon < 1$.

We shall see that while the expression for E^* in (6.2) lends itself to the achievability proof of Theorem 6.1 in Section 6.4, alternative forms are suited better for the converse proof. Using the alternative form in (2.14), the expression (6.2) can be written also as

$$E^* = \min_{\lambda \in \Lambda(\mathcal{A})} \left[\sum_{B \in \mathcal{B}} \lambda_B H\left(X_{B^c}\right) - \left(\lambda_{\mathtt{sum}} - 1\right) H\left(X_{\mathcal{M}}\right) \right], \tag{6.3}$$

which is used in the converse proof for an arbitrary $\mathcal{A} \subseteq \mathcal{M}$ in Section 6.6. The converse proof for the case $\mathcal{A} = \mathcal{M}$ is facilitated by the further simplification pointed out in (2.15).

6.3 Technical tools

The following simple observation relates the number of queries in a query strategy q to the cardinality of an associated set.

Proposition 6.2. *Let q be a query strategy for U given $V = v$, $v \in \mathcal{V}$. Then,*

$$|\{u \in \mathcal{U} : q(u|v) \leq \gamma\}| \leq \gamma.$$

Proof. The claim is straightforward since $q(\cdot|v)$ is a bijection. □

For rvs U, V, finding a lower bound for $q(U|V)$ involves finding a suitable upper bound for the conditional probabilities $\mathrm{P}_{U|V}(\cdot \mid \cdot)$. This idea is formalized by the following lemma.

Lemma 6.3. *Given $\gamma > 0$ and $0 < \delta < 1/2$, let the rvs U, V, satisfy*

$$\mathrm{P}\left(\left\{(u,v) : \mathrm{P}_{U|V}(u|v) \leq \frac{\delta}{\gamma}\right\}\right) \geq 1 - \delta. \tag{6.4}$$

Then for every query strategy q for U given V,

$$\mathrm{P}\left(q(U|V) \geq \gamma\right) \geq 1 - \epsilon', \tag{6.5}$$

for all $\epsilon' \geq 2\delta$.

Conversely, if (6.5) holds for every query strategy q for U given V, with $0 < \epsilon' \leq (1 - \sqrt{\delta})^2$, then

$$\mathrm{P}\left(\left\{(u,v) : \mathrm{P}_{U|V}(u|v) \leq \frac{1}{\gamma}\right\}\right) \geq \delta. \tag{6.6}$$

Proof. Suppose (6.4) holds but not (6.5). Then there exists q with

$$\mathrm{P}\left(q(U|V) < \gamma\right) > \epsilon'. \tag{6.7}$$

From (6.4) and (6.7)

$$\begin{aligned}\mathrm{P}\left(\left\{(u,v) : \mathrm{P}_{U|V}(u|v) \leq \frac{\delta}{\gamma},\ q(u|v) < \gamma\right\}\right) &> 1 - \delta + \epsilon' - 1 \\ &= \epsilon' - \delta. \end{aligned} \tag{6.8}$$

On the other hand, the left side of (6.8) equals

$$\begin{aligned}
&\sum_v \mathrm{P}_V(v) \sum_{u: q(u|v)<\gamma,\ \mathrm{P}_{U|V}(u|v)\leq \frac{\delta}{\gamma}} \mathrm{P}_{U|V}(u|v) \\
&\leq \gamma . \frac{\delta}{\gamma}, \qquad \text{by Proposition 6.2} \\
&= \delta,
\end{aligned}$$

which contradicts (6.8) since $\epsilon' \geq 2\delta$.

For the converse, suppose that (6.6) does not hold; then, we show that a query strategy q_0 exists which violates (6.5) when $0 < \epsilon' \leq (1-\sqrt{\delta})^2$. The negation of (6.6) is

$$\mathrm{P}\left(\left\{(u,v) : \mathrm{P}_{U|V}(u|v) > \frac{1}{\gamma}\right\}\right) > 1 - \delta,$$

which, by a *reverse Markov inequality*[1] [41, p. 157] (see also [26, p. 153]), gives a set $\mathcal{V}_0 \subseteq \mathcal{V}$ with

$$\mathrm{P}_V(\mathcal{V}_0) > 1 - \sqrt{\delta}, \tag{6.9}$$

and

$$\mathrm{P}_{U|V}\left(\left\{(u : \mathrm{P}_{U|V}(u|v) > \frac{1}{\gamma}\right\} \,\middle|\, v\right) > 1 - \sqrt{\delta}, \quad v \in \mathcal{V}_0. \tag{6.10}$$

Denoting by $\mathcal{U}_v$ the set $\{\cdot\}$ in (6.10), we have

$$1 \geq \mathrm{P}_{U|V}(\mathcal{U}_v \mid v) > \frac{|\mathcal{U}_v|}{\gamma},$$

[1] The reverse Markov inequality states that for rvs U, V with $\mathrm{P}((U,V) \in S) \geq 1 - \epsilon$ for some $S \subseteq \mathcal{U} \times \mathcal{V}$, there exists $\mathcal{V}_0 \subseteq \mathcal{V}$ such that $\mathrm{P}((U,V) \in S \mid V = v) \geq 1 - \sqrt{\epsilon}$, $v \in \mathcal{V}_0$, and $\mathrm{P}(V \in \mathcal{V}_0) \geq 1 - \sqrt{\epsilon}$.

so that

$$|\mathcal{U}_v| < \gamma, \quad v \in \mathcal{V}_0. \tag{6.11}$$

For each $v \in \mathcal{V}_0$, order the elements of $\mathcal{U}$ arbitrarily but with the first $|\mathcal{U}_v|$ elements being from $\mathcal{U}_v$. This ordering defines a query strategy $q_0(\cdot|v)$, $v \in \mathcal{V}_0$; for $v \notin \mathcal{V}_0$, let $q_0(\cdot|v)$ be defined arbitrarily. Then for $v \in \mathcal{V}_0$, $u \in \mathcal{U}_v$,

$$q_0(u|v) < \gamma$$

by (6.11), so that

$$\begin{aligned} \mathrm{P}\left(q_0(U|V) < \gamma\right) &\geq \sum_{v \in \mathcal{V}_0} \sum_{u \in \mathcal{U}_v} \mathrm{P}_{U,V}\left(u, v\right) \\ &> (1 - \sqrt{\delta})^2, \end{aligned} \tag{6.12}$$

by (6.9) and (6.10). So, $q = q_0$ violates (6.5) when $\epsilon' \leq (1 - \sqrt{\delta})^2$. □

Finally, the following simple observation will be useful.

Proposition 6.4. *For pmfs Q_1, Q_2, on $\mathcal{V}$,*

$$Q_1\left(\{v : Q_1(v) \geq \delta Q_2(v)\}\right) \geq 1 - \delta, \qquad 0 < \delta < 1.$$

Proof. The claim follows from

$$\sum_{v \in \mathcal{V}: Q_1(v) < \delta Q_2(v)} Q_1(v) < \sum_{v \in \mathcal{V}: Q_1(v) < \delta Q_2(v)} \delta\, Q_2(v) \leq \delta.$$

□

6.4 Proof of achievability

Denoting the right-side of (6.2) by C, we claim, for $0 < \epsilon < 1$, $0 < \delta < 1/2$, $\beta > 0$, the existence of an ϵ-CR $L = X_{\mathcal{M}}^n$ for $\mathcal{A}$ from $\mathbf{F}$ with

$$\mathrm{P}\left(\left\{(x_{\mathcal{M}}^n, \mathbf{i}) : \mathrm{P}_{L|\mathbf{F}}\left(x_{\mathcal{M}}^n \mid \mathbf{i}\right) \leq \delta \exp\left[-n(C-\beta)\right]\right\}\right) \geq 1 - \delta, \tag{6.13}$$

for all n sufficiently large. Then the assertion of the theorem follows by applying the first part of Lemma 6.3 with $U = L$, $V = \mathbf{F}, \gamma = \exp[n(C-\beta)]$, to conclude from (6.5) that

$$E^*(\epsilon) \geq C,$$

since $\beta > 0$ was chosen arbitrarily.

Turning to the mentioned claim, it is shown in [20, Proposition 1], [21, Theorem 3.1] that there exists communication $\mathbf{F}$ such that $L = X_{\mathcal{M}}^n$ is ϵ-CR for $\mathcal{A}$ from $\mathbf{F}$ with

$$\frac{1}{n} \log \|\mathbf{F}\| \leq \max_{\lambda \in \Lambda(\mathcal{A})} \sum_{B \in \mathcal{B}} \lambda_B H\left(X_B \mid X_{B^c}\right) + \frac{\beta}{3}, \tag{6.14}$$

for all n sufficiently large. Using Proposition 6.4 with $Q_1 = P_{\mathbf{F}}$ and Q_2 being the uniform pmf over the range of $\mathbf{F}$, we get

$$\mathrm{P}_{\mathbf{F}}\left(\left\{\mathbf{i} : \mathrm{P}_{\mathbf{F}}\left(\mathbf{i}\right) \geq \frac{\delta}{2\|\mathbf{F}\|}\right\}\right) \geq 1 - \frac{\delta}{2}. \tag{6.15}$$

Also, for $x_{\mathcal{M}}^n$ in the set $\mathcal{T}_n$ of $\mathrm{P}_{X_{\mathcal{M}}}$-typical sequences with constant δ [18, Definition 2.8], we have

$$\mathrm{P}_{X_{\mathcal{M}}^n}\left(x_{\mathcal{M}}^n\right) \leq \exp\left[-n\left(H\left(X_{\mathcal{M}}\right) - \frac{\beta}{3}\right)\right] \tag{6.16}$$

and

$$\mathrm{P}_{X_{\mathcal{M}}^n}\left(\mathcal{T}_n\right) \geq 1-\frac{\delta}{2},$$

for all n sufficiently large. Denoting by $\mathcal{I}_0$ the set on the left-side of (6.15), it follows that

$$\mathrm{P}\left(X_{\mathcal{M}}^n \in \mathcal{T}_n, \mathbf{F} \in \mathcal{I}_0\right) \geq 1-\delta. \tag{6.17}$$

The claim results from (6.15)-(6.17) upon observing that for $(x_{\mathcal{M}}^n, \mathbf{i}) \in \mathcal{T}^n \times \mathcal{I}_0$,

$$\begin{aligned}
\mathrm{P}_{X_{\mathcal{M}}^n \mid \mathbf{F}}\left(x_{\mathcal{M}}^n \mid \mathbf{i}\right) &= \frac{\mathrm{P}_{X_{\mathcal{M}}^n}\left(\left(x_{\mathcal{M}}^n\right)\right) \mathbf{1}\left(\mathbf{F}\left(x_{\mathcal{M}}^n\right)=\mathbf{i}\right)}{\mathrm{P}_{\mathbf{F}}\left(\mathbf{i}\right)} \\
&\leq \frac{2 \exp\left[-n\left(H\left(X_{\mathcal{M}}\right)-\frac{\beta}{3}\right)\right] \|\mathbf{F}\|}{\delta} \\
&\leq \delta \exp[-n(C-\beta)],
\end{aligned}$$

for all n large enough, where the last inequality is by (6.14). □

Remark. The achievability proof brings out a connection between a large probability uniform upper bound κ for P_L, the size $\|\mathbf{F}\|$ of the communication $\mathbf{F}$, and the associated number of queries needed. Loosely speaking, the number of queries is approximately $\frac{1}{\|\mathbf{F}\|\kappa}$, which reduces to $\frac{\|L\|}{\|\mathbf{F}\|}$ if L is nearly uniformly distributed.

6.5 Proof of converse for $\mathcal{A} = \mathcal{M}$

Recalling the expression for E^* in (2.15), given a partition π of $\mathcal{M}$ with $|\pi| = k$, $2 \leq k \leq m$, we observe that for a consolidated source model with k sources and underlying rvs $Y_1, ..., Y_k$ where[2] $Y_i = X_{\pi_i}$, the ϵ-optimum query exponent $E_\pi^*(\epsilon)$ can be no smaller than $E^*(\epsilon)$ (since the terminals in each π_i coalesce, in effect).

[2] For specificity, the elements in each π_i are arranged in increasing order.

Theorem 6.5. *For every partition π of $\mathcal{M}$ with $|\pi| = k$,*

$$E^*_\pi(\epsilon) \leq \frac{1}{k-1} D\bigg(\mathrm{P}_{Y_1,...,Y_k} \| \prod_{i=1}^{k} \mathrm{P}_{Y_i}\bigg), \qquad 0 < \epsilon < 1,$$

and so

$$E^*(\epsilon) \leq \min_\pi E^*_\pi(\epsilon) \leq \min_\pi \frac{1}{|\pi| - 1} D\bigg(\mathrm{P}_{X_\mathcal{M}} \| \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi_i}}\bigg).$$

Theorem 6.5 establishes, in view of (2.15), the converse part of Theorem 6.1 when $\mathcal{A} = \mathcal{M}$.

The proof of Theorem 6.5 relies on the following general result, which holds for queries of CR generated in a multiterminal source model with underlying rvs $Y_1, ..., Y_k$ for $n = 1$.

Theorem 6.6. *Let $L = L\left(Y_1, ..., Y_k\right)$ be ϵ-CR for $\{1, ..., k\}$ from interactive communication $\mathbf{F} = \mathbf{F}\left(Y_1, ..., Y_k\right)$, $0 < \epsilon < 1$. Given $\delta > 0$ such that $\delta + \sqrt{\delta + \epsilon} < 1$, let θ be such that*

$$\mathrm{P}\left(\left\{(y_1, ..., y_k) : \frac{\mathrm{P}_{Y_1,...,Y_k}\left(y_1, ..., y_k\right)}{\prod_{i=1}^{k} \mathrm{P}_{Y_i}\left(y_i\right)} \leq \theta\right\}\right) \geq 1 - \delta. \tag{6.18}$$

Then, there exists a query strategy q_0 for L given $\mathbf{F}$ such that

$$\mathrm{P}\left(q_0(L \mid \mathbf{F}) \leq \left(\frac{\theta}{\delta^2}\right)^{\frac{1}{k-1}}\right) \geq (1 - \delta - \sqrt{\delta + \epsilon})^2. \tag{6.19}$$

Proof of Theorem 6.5. We apply Theorem 6.6 to n i.i.d. repetitions of the rvs $Y_1, ..., Y_k$. Denoting by $\mathcal{T}'_n$ the set of $\mathrm{P}_{Y_1,...,Y_k}$-typical sequences with constant δ, we have

$$\mathrm{P}_{Y_1^n,...,Y_k^n}\left(\mathcal{T}'_n\right) \geq 1 - \delta,$$

and for $(y_1^n, ..., y_k^n) \in \mathcal{T}_n'$,

$$\frac{\mathrm{P}_{Y_1^n,...,Y_k^n}\left(y_1^n, ..., y_k^n\right)}{\prod_{i=1}^k \mathrm{P}_{Y_i^n}\left(y_i^n\right)} \leq \exp\left[n\bigg(\sum_{i=1}^k H\left(Y_i\right) - H\left(Y_1, ..., Y_k\right) + \delta\bigg)\right]$$
$$= \exp\left[n\bigg(D\bigg(\mathrm{P}_{Y_1,...,Y_k} \| \prod_{i=1}^k \mathrm{P}_{Y_i}\bigg) + \delta\bigg)\right],$$

for all n large enough. Thus, the hypothesis of Theorem 6.6 holds with

$$\theta = \theta_n = \exp\left[n\bigg(D\bigg(\mathrm{P}_{Y_1,...,Y_k} \| \prod_{i=1}^k \mathrm{P}_{Y_i}\bigg) + \delta\bigg)\right].$$

If E is an ϵ-achievable query exponent (see Definition 6.2), then there exists an ϵ-CR $L = L\left(Y_1^n, ..., Y_k^n\right)$ from communication $\mathbf{F} = \mathbf{F}\left(Y_1^n, ..., Y_k^n\right)$ such that (6.1) holds for the query strategy q_0 of Theorem 6.6 for this choice of L and $\mathbf{F}$. In particular for $\epsilon' < (1 - \delta - \sqrt{\delta + \epsilon})^2$, we get from (6.19) and (6.1) that

$$\mathrm{P}\Bigg(\exp(nE) \leq q_0(L \mid \mathbf{F}) \leq \delta^{-2/(k-1)} \exp\left[n\left(\frac{1}{k-1}D\bigg(\mathrm{P}_{Y_1,...,Y_k} \| \prod_{i=1}^k \mathrm{P}_{Y_i}\bigg) + \frac{\delta}{k-1}\right)\right]\Bigg)$$
$$\geq (1 - \delta - \sqrt{\delta + \epsilon})^2 - \epsilon' > 0, \tag{6.20}$$

for all n sufficiently large. It follows that

$$E \leq \frac{1}{k-1}D\bigg(\mathrm{P}_{Y_1,...,Y_k} \| \prod_{i=1}^k \mathrm{P}_{Y_i}\bigg) + \frac{2\delta}{k-1}.$$

Since E was any ϵ-achievable query exponent and $\delta > 0$ was chosen arbitrarily, the assertion of Theorem 6.5 is established. □

Proof of Theorem 6.6. Denote by $\mathcal{L}$ the set of values of the CR L. Using the hypothesis (6.18) of the Theorem, we shall show below the existence of a set $\mathcal{I}_o$ of values of $\mathbf{F}$ and associated sets $\mathcal{L}(\mathbf{i}) \subseteq \mathcal{L}$, $\mathbf{i} \in \mathcal{I}_0$, such that for every $\mathbf{i} \in \mathcal{I}_0$

$$\mathrm{P}_{L|\mathbf{F}}\left(\mathcal{L}(\mathbf{i}) \mid \mathbf{i}\right) \geq 1 - \delta - \sqrt{\epsilon + \delta}, \tag{6.21}$$

$$|\mathcal{L}(\mathbf{i})| \leq \left(\frac{\theta}{\delta^2}\right)^{\frac{1}{k-1}}, \tag{6.22}$$

$$\text{and} \qquad \mathrm{P}_{\mathbf{F}}\left(\mathcal{I}_0\right) \geq 1 - \delta - \sqrt{\epsilon + \delta}. \tag{6.23}$$

Then, we consider a query strategy q_0 for L given $\mathbf{F}$ as in the proof of converse part of Lemma 6.3, with L, $\mathbf{F}$, $\mathcal{I}_0$, $\mathcal{L}(\mathbf{i})$ in the roles of U, V, $\mathcal{V}_0$, $\mathcal{U}_v$, respectively. Thus, for all $\mathbf{i} \in \mathcal{I}_0$, $l \in \mathcal{L}(\mathbf{i})$,

$$q_0(l \mid \mathbf{i}) \leq |\mathcal{L}(\mathbf{i})| \leq \left(\frac{\theta}{\delta^2}\right)^{\frac{1}{k-1}},$$

and so, as in (6.12), we get by (6.21)-(6.23),

$$\mathrm{P}\left(q_0(L \mid \mathbf{F}) \leq \left(\frac{\theta}{\delta^2}\right)^{\frac{1}{k-1}}\right) \geq (1 - \delta - \sqrt{\delta + \epsilon})^2,$$

thereby establishing the assertion (6.19).

The existence of the sets $\mathcal{I}_0$ and $\{\mathcal{L}(\mathbf{i}), \mathbf{i} \in \mathcal{I}_0\}$ satisfying (6.21)-(6.23) is argued in three steps below.

Step 1. First, we note the following simple property of interactive communication: if rvs $Y_1, ..., Y_k$ are mutually independent, they remain mutually independent when conditioned on an interactive communication $\mathbf{F}$.

Lemma 6.7. *Let the pmf* $\tilde{\mathrm{P}}_{Y_1,...,Y_k}$ *be such that*

$$\tilde{\mathrm{P}}_{Y_1,...,Y_k} = \prod_{j=1}^{k} \tilde{\mathrm{P}}_{Y_j}. \tag{6.24}$$

Then, for $\mathbf{i} = \mathbf{F}\left(y_1, ..., y_k\right)$*, we have*

$$\tilde{\mathrm{P}}_{Y_1,...,Y_k|\mathbf{F}}\left(y_1, ..., y_k \mid \mathbf{i}\right) = \prod_{j=1}^{k} \tilde{\mathrm{P}}_{Y_j|\mathbf{F}}\left(y_j \mid \mathbf{i}\right). \tag{6.25}$$

Proof. The proof follows upon observing that

$$I_{\tilde{\mathrm{P}}}\left(Y_j \wedge Y_1, ..., Y_{j-1}, Y_{j+1}, ..., Y_k \mid \mathbf{F}\right) \leq I_{\tilde{\mathrm{P}}}\left(Y_j \wedge Y_1, ..., Y_{j-1}, Y_{j+1}, ..., Y_k\right) = 0, \quad j = 1, ..., k, \tag{6.26}$$

where the first inequality is by [1, Lemma 2.2] upon choosing $U = Y_j$, $V = (Y_1, ..., Y_{j-1}, Y_{j+1}, ..., Y_k)$, Φ to be the communication from terminal j, and Ψ to be the communication from the remaining terminals.

Hereafter in this proof, we shall select

$$\tilde{\mathrm{P}}_{Y_j} = \mathrm{P}_{Y_j}, \quad j = 1, ..., k. \tag{6.27}$$

Step 2. In this step, we select the aforementioned set of communication values $\mathcal{I}_0$. Let $L_j = L_j\left(Y_j, \mathbf{F}\right)$ denote an estimate of CR L at terminal j, $j = 1, ..., k$ (see Definition 2.2). Denote by $\mathcal{T}_0$ the set $\{\cdot\}$ on the left side of (6.18). For each realization $(l, \mathbf{i})$ of $(L, \mathbf{F})$, denote by $A_{l,\mathbf{i}} \subseteq \mathcal{Y}_1 \times ... \times \mathcal{Y}_k$ the set

$$A_{l,\mathbf{i}} = \mathcal{T}_0 \cap \left\{(y_1, ..., y_k) : \mathbf{F}\left(y_1, ..., y_k\right) = \mathbf{i}, L_j\left(y_j, \mathbf{i}\right) = L\left(y_1, ..., y_k\right) = l, j = 1, ..., k\right\}. \tag{6.28}$$

Since L is ϵ-CR from $\mathbf{F}$, we have from (2.1) and (6.18) that

$$\mathrm{P}\left(\left(Y_1, ..., Y_k\right) \in A_{L,\mathbf{F}}\right) \geq 1 - \epsilon - \delta.$$

By a *reverse Markov inequality*, there exists a set $\mathcal{I}_1$ of values of $\mathbf{F}$ with

$$\mathrm{P}_{\mathbf{F}}\left(\mathcal{I}_1\right) \geq 1 - \sqrt{\epsilon + \delta}, \tag{6.29}$$

and

$$\mathrm{P}\left(\left(Y_1, ..., Y_k\right) \in A_{L,\mathbf{F}} \mid \mathbf{F} = \mathbf{i}\right) \geq 1 - \sqrt{\epsilon + \delta}, \qquad \mathbf{i} \in \mathcal{I}_1. \tag{6.30}$$

Next, denote by $\mathcal{I}_2$ the set of values of $\mathbf{F}$ such that

$$\delta\tilde{\mathrm{P}}_{\mathbf{F}}\left(\mathbf{i}\right) \leq \mathrm{P}_{\mathbf{F}}\left(\mathbf{i}\right), \qquad \mathbf{i} \in \mathcal{I}_2, \tag{6.31}$$

where $\tilde{\mathrm{P}}_{\mathbf{F}}$ is, as usual, the distribution of $\mathbf{F}$ under $\tilde{\mathrm{P}}$. From Proposition 6.4 with $Q_1 = \mathrm{P}_{\mathbf{F}}$, $Q_2 = \tilde{\mathrm{P}}_{\mathbf{F}}$, we have

$$\mathrm{P}_{\mathbf{F}}\left(\mathcal{I}_2\right) \geq 1 - \delta. \tag{6.32}$$

Thus, by (6.29) and (6.32), $\mathcal{I}_0 \triangleq \mathcal{I}_1 \cap \mathcal{I}_2$ satisfies (6.23).

Step 3. In this step, we identify sets $\mathcal{L}(\mathbf{i})$ that satisfy (6.21) and (6.22). For each $\mathbf{i} \in \mathcal{I}_0$, the sets $A_{l,\mathbf{i}}$ corresponding to different values l are disjoint. Upon defining the nonnegative measure[3] μ on $\mathcal{L}$ for each $\mathbf{i} \in \mathcal{I}_0$ by

$$\mu(l) \triangleq \mathrm{P}_{Y_1,...,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right), \qquad l \in \mathcal{L}, \tag{6.33}$$

we get

$$\begin{aligned}\mu(\mathcal{L}) &= \sum_{l \in \mathcal{L}} \mathrm{P}_{Y_1,...,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right) \\ &= \mathrm{P}\left((Y_1, ..., Y_k) \in A_{L,\mathbf{i}} \mid \mathbf{F} = \mathbf{i}\right) \\ &\geq 1 - \sqrt{\epsilon + \delta},\end{aligned}$$

by (6.30). Applying Lemma 2.8 (i) with $\mathcal{L}$ in the role of $\mathcal{U}$, we set $\mathcal{L}(\mathbf{i}) = \mathcal{U}_\delta$, and so

$$\begin{aligned}\mu(\mathcal{L}(\mathbf{i})) &\geq \mu(\mathcal{L}) - \delta \\ &\geq 1 - \delta - \sqrt{\epsilon + \delta}\end{aligned} \tag{6.34}$$

[3] Although μ depends on $\mathbf{i}$, our notation will suppress this dependence.

and

$$|\mathcal{L}(\mathbf{i})| \leq \delta^{-\alpha/(1-\alpha)} \exp\left(H_\alpha(\mu)\right), \quad 0 \leq \alpha < 1. \tag{6.35}$$

It follows from (6.34) that

$$\begin{aligned} \mathrm{P}_{L|\mathbf{F}}\left(\mathcal{L}(\mathbf{i}) \mid \mathbf{i}\right) &\geq \sum_{l \in \mathcal{L}(\mathbf{i})} \mathrm{P}_{Y_1,...,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right) \\ &= \mu(\mathcal{L}(\mathbf{i})) \\ &\geq 1 - \delta - \sqrt{\epsilon + \delta}, \end{aligned} \tag{6.36}$$

which establishes (6.21).

Finally, we obtain an upper bound on $\exp\left(H_\alpha(\mu)\right)$ for $\alpha = \frac{1}{k}$, which will lead to (6.22). Denote by $A^j_{l,\mathbf{i}} \subseteq \mathcal{Y}_j$ the projection of the set $A_{l,\mathbf{i}} \subseteq \mathcal{Y}_1 \times ... \times \mathcal{Y}_k$ along the jth coordinate, $j = 1, ..., k$. The sets $A^j_{l,\mathbf{i}}$ are disjoint for different values of l, by definition (see (6.28)). Thus, for the pmf $\tilde{\mathrm{P}}_{Y_1,..,Y_k}$ in (6.24), (6.27), we have

$$\begin{aligned} 1 &\geq \prod_{j=1}^{k} \left[\sum_{l \in \mathcal{L}} \tilde{\mathrm{P}}_{Y_j|\mathbf{F}}\left(A^j_{l,\mathbf{i}} \mid \mathbf{i}\right)\right] \\ &\geq \left[\sum_{l \in \mathcal{L}} \left(\prod_{j=1}^{k} \tilde{\mathrm{P}}_{Y_j|\mathbf{F}}\left(A^j_{l,\mathbf{i}} \mid \mathbf{i}\right)^{\frac{1}{k}}\right)\right]^k, \end{aligned} \tag{6.37}$$

where the last step follows from *Hölder's inequality*[4] [35, Section 2.7]. Using (6.25), the right-side of (6.37) is the same as

$$\left[\sum_{l \in \mathcal{L}} \tilde{\mathrm{P}}_{Y_1,...,Y_k|\mathbf{F}}\left(A^1_{l,\mathbf{i}} \times ... \times A^k_{l,\mathbf{i}} \mid \mathbf{i}\right)^{\frac{1}{k}}\right]^k,$$

which is bounded below by

$$\left[\sum_{l \in \mathcal{L}} \tilde{\mathrm{P}}_{Y_1,...,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right)^{\frac{1}{k}}\right]^k, \tag{6.38}$$

[4] See [71, equation (33)] for an early use of *Hölder's inequality* in a CR converse proof.

since

$$A_{l,\mathbf{i}} \subseteq A_{l,\mathbf{i}}^1 \times ... \times A_{l,\mathbf{i}}^k. \tag{6.39}$$

Upon noting that $A_{l,\mathbf{i}} \subseteq \mathcal{T}_0$, for all $(y_1, ..., y_k) \in A_{l,\mathbf{i}}$, it follows that

$$\begin{aligned}
\tilde{\mathrm{P}}_{Y_1,...,Y_k|\mathbf{F}}\left(y_1, ..., y_k \mid \mathbf{i}\right) &= \frac{\tilde{\mathrm{P}}_{Y_1,...,Y_k}\left(y_1, ..., y_k\right)}{\tilde{\mathrm{P}}_{\mathbf{F}}\left(\mathbf{i}\right)} \\
&= \frac{\prod_{j=1}^k \tilde{\mathrm{P}}_{Y_j}\left(y_j\right)}{\tilde{\mathrm{P}}_{\mathbf{F}}\left(\mathbf{i}\right)} \\
&= \frac{\prod_{j=1}^k \mathrm{P}_{Y_j}\left(y_j\right)}{\tilde{\mathrm{P}}_{\mathbf{F}}\left(\mathbf{i}\right)} \\
&\geq \frac{\mathrm{P}_{Y_1,...,Y_k}\left(y_1, ..., y_k\right)}{\theta\ \tilde{\mathrm{P}}_{\mathbf{F}}\left(\mathbf{i}\right)} \\
&\geq \frac{\mathrm{P}_{Y_1,...,Y_k|\mathbf{F}}\left(y_1, ..., y_k \mid \mathbf{i}\right)}{\delta^{-1}\ \theta},
\end{aligned}$$

where the third equality and the subsequent inequalities are by (6.27), (6.18) and (6.31), respectively. Combining the observations above with (6.37) and (6.38), we get

$$\begin{aligned}
1 &\geq \left[\sum_{l\in\mathcal{L}} \left(\frac{\mathrm{P}_{Y_1,...,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right)}{\delta^{-1}\ \theta}\right)^{\frac{1}{k}}\right]^k, \\
&= \frac{\delta}{\theta}\left[\sum_{l\in\mathcal{L}} \mu(l)^{\frac{1}{k}}\right]^k,
\end{aligned}$$

which, recalling Definition 2.6, further yields

$$\begin{aligned}
\exp\left(H_{\frac{1}{k}}(\mu)\right) &= \left[\sum_{l\in\mathcal{L}} \mu(l)^{\frac{1}{k}}\right]^{\frac{k}{k-1}} \\
&\leq \left(\frac{\theta}{\delta}\right)^{\frac{1}{k-1}}.
\end{aligned}$$

The previous bound, along with (6.35), gives (6.22). □

6.6 Proof of converse for arbitrary $\mathcal{A} \subseteq \mathcal{M}$

The converse technique of the previous section for $\mathcal{A} = \mathcal{M}$ can be extended to an arbitrary $\mathcal{A} \subseteq \mathcal{M}$, yielding an analogous upper bound for $E^*(\epsilon)$ in terms of divergences. However, the resulting upper bound is inadequate as it is known to exceed the expression in the right-side of (6.3) (see [10]). In this section, we develop a new converse technique that targets directly the latter.

The main steps of the general converse proof for the case $\mathcal{A} \subseteq \mathcal{M}$ are analogous to those in the previous section. The central step is the counterpart of Theorem 6.6, which is given next. Given a *fractional partition* λ as in (2.11), its dual partition is $\overline{\lambda} = \overline{\lambda}(\lambda) = \left\{\overline{\lambda}_{B^c}, B \in \mathcal{B}\right\}$ with

$$\overline{\lambda}_{B^c} = \frac{\lambda_B}{\lambda_{\texttt{sum}} - 1}, \quad B \in \mathcal{B}, \tag{6.40}$$

where $\mathcal{B}$ is defined in (2.10) and $\lambda_{\texttt{sum}}$ is given by (2.13). It is known from [46], and can be seen also from (2.11) and (2.13), that

$$\begin{aligned}\sum_{B \in \mathcal{B}: B^c \ni i} \overline{\lambda}_{B^c} &= \frac{1}{\lambda_{\texttt{sum}} - 1} \sum_{B \in \mathcal{B}: B^c \ni i} \lambda_B \\ &= \frac{1}{\lambda_{\texttt{sum}} - 1} \left[\sum_{B \in \mathcal{B}} \lambda_B - \sum_{B \in \mathcal{B}: B \ni i} \lambda_B\right] \\ &= \frac{1}{\lambda_{\texttt{sum}} - 1} \left[\lambda_{\texttt{sum}} - 1\right] = 1, \quad i \in \mathcal{M},\end{aligned} \tag{6.41}$$

so that $\overline{\lambda}$, too, is a fractional partition of $\mathcal{M}$.

Theorem 6.8. *Let $L = L(Y_1, ..., Y_m)$ be ϵ-CR for $\mathcal{A}$ from interactive communication $\mathbf{F} = \mathbf{F}(Y_1, ..., Y_m)$, $0 < \epsilon < 1$. Given $\delta > 0$ such that $\delta + \sqrt{\delta + \epsilon} < 1$ and a fractional* *partition $\lambda \in \Lambda(\mathcal{A})$, let $\theta_{B^c}, B \in \mathcal{B}$, and θ_0 be such that*

$$\mathrm{P}\left(\left\{y_{\mathcal{M}} : \mathrm{P}_{Y_{\mathcal{M}}}\left(y_{\mathcal{M}}\right) \leq \frac{1}{\theta_0},\ \mathrm{P}_{Y_{B^c}}\left(y_{B^c}\right) \geq \frac{1}{\theta_{B^c}}, B \in \mathcal{B}\right\}\right) \geq 1 - \delta. \tag{6.42}$$

Then, with

$$\theta = \frac{\prod_{B \in \mathcal{B}} \theta_{B^c}^{\bar{\lambda}_{B^c}}}{\theta_0}, \tag{6.43}$$

there exists a query strategy q_0 for L given $\mathbf{F}$ such that

$$\mathrm{P}\left(q_0(L \mid \mathbf{F}) \leq \left(\frac{\theta}{\kappa(\delta)}\right)^{\lambda_{\mathtt{sum}}-1}\right) \geq (1 - \delta - \sqrt{\delta + \epsilon})^2, \tag{6.44}$$

where $\kappa(\delta) = \left(m 2^m\right)^{-m} \delta^{m+1}$.

Proof. As in the proof of Theorem 6.6, the assertion (6.44) will follow upon showing the existence of sets $\mathcal{I}_0$ and $\mathcal{L}(\mathbf{i}) \subseteq \mathcal{L}$, $\mathbf{i} \in \mathcal{I}_0$, such that (6.21) and (6.23) are satisfied, along with the following replacement for (6.22):

$$|\mathcal{L}(\mathbf{i})| \leq \left(\frac{\theta}{\kappa(\delta)}\right)^{\lambda_{\mathtt{sum}}-1}, \qquad \mathbf{i} \in \mathcal{I}_0. \tag{6.45}$$

To this end, we provide here appropriate replacements for the three steps in the proof of Theorem 6.6.

Step 1. For each $B \subsetneq \mathcal{M}$, consider the pmf $\tilde{\mathrm{P}}^B_{Y_{\mathcal{M}}}$ defined by

$$\tilde{\mathrm{P}}^B_{Y_{\mathcal{M}}}\left(y_{\mathcal{M}}\right) = \mathrm{P}_{Y_B}\left(y_B\right) \mathrm{P}_{Y_{B^c}}\left(y_{B^c}\right) \tag{6.46}$$

Note that $\tilde{\mathrm{P}}^B \equiv \tilde{\mathrm{P}}^{B^c}$. The collection of pmfs $\left\{\tilde{\mathrm{P}}^{B^c},\ B \in \mathcal{B}\right\}$ serve as a replacement for the pmf $\tilde{\mathrm{P}}$ in (6.24).

For the pmf $\tilde{\mathrm{P}}^B$ in (6.46), we note that

$$I_{\tilde{\mathrm{P}}^B}\left(Y_B \wedge F_{kj} \mid \Phi_{kj}\right) = 0, \qquad j \in B^c, \tag{6.47}$$

since $F_{kj} = f_{kj}\left(Y_j, \Phi_{kj}\right)$ and Y_{B^c} is independent of Y_B conditioned on Φ_{kj}. The following Lemma serves the role of Lemma 6.7.

Lemma 6.9. *For $B \subsetneq \mathcal{M}$ and $\mathbf{i} = \mathbf{F}\left(y_{\mathcal{M}}\right)$, we have*

$$\tilde{\mathrm{P}}^B_{Y_B|\mathbf{F}}\left(y_B \mid \mathbf{i}\right) = \frac{\mathrm{P}_{Y_B}\left(y_B\right)}{\prod_{k=1}^r \prod_{j\in B} \tilde{\mathrm{P}}^B_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^-_{kj}\right)}, \tag{6.48}$$

where i^-_{kj} denotes the past values of communication in $\mathbf{i}$ for round k and terminal j.

Proof. Note that

$$\begin{aligned}\tilde{\mathrm{P}}^B_{Y_B|\mathbf{F}}\left(y_B \mid \mathbf{i}\right) &= \frac{\tilde{\mathrm{P}}^B_{\mathbf{F}|Y_B}\left(\mathbf{i} \mid y_B\right)\tilde{\mathrm{P}}^B_{Y_B}\left(y_B\right)}{\tilde{\mathrm{P}}^B_{\mathbf{F}}\left(\mathbf{i}\right)} \\ &= \frac{\tilde{\mathrm{P}}^B_{\mathbf{F}|Y_B}\left(\mathbf{i} \mid y_B\right)\mathrm{P}_{Y_B}\left(y_B\right)}{\tilde{\mathrm{P}}^B_{\mathbf{F}}\left(\mathbf{i}\right)},\end{aligned} \tag{6.49}$$

where the previous step is by (6.46). Furthermore,

$$\begin{aligned}\tilde{\mathrm{P}}^B_{\mathbf{F}|Y_B}\left(\mathbf{i} \mid y_B\right) &= \prod_{k=1}^r \prod_{j=1}^m \tilde{\mathrm{P}}^B_{F_{kj}|Y_B,\Phi_{kj}}\left(i_{kj} \mid y_B, i^-_{kj}\right) \\ &= \prod_{k=1}^r \prod_{j\in B^c} \tilde{\mathrm{P}}^B_{F_{kj}|Y_B,\Phi_{kj}}\left(i_{kj} \mid y_B, i^-_{kj}\right) \\ &= \prod_{k=1}^r \prod_{j\in B^c} \tilde{\mathrm{P}}^B_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^-_{kj}\right),\end{aligned} \tag{6.50}$$

where the last step uses (6.47). Next,

$$\begin{aligned}\tilde{\mathrm{P}}^B_{\mathbf{F}}\left(\mathbf{i}\right) &= \prod_{k=1}^r \prod_{j=1}^m \tilde{\mathrm{P}}^B_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^-_{kj}\right) \\ &= \prod_{k=1}^r \left(\prod_{j\in B} \tilde{\mathrm{P}}^B_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^-_{kj}\right) \prod_{j\in B^c} \tilde{\mathrm{P}}^B_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^-_{kj}\right)\right).\end{aligned} \tag{6.51}$$

Then (6.49), along with (6.50) and (6.51), gives (6.48).

Step 2. Denoting by $\mathcal{T}_0$ the set $\{\cdot\}$ on the left-side of (6.42), for each $L = l, \mathbf{F} = \mathbf{i}$, define

$$A_{l,\mathbf{i}} = \mathcal{T}_0 \cap \left\{y_{\mathcal{M}} : \mathbf{F}\left(y_{\mathcal{M}}\right) = \mathbf{i}, L_j\left(y_j, \mathbf{i}\right) = L\left(y_{\mathcal{M}}\right) = l, j \in \mathcal{A}\right\}. \tag{6.52}$$

Analogous to the proof of Theorem 6.6, the set $\mathcal{I}_1$ of values of $\mathbf{F}$ with

$$\mathrm{P}\left(Y_{\mathcal{M}} \in A_{L,\mathbf{F}} \mid \mathbf{F} = \mathbf{i}\right) \geq 1 - \sqrt{\epsilon + \delta}, \qquad \mathbf{i} \in \mathcal{I}_1,$$

satisfies

$$\mathrm{P}_{\mathbf{F}}\left(\mathcal{I}_1\right) \geq 1 - \sqrt{\epsilon + \delta}.$$

For $j \in \mathcal{M}$ and $B \subsetneq \mathcal{M}$, denote by $\mathcal{I}_{j,B}$ the set of $\mathbf{i}$ such that

$$\left(m2^m\right)^{-1} \delta \prod_{k=1}^{r} \tilde{\mathrm{P}}^{B}_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^{-}_{kj}\right) \leq \prod_{k=1}^{r} \mathrm{P}_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^{-}_{kj}\right). \tag{6.53}$$

The following simple extension of Proposition 6.4 holds:

$$\begin{aligned}
\mathrm{P}_{\mathbf{F}}\left(\mathcal{I}^c_{j,B}\right) &= \sum_{\mathbf{i} \in \mathcal{I}^c_{j,B}} \mathrm{P}_{\mathbf{F}}\left(\mathbf{i}\right) \\
&= \sum_{\mathbf{i} \in \mathcal{I}^c_{j,B}} \prod_{l=1}^{m} \prod_{k=1}^{r} \mathrm{P}_{F_{kl}|\Phi_{kl}}\left(i_{kl} \mid i^{-}_{kl}\right) \\
&= \sum_{\mathbf{i} \in \mathcal{I}^c_{j,B}} \left(\prod_{l \neq j} \prod_{k=1}^{r} \mathrm{P}_{F_{kl}|\Phi_{kl}}\left(i_{kl} \mid i^{-}_{kl}\right)\right) \prod_{k=1}^{r} \mathrm{P}_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^{-}_{kj}\right) \\
&< \left(m2^m\right)^{-1} \delta \sum_{\mathbf{i} \in \mathcal{I}^c_{j,B}} \left(\prod_{l \neq j} \prod_{k=1}^{r} \mathrm{P}_{F_{kl}|\Phi_{kl}}\left(i_{kl} \mid i^{-}_{kl}\right)\right) \prod_{k=1}^{r} \tilde{\mathrm{P}}^{B}_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^{-}_{kj}\right) \\
&\leq \left(m2^m\right)^{-1} \delta \sum_{\mathbf{i}} \left(\prod_{l \neq j} \prod_{k=1}^{r} \mathrm{P}_{F_{kl}|\Phi_{kl}}\left(i_{kl} \mid i^{-}_{kl}\right)\right) \prod_{k=1}^{r} \tilde{\mathrm{P}}^{B}_{F_{kj}|\Phi_{kj}}\left(i_{kj} \mid i^{-}_{kj}\right) \\
&= \left(m2^m\right)^{-1} \delta,
\end{aligned} \tag{6.54}$$

where the first inequality is by (6.53), and (6.54) holds since the summand is a pmf for $\mathbf{F}$, as can be seen by directly computing the sum. Defining $\mathcal{I}_2 = \bigcap_{j=1}^{m} \bigcap_{B \subsetneq \mathcal{M}} \mathcal{I}_{j,B}$, we get

$$\mathrm{P}_{\mathbf{F}}\left(\mathcal{I}_2\right) \geq 1 - \delta.$$

The set $\mathcal{I}_0$ is defined as $\mathcal{I}_1 \cap \mathcal{I}_2$, and satisfies (6.23).

Step 3. Finally, we define sets $\mathcal{L}(\mathbf{i}) \subseteq \mathcal{L}$, $\mathbf{i} \in \mathcal{I}_0$ that satisfy (6.21) and (6.45). For each $\mathbf{i} \in \mathcal{I}_0$, let

$$\mu(l) = \mathrm{P}_{Y_{\mathcal{M}}|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right), \qquad l \in \mathcal{L}. \tag{6.55}$$

Then, the sets $\mathcal{L}(\mathbf{i})$ satisfying (6.21) are obtained by an application of Lemma 2.8 (i) as in (6.34) and (6.35) above.

The condition (6.45) will be obtained upon showing that for

$$\alpha = \frac{\lambda_{\mathtt{sum}} - 1}{\lambda_{\mathtt{sum}}}, \tag{6.56}$$

it holds that

$$\delta^{-\alpha/(1-\alpha)} \exp\left(H_\alpha(\mu)\right) \le \left(\frac{\theta}{\kappa(\delta)}\right)^{\lambda_{\mathtt{sum}}-1}. \tag{6.57}$$

To do so, first note that for each $B \in \mathcal{B}$, the set $B^c \cap \mathcal{A}$ is nonempty. Thus, by (6.52), the projections $A_{l,\mathbf{i}}^{B^c}$ of $A_{l,\mathbf{i}}$ along the coordinates in $B^c \subsetneq \mathcal{M}$ are disjoint across $l \in \mathcal{L}$. Thus,

$$1 \ge \prod_{B \in \mathcal{B}} \left(\sum_{l \in \mathcal{L}} \tilde{\mathrm{P}}_{Y_{\mathcal{M}}|\mathbf{F}}^{B^c}\left(A_{l,\mathbf{i}}^{B^c} \mid \mathbf{i}\right)\right)^{\lambda_B}.$$

Using Hölder's inequality [35, Section 2.7], and recalling (6.40) and (2.13) we get

$$1 \ge \left[\sum_{l \in \mathcal{L}} \left(\prod_{B \in \mathcal{B}} \tilde{\mathrm{P}}_{Y_{\mathcal{M}}|\mathbf{F}}^{B^c}\left(A_{l,\mathbf{i}}^{B^c} \mid \mathbf{i}\right)^{\overline{\lambda}_{B^c}}\right)^{\alpha}\right]^{\frac{1}{1-\alpha}}. \tag{6.58}$$

Next, note from Lemma 6.9 that

$$\tilde{\mathrm{P}}_{Y_{\mathcal{M}}|\mathbf{F}}^{B^c}\left(A_{l,\mathbf{i}}^{B^c} \mid \mathbf{i}\right) = \frac{\displaystyle\sum_{y_{B^c} \in A_{l,\mathbf{i}}^{B^c}} \mathrm{P}_{Y_{B^c}}\left(y_{B^c}\right)}{\displaystyle\prod_{k=1}^{r} \prod_{j \in B^c} \tilde{\mathrm{P}}_{F_{kj}|\Phi_{kj}}^{B^c}\left(i_{kj} \mid i_{kj}^{-}\right)},$$

which, since the order of products can be interchanged, and upon using (6.53), is bounded below by

$$\frac{\displaystyle\sum_{y_{B^c}\in A_{l,\mathbf{i}}^{B^c}} \mathrm{P}_{Y_{B^c}}\left(y_{B^c}\right)}{\displaystyle\prod_{j\in B^c}\left(m2^m\right)\delta^{-1}\prod_{k=1}^{r}\left[\mathrm{P}_{F_{kj}|\Phi_{kj}}\left(i_{kj}\mid i_{kj}^{-}\right)\right]}.$$

It follows that

$$\prod_{B\in\mathcal{B}}\tilde{\mathrm{P}}_{Y_{\mathcal{M}}|\mathbf{F}}^{B^c}\left(A_{l,\mathbf{i}}^{B^c}\mid\mathbf{i}\right)^{\overline{\lambda}_{B^c}} \geq \frac{\displaystyle\prod_{B\in\mathcal{B}}\left[\sum_{y_{B^c}\in A_{l,\mathbf{i}}^{B^c}}\mathrm{P}_{Y_{B^c}}\left(y_{B^c}\right)\right]^{\overline{\lambda}_{B^c}}}{\displaystyle\prod_{B\in\mathcal{B}}\prod_{j\in B^c}\left[\left(m2^m\right)\delta^{-1}\prod_{k=1}^{r}\mathrm{P}_{F_{kj}|\Phi_{kj}}\left(i_{kj}\mid i_{kj}^{-}\right)\right]^{\overline{\lambda}_{B^c}}}. \qquad (6.59)$$

The right-side of (6.59) can be simplified by noting that

$$\begin{aligned}
&\prod_{B\in\mathcal{B}}\prod_{j\in B^c}\left[\left(m2^m\right)\delta^{-1}\prod_{k=1}^{r}\mathrm{P}_{F_{kj}|\Phi_{kj}}\left(i_{kj}\mid i_{kj}^{-}\right)\right]^{\overline{\lambda}_{B^c}}\\
&=\prod_{j=1}^{m}\left[\left(m2^m\right)\delta^{-1}\prod_{k=1}^{r}\mathrm{P}_{F_{kj}|\Phi_{kj}}\left(i_{kj}\mid i_{kj}^{-}\right)\right]^{\sum_{B\in\mathcal{B}:B^c\ni j}\overline{\lambda}_{B^c}}\\
&=\left(\frac{m2^m}{\delta}\right)^m\mathrm{P}_{\mathbf{F}}\left(\mathbf{i}\right), \qquad (6.60)
\end{aligned}$$

where the previous step uses (6.41). The definition of $\mathcal{T}_0$, along with (6.59) and (6.60), gives

$$\prod_{B\in\mathcal{B}}\tilde{\mathrm{P}}_{Y_{\mathcal{M}}|\mathbf{F}}^{B^c}\left(A_{l,\mathbf{i}}^{B^c}\mid\mathbf{i}\right)^{\overline{\lambda}_{B^c}} \geq \frac{\delta^m}{\left(m2^m\right)^m\mathrm{P}_{\mathbf{F}}\left(\mathbf{i}\right)}\prod_{B\in\mathcal{B}}\left(\frac{\left|A_{l,\mathbf{i}}^{B^c}\right|}{\theta_{B^c}}\right)^{\overline{\lambda}_{B^c}}. \qquad (6.61)$$

Also, since $A_{l,\mathbf{i}}\subseteq\mathcal{T}_0$, we have

$$\mathrm{P}_{Y_{\mathcal{M}}}\left(A_{l,\mathbf{i}}\right)\leq\frac{\left|A_{l,\mathbf{i}}\right|}{\theta_0},$$

which, with (6.43) and (6.61), gives

$$\prod_{B\in\mathcal{B}}\tilde{\mathrm{P}}_{Y_{\mathcal{M}}|\mathbf{F}}^{B^c}\left(A_{l,\mathbf{i}}^{B^c}\mid\mathbf{i}\right)^{\overline{\lambda}_{B^c}} \geq \frac{\delta^m}{\left(m2^m\right)^m\theta}\left(\frac{\prod_{B\in\mathcal{B}}\left|A_{l,\mathbf{i}}^{B^c}\right|^{\overline{\lambda}_{B^c}}}{\left|A_{l,\mathbf{i}}\right|}\right)\mathrm{P}_{Y_{\mathcal{M}}|\mathbf{F}}\left(A_{l,\mathbf{i}}\mid\mathbf{i}\right). \qquad (6.62)$$

Since $\overline{\lambda}$ is a fractional partition, [45, Corollary 3.4] implies

$$\left(\frac{\prod_{B\in\mathcal{B}}\left|A_{l,\mathbf{i}}^{B^c}\right|^{\overline{\lambda}_{B^c}}}{\left|A_{l,\mathbf{i}}\right|}\right)\geq 1, \tag{6.63}$$

which combined with (6.58)-(6.63) yields

$$1\geq\left(\frac{\delta^m}{\left(m2^m\right)^m\theta}\right)^{\frac{\alpha}{1-\alpha}}\left[\sum_{l\in\mathcal{L}}\mu(l)^\alpha\right]^{\frac{1}{1-\alpha}}.$$

The previous inequality implies (6.57) since

$$\frac{\alpha}{1-\alpha}=\lambda_{\mathtt{sum}}-1.$$

□

6.7 Strong converse for secret key capacity

A byproduct of Theorem 6.1 is a new result that establishes a strong converse for the SK capacity of a multiterminal source model, for the terminals in $\mathcal{A}\subseteq\mathcal{M}$. In this context, we shall consider the security requirement (2.9), i.e.,

$$\lim_n ns_{var}(K;\mathbf{F})=0.$$

Definition 6.3. Given $0<\epsilon<1$, $R\geq 0$ is an ϵ-achievable SK rate for $\mathcal{A}\subseteq\mathcal{M}$ if for every $\rho>0$, there is an $N=N(\epsilon,\rho)$ such that for every $n\geq N$, there exists an ϵ-CR $K=K\left(X_{\mathcal{M}}^n\right)$ for $\mathcal{A}$ from $\mathbf{F}$ satisfying

$$\frac{1}{n}\log\|K\|\geq R-\rho, \tag{6.64}$$

and

$$s_{var}(K;\mathbf{F})\leq\frac{\rho}{n}. \tag{6.65}$$

The supremum of ϵ-achievable SK rates is the ϵ-SK capacity, denoted $C(\epsilon)$. The SK capacity is the infimum of $C(\epsilon)$ for $0 < \epsilon < 1$. We recall the following.

Theorem 6.10. *[20] The secret key capacity for $\mathcal{A} \subseteq \mathcal{M}$ is*

$$C = E^* = H\left(X_{\mathcal{M}}\right) - \max_{\lambda \in \Lambda(\mathcal{A})} \sum_{B \in \mathcal{B}} \lambda_B H\left(X_B \mid X_{B^c}\right), \qquad 0 < \epsilon < 1.$$

Remark. The (new) secrecy requirement (2.9) is not unduly restrictive. Indeed, the achievability proof of Theorem 6.10 [20] holds with $s_{in}(K; \mathbf{F})$ vanishing to zero exponentially rapidly in n, which, by Pinsker's inequality (cf. [18]), implies (2.9). The converse proof in [20] was shown under the weak secrecy condition

$$\lim_n \frac{1}{n} I(K \wedge \mathbf{F}) = 0, \tag{6.66}$$

which, in turn, is implied by (2.9) by a simple application of [20, Lemma 1].

The strong converse for SK capacity, valid under (2.9), is given next.

Theorem 6.11. *For every $0 < \epsilon < 1$, it holds that*

$$C(\epsilon) = C. \tag{6.67}$$

Remark. It is not known if the strong converse in Theorem 6.11 holds under (6.66).

Proof. Theorem 6.10 [20] already provides the proof of achievability, i.e., $C(\epsilon) \geq C$. The converse proof below shows that if R is an ϵ-achievable SK rate, then R is an ϵ-achievable query exponent. Therefore,

$$R \leq E^*(\epsilon) = C, \quad 0 < \epsilon < 1, \tag{6.68}$$

where the equality is by (6.2). Specifically, for every $\rho > 0$, suppose that there exists $K = K\left(X_{\mathcal{M}}^n\right)$ and communication $\mathbf{F}$ satisfying (6.64) and (6.65) for all n sufficiently large. We claim that the hypothesis (6.4) of Lemma 6.3 holds with $U = K$, $V = \mathbf{F}$ and $\gamma = \exp[n(R-2\rho)]$ for every $0 < \delta < 1/2$, when ρ is sufficiently small. Therefore, by (6.5), $R - 2\rho$ is an ϵ-achievable query exponent which leads to (6.68) since ρ can be chosen arbitrarily small.

Turning to the claim, observe that

$$\begin{aligned}
&\mathrm{P}\left(\left\{(k, \mathbf{i}) : \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) > \frac{2}{\exp[n(R-\rho)]}\right\}\right) \\
&\leq \mathrm{P}\left(\left\{(k, \mathbf{i}) : \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) > \frac{2}{\|K\|}\right\}\right) \\
&\leq \mathrm{P}\left(\left\{(k, \mathbf{i}) : \left|\log \|K\| \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right)\right| > 1\right\}\right) \\
&\leq \mathbb{E}\left[\left|\log \|K\| \mathrm{P}_{K|\mathbf{F}}\left(K \mid \mathbf{F}\right)\right|\right],
\end{aligned}$$

where the first and the last inequality above follow from (6.64) and the Markov inequality, respectively.

Next, we show that

$$\mathbb{E}\left[\left|\log \|K\| \mathrm{P}_{K|\mathbf{F}}\left(K \mid \mathbf{F}\right)\right|\right] \leq s_{var}(K; \mathbf{F}) \log \frac{\|K\|^2}{s_{var}(K; \mathbf{F})}. \tag{6.69}$$

Then, the right-side can be bounded above by

$$\frac{\rho}{n} \log \frac{n}{\rho} + 2\rho \log \left|X_{\mathcal{M}}\right|, \tag{6.70}$$

for all n sufficiently large; the claim follows upon taking $n \to \infty$ and $\rho \to 0$. To see (6.69), note that for t_1, t_2, $|t_1 - t_2| < 1$, $f(t) \triangleq -t \log t$ satisfies (cf. [18, Lemma 2.7])

$$\left|f(t_1) - f(t_2)\right| \leq \left|t_1 - t_2\right| \log \frac{1}{\left|t_1 - t_2\right|}. \tag{6.71}$$

Then, for $\mathbf{F} = \mathbf{i}$,

$$\sum_k \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) \left|\log \|K\| \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right)\right|$$

$$= \sum_k \left| \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) \log \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) + \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) \log \|K\| + \frac{1}{\|K\|} \log \|K\| - \frac{1}{\|K\|} \log \|K\| \right|$$

$$\leq \sum_k \left[\left| \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) \log \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) - \frac{1}{\|K\|} \log \frac{1}{\|K\|} \right| + \left| \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) - \frac{1}{\|K\|} \right| \log \|K\| \right]$$

$$\leq \sum_k \left| \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) - \frac{1}{\|K\|} \right| \log \frac{\|K\|}{\left| \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right) - \frac{1}{\|K\|} \right|}, \tag{6.72}$$

where the previous inequality uses (6.71) with $t_1 = \mathrm{P}_{K|\mathbf{F}}\left(k \mid \mathbf{i}\right)$ and $t_2 = \|K\|^{-1}$ for every value k of K. Finally, (6.69) follows upon multiplying both sides by $\mathrm{P}_{\mathbf{F}}\left(\mathbf{i}\right)$, summing over $\mathbf{i}$ and using the log-sum inequality [18]. □

Observe that the proof of Theorem 6.11 does not rely on the form of the rvs $K, \mathbf{F}$, and is, in effect, a statement relating the *size* of any achievable SK rate under the s_{var}-secrecy requirement (2.9) to the query exponent. As a consequence, also the SK capacity for more complex models in which the eavesdropper has additional access to side information can be bounded above by the optimum query exponent when the querier, too, is given access to the same side information.

6.8 General alphabet converse for $\mathcal{A} = \mathcal{M}$

In this section, we present a converse technique for the optimum query exponent for rvs with general alphabets, with jointly Gaussian rvs as a special case. No corresponding general claim is made regarding achievability of the exponent. Our technique also leads to a new strong converse for Gaussian SK capacity [51].

Let $\mathcal{Y}_i$ be a complete separable metric space, with associated Borel σ-field σ_i,

$1 \leq i \leq k$; a special case of interest is $\mathcal{Y}_i = \mathbb{R}^{n_i}$. Denote by $\mathcal{Y}^k$ the set $\mathcal{Y}_1 \times ... \times \mathcal{Y}_k$ and by σ^k the product σ-field[5] $\sigma_1 \times ... \times \sigma_k$ on $\mathcal{Y}^k$. Let $\mathrm{P} = \mathrm{P}_{Y_1,...,Y_k}$ be a probability measure on $\left(\mathcal{Y}^k, \sigma^k\right)$. The interactive communication $\{F_{ji} : 1 \leq j \leq r, 1 \leq i \leq k\}$ is specified as in Definition 2.1, with the rv F_{ji} taking values in, say, $(\mathcal{Z}_{ji}, \mathcal{F}_{ji})$, and being σ_i-measurable for each fixed value of the preceding communication

$$\Phi_{ji} = \left(F_{st} : 1 \leq s < j, 1 \leq t \leq k \text{ or } s = j, 1 \leq t < i\right).$$

Then, there exists a unique regular conditional probability measure on $\left(\mathcal{Y}^k, \sigma^k\right)$ conditioned on $\sigma(\mathbf{F})$, denoted $\mathrm{P}_{Y_1,...,Y_k|\mathbf{F}}$ (cf. [6, Chapter 6]). The notation $Q_\mathbf{i}$ will be used interchangeably for the probability measure $\mathrm{P}_{Y_1,...,Y_k|\mathbf{F}}\left(\cdot \mid \mathbf{i}\right)$. We make the following basic assumption of absolute continuity:

$$Q_\mathbf{i} << \mathrm{P}_{Y_1,...,Y_k}, \qquad \mathrm{P}_\mathbf{F} \text{ a.s. in } \mathbf{i}, \tag{6.73}$$

i.e., (6.73) holds over a set of $\mathbf{i}$ with $\mathrm{P}_\mathbf{F}$-probability 1. Assumption (6.73) is satisfied by a large class of interactive communication protocols including $\mathbf{F}$ taking countably many values. Moreover, we can assume the following without loss of generality:

$$Q_\mathbf{i}\left(\mathbf{F}^{-1}(\mathbf{i})^c\right) = 0, \qquad \mathrm{P}_\mathbf{F} \text{ a.s. in } \mathbf{i}, \tag{6.74}$$

$$\frac{d\,Q_\mathbf{i}}{d\,\mathrm{P}}(y^k) = 0, \quad \text{for } y^k \in \mathbf{F}^{-1}(\mathbf{i})^c, \ \mathrm{P}_\mathbf{F} \text{ a.s. in } \mathbf{i}. \tag{6.75}$$

Next, we define ϵ-CR L from $\mathbf{F}$ and its local estimates L_i, respectively, as rvs *taking countably many values*, measurable with respect to σ^k and $\sigma_i \times \sigma(\mathbf{F})$,

[5] Hereafter, the term "product σ-field" of σ-fields $\sigma_1, ..., \sigma_k$, will mean the smallest σ-field containing sets from $\sigma_1 \times ... \times \sigma_k$, and will be denoted, with an abuse of notation, simply as $\sigma^k = \sigma_1 \times ... \times \sigma_k$.

$1 \leq i \leq k$, and satisfying

$$\mathrm{P}\left(L = L_i,\, 1 \leq i \leq k\right) \geq 1 - \epsilon.$$

The main result of this section, given below, extends Theorem 6.6 to general measures as above.

Theorem 6.12. *For $0 < \epsilon < 1$, let L be ϵ-CR from interactive communication $\mathbf{F}$. Let $\tilde{\mathrm{P}} = \tilde{\mathrm{P}}_{Y_1,...,Y_k}$ be a probability measure on $\left(\mathcal{Y}^k, \sigma^k\right)$ with*

$$\tilde{\mathrm{P}}\left(A_1 \times ... \times A_k\right) = \prod_{i=1}^{k} \mathrm{P}_{Y_i}\left(A_i\right) \qquad A_i \in \sigma_i,\, 1 \leq i \leq k. \tag{6.76}$$

Assuming that $\mathrm{P} << \tilde{\mathrm{P}}$, and given $\delta > 0$ such that $\delta + \sqrt{\delta + \epsilon} < 1$, let θ be such that

$$\mathrm{P}\left(\left\{y^k : \frac{d\,\mathrm{P}}{d\,\tilde{\mathrm{P}}}(y^k) \leq \theta\right\}\right) \geq 1 - \delta. \tag{6.77}$$

Then, there exists a query strategy q_0 for L given $\mathbf{F}$ such that

$$\mathrm{P}\left(q_0(L \mid \mathbf{F}) \leq \left(\frac{\theta}{\delta^2}\right)^{\frac{1}{k-1}}\right) \geq (1 - \delta - \sqrt{\delta + \epsilon})^2. \tag{6.78}$$

The proof of Theorem 6.12 is deferred to the end of this section. At this point, we present its implications for a Gaussian setup. Let $X_i^{(n)}$ be an $\mathbb{R}^n$-valued rv, $i = 1, ..., m$, and let $X_{\mathcal{M}}^{(n)} = \left(X_1^{(n)}, ..., X_m^{(n)}\right)$ be jointly Gaussian $\mathcal{N}(\mathbf{0}, \Sigma^{(n)})$, where $\Sigma^{(n)}$ is a positive definite matrix. We remark that $X_{\mathcal{M}}^{(n)}$ need not be independent or identically distributed across n. The notion of an ϵ-optimum query exponent $E^*(\epsilon)$, $0 < \epsilon < 1$, is exactly as in Definition 6.2, even though the underlying CR now can take countably many values. Also, given a partition π of $\mathcal{M}$ with $|\pi| = k$, $2 \leq k \leq m$, the quantity $E_\pi^*(\epsilon)$ is defined as in Section 6.5.

Proposition 6.13. *For $X_{\mathcal{M}}^{(n)} \sim \mathcal{N}(\mathbf{0}, \Sigma^{(n)})$ with $\Sigma^{(n)}$ being positive definite, it holds that*

$$E^*(\epsilon) \leq \min_{\pi} E_{\pi}^*(\epsilon) \leq \min_{\pi} \frac{1}{2(|\pi|-1)} \limsup_{n} \frac{1}{n} \log \frac{\prod_{i=1}^{|\pi|} \left|\Sigma_{\pi_i}^{(n)}\right|}{|\Sigma^{(n)}|}, \quad 0 < \epsilon < 1,$$

where $\Sigma_{\pi_i}^{(n)}$ is the covariance matrix of $X_{\pi_i}^{(n)}$, $i = 1, ..., |\pi|$, and $|\cdot|$ denotes determinant.

Corollary 6.14. *When $X_{\mathcal{M}}^{(n)}$ is i.i.d. in n with $X_{\mathcal{M}} \sim \mathcal{N}(\mathbf{0}, \Sigma)$,*

$$E^*(\epsilon) \leq \min_{\pi} \frac{1}{2(|\pi|-1)} \log \frac{\prod_{i=1}^{|\pi|} \left|\Sigma_{\pi_i}\right|}{|\Sigma|}, \quad 0 < \epsilon < 1.$$

Proof. Proceeding as in the proof of Theorem 6.5, we apply Theorem 6.12 to the rvs $Y_i = X_{\pi_i}^{(n)}$, $1 \leq i \leq |\pi|$. Specifically, we show that the hypothesis (6.77) is satisfied with

$$\theta = \theta_n = \left(\frac{\prod_{i=1}^{|\pi|} \left|\Sigma_{\pi_i}^{n}\right|}{|\Sigma^{(n)}|} \right)^{1/2} \exp(n\delta), \tag{6.79}$$

where $0 < \delta < 1/2$ is arbitrary. Then, the Proposition follows from the definition of $E^*(\epsilon)$ and (6.79) as in the proof of Theorem 6.5. The Corollary results by a straightforward calculation. It remains to verify that (6.77) holds for θ in (6.79). For $B \subsetneq \mathcal{M}, B \neq \emptyset$, let g_B denote the density of the Gaussian rv $X_B^{(n)}$. From the AEP for Gaussian rvs [15, equation (47)] (see also [8]),

$$\mathrm{P}\left(\left| -\frac{1}{n} \log g_B\left(X_B^{(n)}\right) - \frac{1}{n} h\left(X_B^{(n)}\right) \right| > \tau, \text{ for some } \emptyset \neq B \subseteq \mathcal{M} \right) < 2^m \exp(-c(\tau)n), \quad \tau > 0, \tag{6.80}$$

where h denotes differential entropy and $c(\tau) > 0$ is a positive constant that does not depend on n. Since

$$\frac{d\,\mathrm{P}}{d\,\tilde{\mathrm{P}}} = \frac{g_{\mathcal{M}}}{\prod_{i=1}^{|\pi|} g_{\pi_i}}, \qquad \mathrm{P} \text{ a.s.}$$

and

$$h\left(X_{\mathcal{M}}^{(n)}\right) = \frac{1}{2}\log(2\pi e)^{mn}|\Sigma^{(n)}|, \qquad h\left(X_{\pi_i}^{(n)}\right) = \frac{1}{2}\log(2\pi e)^{|\pi_i|n}\left|\Sigma_{\pi_i}^{(n)}\right|, \qquad 1 \leq i \leq |\pi|,$$

using the upper and lower bounds from (6.80) that hold with significant probability for all n sufficiently large, we get that (6.77) holds with θ as in (6.79), for $0 < \delta < 1/2$. □

As an application of the Corollary above, we establish a new strong converse for SK capacity when the underlying rvs $X_{\mathcal{M}}^{(n)}$ are i.i.d. Gaussian in n; for this model, the SK capacity was established in [51]. The notions of ϵ-achievable SK rate, ϵ-SK capacity $C(\epsilon)$ and SK capacity C are as in Definition 6.3, with condition (6.64) replaced by

$$\texttt{range}(K) = \{1, ..., \lfloor \exp(nR) \rfloor\}, \tag{6.81}$$

which rules out such rvs K as take infinitely many values.

Proposition 6.15. *When $X_{\mathcal{M}}^{(n)}$ is i.i.d. in n with $X_{\mathcal{M}} \sim \mathcal{N}(\mathbf{0}, \Sigma)$,*

$$C(\epsilon) = \min_{\pi} \frac{1}{2(|\pi|-1)} \log \frac{\prod_{i=1}^{|\pi|} \left|\Sigma_{\pi_i}\right|}{|\Sigma|}, \quad 0 < \epsilon < 1. \tag{6.82}$$

Proof. That $C(\epsilon)$ is no smaller than the right-side of (6.82) follows from the achievability proof in [51].

The proof of the reverse inequality is along the lines of the proof of Theorem

6.11 and is obtained upon replacing the upper bound (6.70) by

$$\frac{\rho}{n}\log\frac{n}{\rho}+2\rho R,$$

and noting that Lemma 6.3 can be extended straightforwardly to an arbitrary rv V (with the explicit summations in the proof of that Lemma written as expectations), provided that the rv U is finite-valued. □

Proof of Theorem 6.12. In the manner of the proof of Theorem 6.6, it suffices to identify measurable sets $\mathcal{I}_0$ and $\mathcal{L}(\mathbf{i}) \subseteq \mathcal{L}$, $\mathbf{i} \in \mathcal{I}_0$, such that (6.21)-(6.23) are satisfied. Below we generalize appropriately the steps 1-3 in the proof of Theorem 6.6.

Step 1. The following claim is an extension of Lemma 6.7.

Lemma 6.16. *Given measurable sets $A_i \in \sigma_i$, $1 \leq i \leq k$, for $\tilde{\mathrm{P}}$ in (6.76),*

$$\tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_1 \times \ldots \times A_k \mid \mathbf{i}\right) = \prod_{j=1}^{k} \tilde{\mathrm{P}}_{Y_j|\mathbf{F}}\left(A_j \mid \mathbf{i}\right), \qquad \mathrm{P}_{\mathbf{F}} \text{ a.s. in } \mathbf{i}, \tag{6.83}$$

where $\tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}$ is the regular conditional probability on $(\mathcal{Y}^k, \sigma^k)$ conditioned on $\sigma(\mathbf{F})$.

The proof uses the interactive property of the communication and is relegated to the Appendix F.

Step 2. Next, we identify the set $\mathcal{I}_0$. The following technical observation will be used.

Lemma 6.17. *For every $A_0 \in \sigma^k$ such that*

$$\frac{d\,\mathrm{P}}{d\,\tilde{\mathrm{P}}}(y^k) > 0, \qquad y^k \in A_0, \tag{6.84}$$

it holds that

$$\tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_0 \mid \mathbf{i}\right) = \frac{d\,\mathrm{P}_{\mathbf{F}}}{d\,\tilde{\mathrm{P}}_{\mathbf{F}}}(\mathbf{i}) \int_{A_0} \frac{d\,Q_{\mathbf{i}}}{d\,\mathrm{P}}\, d\,\tilde{\mathrm{P}}, \qquad \tilde{\mathrm{P}}_{\mathbf{F}} \textit{ a.s. in } \mathbf{i} \tag{6.85}$$

The proof is given in the Appendix F. Denoting by $\mathcal{T}_0$ the set $\left\{y^k \in \mathcal{Y}^k : 0 < \frac{d\,\mathrm{P}}{d\,\tilde{\mathrm{P}}}(y^k) \leq \theta\right\}$, let

$$A_l = \mathcal{T}_0 \cap \left\{y^k : L_j\big(y_j, \mathbf{F}(y^k)\big) = L(y^k) = l, 1 \leq j \leq k\right\}, \qquad l \in \mathcal{L}.$$

Then, for $A_{l,\mathbf{i}} \triangleq A_l \cap \mathbf{F}^{-1}(\mathbf{i})$, (6.74), (6.75) and Lemma 6.17 imply

$$\tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right) = \frac{d\,\mathrm{P}_{\mathbf{F}}}{d\,\tilde{\mathrm{P}}_{\mathbf{F}}}(\mathbf{i}) \int_{A_{l,\mathbf{i}}} \frac{d\,Q_{\mathbf{i}}}{d\,\mathrm{P}}\, d\,\tilde{\mathrm{P}}, \qquad \tilde{\mathrm{P}}_{\mathbf{F}} \text{ a.s. in } \mathbf{i}. \tag{6.86}$$

Below we restrict attention to the set of values of $\mathbf{F}$ for which (6.86) holds for every $l \in \mathcal{L}$; this set has $\tilde{\mathrm{P}}_{\mathbf{F}}$ measure 1 by (6.85) since the set $\mathcal{L}$ is countable. Proceeding along the lines of the proof of Theorem 6.6, we define $\mathcal{I}_1$ as the set of those $\mathbf{i}$ for which

$$\mathrm{P}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right) \geq 1 - \sqrt{\epsilon + \delta}. \tag{6.87}$$

Since L is an ϵ-CR from $\mathbf{F}$, it follows from (6.77), the fact that

$$\mathrm{P}\left(\left\{y^k : \frac{d\,\mathrm{P}}{d\,\tilde{\mathrm{P}}}(y^k) = 0\right\}\right) = 0,$$

and by a *reverse Markov inequality*, that

$$\mathrm{P}_{\mathbf{F}}\left(\mathcal{I}_1\right) \geq 1 - \sqrt{\epsilon + \delta}. \tag{6.88}$$

Furthermore, for the set $\mathcal{I}_2$ of values $\mathbf{i}$ of $\mathbf{F}$ satisfying

$$\frac{d\,\mathrm{P}_{\mathbf{F}}}{d\,\tilde{\mathrm{P}}_{\mathbf{F}}}(\mathbf{i}) \geq \delta, \tag{6.89}$$

it holds that

$$\mathrm{P}_{\mathbf{F}}\left(\mathcal{I}_2\right) \geq 1-\delta, \tag{6.90}$$

since

$$\begin{aligned}\int_{\mathcal{I}_2^c} d\,\mathrm{P}_{\mathbf{F}} &= \int_{\mathcal{I}_2^c} \frac{d\,\mathrm{P}_{\mathbf{F}}}{d\,\tilde{\mathrm{P}}_{\mathbf{F}}} d\,\tilde{\mathrm{P}}_{\mathbf{F}} \\ &< \delta.\end{aligned}$$

Define $\mathcal{I}_0 = \mathcal{I}_1 \cap \mathcal{I}_2$; (6.23) follows from (6.88) and (6.90).

Step 3. Since Lemma 2.8 (i) applies to a countable set $\mathcal{U} = \mathcal{L}$, defining the non-negative measure μ on $\mathcal{L}$ as in (6.33) for each $\mathbf{i} \in \mathcal{I}_0$ and using (6.87), the sets $\mathcal{L}(\mathbf{i})$ obtained in (6.34)-(6.36) satisfy (6.21). Also, condition (6.22) will follow from (6.35) upon showing that

$$\exp\left(H_\alpha(\mu)\right) \leq \left(\frac{\theta}{\delta}\right)^{\frac{1}{k-1}}. \tag{6.91}$$

To do so, denote by $A_{l,\mathbf{i}}^j$ the projection of $A_{l,\mathbf{i}}$ along the jth coordinate, $1 \leq j \leq k$. As before, the sets $A_{l,\mathbf{i}}^j$ are disjoint across $l \in \mathcal{L}$. Then, *Hölder's inequality* [35] implies that

$$\begin{aligned}1 &\geq \prod_{j=1}^{k}\left[\sum_{l\in\mathcal{L}} \tilde{\mathrm{P}}_{Y_j|\mathbf{F}}\left(A_{l,\mathbf{i}}^j \mid \mathbf{i}\right)\right] \\ &\geq \left[\sum_{l\in\mathcal{L}}\left(\prod_{j=1}^{k} \tilde{\mathrm{P}}_{Y_j|\mathbf{F}}\left(A_{l,\mathbf{i}}^j \mid \mathbf{i}\right)^{\frac{1}{k}}\right)\right]^k \\ &= \left[\sum_{l\in\mathcal{L}} \tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}}^1 \times \ldots \times A_{l,\mathbf{i}}^k \mid \mathbf{i}\right)^{\frac{1}{k}}\right]^k,\end{aligned} \tag{6.92}$$

where the previous step uses Lemma 6.16. The right-side of (6.92) is bounded below by

$$\left[\sum_{l\in\mathcal{L}} \tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right)^{\frac{1}{k}}\right]^k,$$

since $A_{l,\mathbf{i}} \subseteq A^1_{l,\mathbf{i}} \times \ldots \times A^k_{l,\mathbf{i}}$, which by (6.86) equals

$$\left[\sum_{l\in\mathcal{L}} \left(\frac{d\,\mathrm{P}_{\mathbf{F}}}{d\,\tilde{\mathrm{P}}_{\mathbf{F}}}(\mathbf{i}) \int_{A_{l,\mathbf{i}}} \frac{d\,Q_{\mathbf{i}}}{d\,\mathrm{P}} d\,\tilde{\mathrm{P}}\right)^{\frac{1}{k}}\right]^k.$$

From the definition of the set $\mathcal{I}_2$ in (6.89), the expression above exceeds

$$\left[\sum_{l\in\mathcal{L}} \left(\delta \int_{A_{l,\mathbf{i}}} \frac{d\,Q_{\mathbf{i}}}{d\,\mathrm{P}} d\,\tilde{\mathrm{P}}\right)^{\frac{1}{k}}\right]^k,$$

which is the same as

$$\left[\sum_{l\in\mathcal{L}} \left(\delta \int_{A_{l,\mathbf{i}}} \frac{d\,Q_{\mathbf{i}}}{d\,\mathrm{P}} \frac{d\,\mathrm{P}/d\,\tilde{\mathrm{P}}}{d\,\mathrm{P}/d\,\tilde{\mathrm{P}}} d\,\tilde{\mathrm{P}}\right)^{\frac{1}{k}}\right]^k. \tag{6.93}$$

Since $A_{l,\mathbf{i}} \subseteq \mathcal{T}_0$, the sum in (6.93) is bounded below further by

$$\begin{aligned}
&\left[\sum_{l\in\mathcal{L}} \left(\frac{\delta}{\theta} \int_{A_{l,\mathbf{i}}} \frac{d\,Q_{\mathbf{i}}}{d\,\mathrm{P}} \frac{d\,\mathrm{P}}{d\,\tilde{\mathrm{P}}} d\,\tilde{\mathrm{P}}\right)^{\frac{1}{k}}\right]^k \\
&= \frac{\delta}{\theta}\left[\sum_{l\in\mathcal{L}} \left(\int_{A_{l,\mathbf{i}}} d\,Q_{\mathbf{i}}\right)^{\frac{1}{k}}\right]^k \\
&= \frac{\delta}{\theta}\left[\sum_{l\in\mathcal{L}} \mathrm{P}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right)^{\frac{1}{k}}\right]^k.
\end{aligned}$$

Combining the observations above from (6.92) onward, we have

$$\frac{\theta}{\delta} \geq \left[\sum_{l\in\mathcal{L}} \mathrm{P}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_{l,\mathbf{i}} \mid \mathbf{i}\right)^{\frac{1}{k}}\right]^k,$$

which is the same as (6.91) with $\alpha = 1/k$. □

6.9 Discussion

The description of the optimum query exponent in Definition 6.2 can be refined to display an explicit dependence on ϵ'. Let $E^*(\epsilon, \epsilon')$ denote the optimum query exponent for fixed $0 < \epsilon, \epsilon' < 1$. Our proofs establish $E^*(\epsilon, \epsilon')$ equals the right side of (6.2) for $\epsilon' < (1 - \sqrt{\epsilon})^2$ (see (6.20)). For $\epsilon' > 1 - \epsilon$, the following construction of L renders $E^*(\epsilon, \epsilon')$ unbounded: Choose $L = 0$ with probability $(1 - \epsilon)$ and uniformly distributed on a sufficiently large set with probability ϵ. For the remaining values of ϵ, ϵ', $E^*(\epsilon, \epsilon')$ is not known.

A less restrictive model for querying than that in Section 6.2 can be considered, allowing general queries with binary answers. Such a query strategy can be represented as a search on a binary tree whose leaves correspond to the values of the CR L. The query strategies considered in this dissertation correspond to the case where the search tree is a path with leaves attached to each node. For a general tree model, our results can be adapted to show that the maximum number of queries that can be inflicted on a querier grows only linearly in n at a rate that is equal to the expression for E^* in (6.2).

We remark also that allowing randomness at the terminals in $\mathcal{M}$ for interactive communication and CR recovery, does not improve the optimum query exponent. Such randomization is described by mutually independent rvs $W_1, ..., W_m$, where each W_i is distributed uniformly on the (finite) set $\{1, ..., w_i\}$, and the rvs $W_1, ..., W_m$ are independent of $X_{\mathcal{M}}^n$. The claim of the remark is seen from the converse result in Theorem 6.8. Indeed, the assertion (6.44) of Theorem 6.8 remains unchanged upon replacing Y_i by (Y_i, W_i), $i \in \mathcal{M}$, θ_0 by $\theta_0\left(\prod_{i\in\mathcal{M}} w_i\right)$, and θ_{B^c} by $\theta_{B^c}\left(\prod_{i\in B^c} w_i\right)$, $B \in \mathcal{B}$; and observing that in (6.43), the w_i- terms cancel in the numerator and the denominator.

Finally, Lemma 6.3, which considered rvs U, V, can be used to characterize the optimum query exponent Γ^* for a family of finite-valued rvs $\{U_n, V_n\}_{n=1}^{\infty}$ with associated probability measures $\{\mathrm{P}_n\}_{n=1}^{\infty}$ (which are not necessarily consistent). Here, Γ^* is described analogously as E^* in Definition 6.2. An application of Lemma 6.3 yields that

$$\mathrm{P}_n\text{-}\liminf_n \frac{-\log \mathrm{P}_{U_n|V_n}\left(U_n \mid V_n\right)}{n} \leq \Gamma^* \leq \mathrm{P}_n\text{-}\limsup_n \frac{-\log \mathrm{P}_{U_n|V_n}\left(U_n \mid V_n\right)}{n}$$

where the left and right limits equal, respectively, the left- and right-sides of (2.33) with $\mu_n = \mathrm{P}_n$ and

$$Z_n = \frac{-\log \mathrm{P}_{U_n|V_n}\left(U_n \mid V_n\right)}{n}.$$

CHAPTER 7

Conclusion: Principles of Secrecy Generation

Would I not have to be a barrel of memory if

I wanted to carry my reasons around with me?

- Nietzsche's Zarathustra

We conclude this dissertation by hypothesizing three principles of multiterminal secrecy generation that have emerged from our research. The results reported in the previous chapters provide important instances that affirmatively support these principles. We conjecture that these basic principles go beyond the models studied here and will apply in a broader setting for appropriately defined notions of security.

The first principle we state applies to secure computing with public communication when the privacy of a given function g_0 must be maintained.

Principle 1. *If the value of a function g_0 of the data can be recovered securely at a terminal, then the entire data can be recovered at the terminal using a communication that does not give away the value of g_0.*

If the entire data can be recovered at a terminal using communication that does not give away the value of g_0 then clearly the function g_0 can be securely computed at that terminal. The principle above claims that the converse is also true. Indeed, once the value of g_0 is recovered at the terminal, the terminals can communicate further to attain omniscience at that terminal using communication that is independent jointly of the previous communication and the function g_0. Specifically, for the cases when we have a single-letter characterization for secure computability, the quantity R^* in Theorem 4.1 remains unchanged if the terminal computing the function g_0 is required to recover the entire data; this includes the special case studied in Chapter 3. In fact, R^* remains roughly the same with such a replacement for all the cases studied in Chapter 4, thus providing credence to the conjecture that computing the private function securely at a terminal is as hard as recovering the entire data at that terminal without giving away the value of the private function to the eavesdropper.

The next principle captures the structure of all protocols that can generate an optimum rate SKs for two terminals by characterizing the CR that is established when such a protocol is executed.

Principle 2. *A CR corresponds to an optimum rate SK if and only if it renders the observations of the two terminals (almost) conditionally independent.*

Clearly, a CR resulting from an optimum rate SK must render the observations almost conditionally independent as otherwise we can exploit the residual correlation to enhance further the SK rate. In Chapter 5 we have established the converse statement for i.i.d. observations[1]. We conjecture that in fact, such a structural equivalence between optimum rate SKs and a CR that renders the observations conditionally independent holds for more general distributions.

[1] Although our actual proof entailed going to multiple blocks, an alternative single-shot proof can also be provided using a very interesting construction suggested by Braverman and Rao in [9, Theorem 4.1] in place of Lemma 5.4.

Finally, our last principle conjectures that the almost independence secrecy criterion imposed on an SK is equivalent to an alternative criterion requiring a large number of queries with probability close to 1 for resolving the value of the SK based on the public communication (see 6.1).

Principle 3. *Imposing an almost independence secrecy criterion is equivalent to imposing a lower bound on the complexity of a querier of the secret.*

Thus, a largest size SK makes the task of a querier the most onerous. It is clear that almost independence between a CR and the associated public communication will ensure that a querier has no option but to retort to exhaustive search of the CR space. It is the converse that we assert here, i.e, forming a CR that necessitates a certain number of queries is tantamount to generating an SK of size equal to the number of queries. In Chapter 6, we proved this principle for i.i.d. observations in an asymptotic sense as the number of observations tends to infinity. We conjecture that a direct correspondence can be obtained between almost independence and lower bounds on the complexity of a querier by connecting both these notions to the bounds on conditional probability in Lemma 6.3; this is work in progress.

Appendix A

Maximum Common Function

We prove Lemma 3.2 based on [26, Lemma, Section 4], which is paraphrased first. Let the rvs Q and R take values in the finite set $\mathcal{Q}$ and $\mathcal{R}$, respectively. For a stochastic matrix $W : \mathcal{Q} \to \mathcal{Q}$, let $\{\tilde{\mathcal{D}}_1, ..., \tilde{\mathcal{D}}_l\}$ be the ergodic decomposition (into communicating classes) of $\mathcal{Q}$ based on W (cf. e.g., [41, pp. 157 and 28–42]). Let $\tilde{\mathcal{D}}^{(n)}$ denote a fixed ergodic class of $\mathcal{Q}^n$ (the n-fold Cartesian product of $\mathcal{Q}$) on the basis of W^n (the n-fold product of W). Let $\mathcal{D}^{(n)}$ and $\mathcal{R}^{(n)}$ be any (nonempty) subsets of $\tilde{\mathcal{D}}^{(n)}$ and $\mathcal{R}^n$, respectively.

Lemma GK. [26] For $\tilde{\mathcal{D}}^{(n)}, \mathcal{D}^{(n)}, \mathcal{R}^{(n)}$ as above, assume that

$$\begin{aligned} \mathrm{P}\left(Q^n \in \mathcal{D}^{(n)} \mid R^n \in \mathcal{R}^{(n)}\right) &\geq \exp[-n\epsilon_n], \\ \mathrm{P}\left(R^n \in \mathcal{R}^{(n)} \mid Q^n \in \mathcal{D}^{(n)}\right) &\geq \exp[-n\epsilon_n], \end{aligned} \tag{A.1}$$

where $\lim_n \epsilon_n = 0$. Then (as stated in [26, bottom of p. 157]),

$$\frac{\mathrm{P}\left(Q^n \in \mathcal{D}^{(n)}\right)}{\mathrm{P}\left(Q^n \in \tilde{\mathcal{D}}^{(n)}\right)} \geq \exp[-n\kappa\epsilon_n \log^2 \epsilon_n], \tag{A.2}$$

for a (positive) constant κ that depends only on the pmf of (Q, R) and on W.

A simple consequence of (A.2) is that for a given ergodic class $\tilde{\mathcal{D}}^{(n)}$ and disjoint subsets $\mathcal{D}_1^{(n)}, ..., \mathcal{D}_t^{(n)}$ of it, and subsets $\mathcal{R}_1^{(n)}, ..., \mathcal{R}_t^{(n)}$ (not necessarily distinct) of $\mathcal{R}^n$, such that $\mathcal{D}_{t'}^{(n)}, \mathcal{R}_{t'}^{(n)}, t' = 1, ..., t$, satisfy (A.1), it holds that

$$t \leq \exp[n\kappa\epsilon_n \log^2 \epsilon_n]. \tag{A.3}$$

Note that the ergodic decomposition of Q^n on the basis of W^n for the specific choice

$$W(q|q') = \sum_{r \in \mathcal{R}} \mathrm{P}\left(Q = q \mid R = r\right) \mathrm{P}\left(R = r \mid Q = q'\right), \quad q, q' \in \mathcal{Q}$$

corresponds to the set of values of $\mathrm{mcf}^n(Q, R)$ defined by (3.10) [26]. Next, pick $Q = Q_m$, $R = (Q_1, ..., Q_{m-1})$, and define the stochastic matrix $W : \mathcal{Q} \to \mathcal{Q}$ by

$$\begin{aligned} W(q|q') = \sum_{\alpha} \mathrm{P}\left(Q = q \mid \mathrm{mcf}(Q_1, ..., Q_{m-1}) = \alpha\right) \times \\ \mathrm{P}\left(\mathrm{mcf}(Q_1, ..., Q_{m-1}) = \alpha \mid Q = q'\right), \quad q, q' \in \mathcal{Q}. \end{aligned} \tag{A.4}$$

The ergodic decomposition of $\mathcal{Q}^n$ on the basis of W^n (with W as in (A.4)) will correspond to the set of values of $\mathrm{mcf}^n(Q_1, ..., Q_m)$, recalling (3.9). Since $\xi^{(n)}$ is ϵ-recoverable from $Q_i^n, i = 1, ..., m$, note that

$$\xi'^{(n)} = \left(\xi^{(n)}, \mathrm{mcf}^n(Q_1, ..., Q_m)\right)$$

also is ϵ-recoverable in the same sense, recalling Definition 3.5. This implies the existence of mappings $\xi_i'^{(n)}, i = 1, ..., m$, satisfying

$$\mathrm{P}\left(\xi_1'^{(n)}(Q_1^n) = ... = \xi_m'^{(n)}(Q_m^n) = \xi'^{(n)}\right) \geq 1 - \epsilon. \tag{A.5}$$

For each fixed value $c = (c_1, c_2)$ of $\xi'^{(n)}$, let

$$\begin{aligned} \mathcal{D}_c^{(n)} &= \left\{q_m^n \in \mathcal{Q}_m^n : \xi_m'^{(n)}(q_m^n) = c\right\}, \\ \mathcal{R}_c^{(n)} &= \left\{(q_1^n, ..., q_{m-1}^n) \in \mathcal{Q}_1^n \times ... \times \mathcal{Q}_{m-1}^n : \xi_i'^{(n)}(q_i^n) = c, i = 1, ..., m-1\right\}. \end{aligned}$$

Let $C(\epsilon)$ denote the set of c's such that

$$\begin{aligned} &\mathrm{P}\left(Q^n \in \mathcal{D}_c^{(n)} \mid R^n \in \mathcal{R}_c^{(n)}\right) \geq 1 - \sqrt{\epsilon}, \\ &\mathrm{P}\left(R^n \in \mathcal{R}_c^{(n)} \mid Q^n \in \mathcal{D}_c^{(n)}\right) \geq 1 - \sqrt{\epsilon}. \end{aligned} \tag{A.6}$$

Then, as in [26, Proposition 1], it follows from (A.5) that

$$\mathrm{P}\left(\xi'^{(n)} \in C(\epsilon)\right) \geq 1 - 4\sqrt{\epsilon}. \tag{A.7}$$

Next, we observe for each fixed c_2, that the disjoint sets $\mathcal{D}_{c_1,c_2}^{(n)}$ lie in a fixed ergodic class of $\mathcal{Q}^n$ (determined by c_2). Since (A.6) are compatible with the assumption (A.1) for all n sufficiently large, we have from (A.3) that

$$\|\{c_1 : (c_1, c_2) \in C(\epsilon)\}\| \leq \exp[n\kappa\epsilon_n \log^2 \epsilon_n], \tag{A.8}$$

where κ depends on the pmf of $(Q_1, ..., Q_m)$ and W in (A.4), and where $\lim_n \epsilon_n = 0$. Finally,

$$\begin{aligned} &\frac{1}{n}H\left(\xi'^{(n)}\right) = \frac{1}{n}H\left(\xi^{(n)}, \mathrm{mcf}^n(Q_1, ..., Q_m\right) \\ &\leq H\left(\mathrm{mcf}(Q_1, ..., Q_m)\right) + \frac{1}{n}H\left(\xi^{(n)}, \mathbf{1}\left(\xi'^{(n)} \in C(\epsilon)\right) \mid \mathrm{mcf}^n(Q_1, ..., Q_m)\right) \\ &\leq H\left(\mathrm{mcf}(Q_1, ..., Q_m)\right) + \frac{1}{n}H\left(\xi^{(n)} \mid \mathrm{mcf}^n(Q_1, ..., Q_m), \mathbf{1}\left(\xi'^{(n)} \in C(\epsilon)\right)\right) \\ &\leq H(\mathrm{mcf}(Q_1, ..., Q_m)) + \delta_n, \end{aligned}$$

where $\lim_n \delta_n = 0$ by (A.7) and (A.8). □

Appendix B

Aided Secret Key Capacity

We prove Theorem 3.13. Considering first the achievability part, fix $\delta > 0$. From the result for a general source network [17, Theorem 3.1.14] it follows, as in the proof of [20, Proposition 1], that for $R_{\mathcal{M}} \in \mathcal{R}\left(\mathcal{A}', Z_{\mathcal{A}'}\right)$ and all n sufficiently large, there exists a noninteractive communication $\mathbf{F}^{(\mathbf{n})} = (F_1^{(n)}, ..., F_m^{(n)})$ with

$$\frac{1}{n} \log \|\mathbf{F}^{(n)}\| \leq \sum_{i=1}^{m} R_i + \delta,$$

such that $\mathcal{X}_{\mathcal{M}}^n$ is ϵ_n-recoverable from $\left(X_i^n, Z_i^n, \mathbf{F}^{(\mathbf{n})}\right), i \in \mathcal{A}'$. Therefore,

$$\left\{\text{mcf}\left(\left(X_{\mathcal{M}t}, Z_{it}\right)_{i \in A'}\right)\right\}_{t=1}^{n}$$

is ϵ_n-recoverable from $\left(X_i^n, Z_i^n, \mathbf{F}^{(\mathbf{n})}\right), i \in \mathcal{A}'$. The last step takes recourse to Lemma 2.7. Specifically, choose

$$\begin{aligned} &U = U' = \left\{\text{mcf}\left(\left(X_{\mathcal{M}t}, \quad Z_{it}\right)_{i \in A'}\right)\right\}_{t=1}^{n}, \quad \mathcal{U}_0 = \mathcal{U}, \quad V = \text{constant}, \\ &h = F^{(n)}, \quad d = n\left[H\left(\text{mcf}\left(\left(X_{\mathcal{M}}, Z_i\right)_{i \in A'}\right)\right) - \delta\right], \end{aligned}$$

whereby the hypothesis (2.23) of Lemma 2.7 is satisfied for all n sufficiently large. Fixing

$$r' = \left\lceil \exp\left[n\left(\sum_{i=1}^{m} R_i + \delta\right)\right]\right\rceil,$$

Lemma 2.7 implies the existence of a ϕ, and thereby an ASK

$$K^{(n)} = \phi\left(\{\mathrm{mcf}\left((X_{\mathcal{M}t}, Z_{it})_{i\in A'}\right)\}_{t=1}^{n}\right)$$

of rate

$$\frac{1}{n}\log r = H\left(\mathrm{mcf}\left((X_{\mathcal{M}}, Z_i)_{i\in A'}\right)\right) - \sum_{i=1}^{m} R_i - 3\delta.$$

In particular, we can choose

$$\sum_{i=1}^{m} R_i \leq R_{CO}\left(\mathcal{A}'; Z_{\mathcal{A}'}\right) + \frac{\delta}{2}.$$

Since δ was arbitrary, this establishes the achievability part.

We shall establish a stronger converse result by requiring the ASK as in Definition 3.3 to satisfy the weaker secrecy condition (2.6), or by allowing the ASK to depend explicitly on the randomization $U_{\mathcal{M}}$ but enforcing the strong secrecy condition (2.2). Let $K = K^{(n)}\left(U_{\mathcal{M}}, X_{\mathcal{M}}^n, Z_{\mathcal{A}'}^n\right)$ be an ϵ_n-ASK for $\mathcal{A}'$, achievable using observations of length n, randomization $U_{\mathcal{M}}$, public communication $\mathbf{F} = \mathbf{F}\left(U_{\mathcal{M}}, X_{\mathcal{M}}^n\right)$ and side information $Z_{\mathcal{A}'}^n$. Then, by (2.6),

$$\frac{1}{n}H(K) \leq \frac{1}{n}H(K \mid \mathbf{F}) + \epsilon_n. \tag{B.1}$$

Let $K_u = K\left(u, X_{\mathcal{M}}^n, Z_{\mathcal{A}'}^n\right)$ denote the random value of the ASK for a fixed $U_{\mathcal{M}} = u$. Since $(X_{\mathcal{M}}^n, K)$ is ϵ_n-recoverable from the rvs $(U_{\mathcal{M}}, X_{\mathcal{M}}^n, Z_i^n)$ for each $i \in \mathcal{A}'$,

$$\mathrm{P}_{U_{\mathcal{M}}}\left(\{u : \left(X_{\mathcal{M}}^n, K_u\right) \text{ is} \sqrt{\epsilon_n}\text{-recoverable from } \left(U_{\mathcal{M}} = u, X_{\mathcal{M}}^n, Z_i^n\right) \text{ for each } i \in \mathcal{A}'\}\right) \geq 1 - \sqrt{\epsilon_n}. \tag{B.2}$$

Also, for each $U_{\mathcal{M}} = u$

$$\frac{1}{n}H\left(X_{\mathcal{M}}^n, K \mid U_{\mathcal{M}} = u\right) = \frac{1}{n}H\left(X_{\mathcal{M}}^n, K_u\right)$$

by independence of $U_{\mathcal{M}}$ and $(X_{\mathcal{M}}^n, Z_{\mathcal{A}'}^n)$, and therefore, by Lemma 3.2, for u in the set in (B.2),

$$\frac{1}{n} H\left(X_{\mathcal{M}}^n, K \mid U_{\mathcal{M}} = u\right) \leq H\left(\mathrm{mcf}\left((X_{\mathcal{M}}, Z_i)_{i \in \mathcal{A}'}\right)\right) + \delta_n, \tag{B.3}$$

for all n sufficiently large and where $\lim_n \delta_n = 0$. Then,

$$\begin{aligned} &\frac{1}{n} H(U_{\mathcal{M}}, X_{\mathcal{M}}^n, K) \\ &\leq \frac{1}{n} H\left(U_{\mathcal{M}}\right) + H\left(\mathrm{mcf}\left((X_{\mathcal{M}}, Z_i)_{i \in \mathcal{A}'}\right)\right) + \delta_n + \sqrt{\epsilon_n} \log\left(|\mathcal{X}_{\mathcal{M}}||\mathcal{Z}_{\mathcal{A}'}|\right), \end{aligned} \tag{B.4}$$

by (B.2) and (B.3). The proof is now completed along the lines of [20, Lemma 2 and Theorem 3]. Specifically, denoting the set of positive integers $\{1, ..., l\}$ by $[1, l]$,

$$\frac{1}{n} H(U_{\mathcal{M}}, X_{\mathcal{M}}^n, K) = \frac{1}{n} H(K \mid \mathbf{F}) + \sum_{i=1}^{m} R_i' + \frac{1}{n} H(U_{\mathcal{M}}),$$

where

$$R_i' = \frac{1}{n} \sum_{\nu : \nu \equiv i \,\texttt{mod}\, m} H(F_\nu \mid F_{[1, \nu-1]}) + \frac{1}{n} H\left(U_i, X_i^n \mid \mathbf{F}, K, U_{[1,i-1]}, X_{[1,i-1]}^n\right) - H(U_i). \tag{B.5}$$

Consider $B \nsubseteq \mathcal{M}$, $\mathcal{A}' \nsubseteq B$. For $j \in \mathcal{A}' \cap B^c$, we have

$$\begin{aligned} \frac{1}{n} H\left(U_B\right) + \frac{1}{n} H\left(X_B \mid X_{B^c}^n, Z_j^n\right) &= \frac{1}{n} H\left(U_B, X_B^n \mid U_{B^c}, X_{B^c}^n, Z_j^n\right) \\ &= \frac{1}{n} H\left(F_1, ..., F_{rm}, K, U_B, X_B^n \mid U_{B^c}, X_{B^c}^n, Z_j^n\right). \end{aligned}$$

Furthermore, since K is ϵ_n-recoverable from $(\mathbf{F}, U_{B^c}, X_{B^c}^n, Z_j^n)$ and $H(F_\nu | U_{B^c}, X_{B^c}^n) = 0$ for $\nu \equiv i \mod m$ with $i \in B^c$,

$$\frac{1}{n} H\left(F_1, ..., F_{rm}, K, U_B, X_B^n \mid U_{B^c}, X_{B^c}^n, Z_j^n\right)$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{\nu=1}^{rm} H\left(F_\nu \mid F_{[1,\nu-1]}, U_{B^c}, X^n_{B^c}, Z^n_j\right) + \frac{1}{n} H\left(K \mid U_{B^c}, X^n_{B^c}, Z^n_j, \mathbf{F}\right) \\
&\quad + \frac{1}{n}\sum_{i\in B} H\big(U_i, X_i^n \mid U_{B^c\cap[i+1,m]}, X^n_{B^c\cap[i+1,m]}, Z^n_j, \mathbf{F}, K, U_{[1,i-1]}, X^n_{[1,i-1]}\big) \\
&\le \frac{1}{n}\sum_{i\in B}\Bigg[\sum_{\nu:\nu\equiv i \,\texttt{mod}\, m} H\left(F_\nu \mid F_{[1,\nu-1]}\right) + H\left(U_i, X_i^n \mid \mathbf{F}, K, U_{[1,i-1]}, X^n_{[1,i-1]}\right)\Bigg] \\
&\quad + \frac{\epsilon_n \log|\mathcal{K}| + 1}{n} \\
&\le \sum_{i\in B} R_i + H(U_B), \qquad \text{(B.6)}
\end{aligned}
$$

where

$$
R_i \triangleq \left(R_i' + \frac{\epsilon_n \log|\mathcal{K}| + 1}{n}\right), \quad i \in \mathcal{M}.
$$

It follows from (B.1) and (B.4)-(B.6) that

$$
\begin{aligned}
\frac{1}{n}H(K) \le & H\left(\mathrm{mcf}\left((X_{\mathcal{M}}, Z_i)_{i\in\mathcal{A}'}\right)\right) - \sum_{i=1}^{m} R_i + \epsilon_n + \delta_n \\
& + \frac{\epsilon_n \log|\mathcal{K}| + 1}{n} + \sqrt{\epsilon_n}\log\left(|\mathcal{X}_{\mathcal{M}}||\mathcal{Z}_{\mathcal{A}'}|\right), \qquad \text{(B.7)}
\end{aligned}
$$

where $R_{\mathcal{M}} \in \mathcal{R}\left(\mathcal{A}', Z_{\mathcal{A}'}\right)$ from (B.6), and therefore

$$
\sum_{i=1}^{m} R_i \ge R_{CO}\left(\mathcal{A}', Z_{\mathcal{A}'}\right). \qquad \text{(B.8)}
$$

Then, (B.7), (B.8) imply

$$
\frac{1}{n}H(K) \le C\left(\mathcal{A}', Z_{\mathcal{A}'}\right) + \epsilon_n + \delta_n + \frac{\epsilon_n \log|\mathcal{K}| + 1}{n} + \sqrt{\epsilon_n}\log\left(|\mathcal{X}_{\mathcal{M}}||\mathcal{Z}_{\mathcal{A}'}|\right). \qquad \text{(B.9)}
$$

If $K = K\left(X^n_{\mathcal{M}}, Z^n_{\mathcal{A}'}\right)$ as in Definition 3.3, then $|\mathcal{K}| \le \left(|\mathcal{X}||\mathcal{Z}_{\mathcal{A}'}|\right)^n$ and the converse part follows from (B.9). On the other hand, for $K = K\left(U_{\mathcal{M}}, X^n_{\mathcal{M}}, Z^n_{\mathcal{A}'}\right)$, the proof is completed using (2.2) in the manner of [20, Theorem 3]. This completes the converse part. □

Appendix C

Balanced Coloring Lemma

We provide a proof of Lemma 2.7. Using the condition (i) in the definition of $\mathcal{U}_0$, the left-side of (2.24) is bounded above by

$$2\lambda^2 + \sum_{j=1}^{r'} \sum_{v \in \mathcal{V}} \mathrm{P}\left(h(U) = j, V = v, U \in \mathcal{U}_0\right) \times \\ \sum_{i=1}^{r} \left| \sum_{\substack{u' \in \mathcal{U}': \\ \phi(u') = i}} \mathrm{P}\left(U' = u' \mid h(U) = j, V = v, U \in \mathcal{U}_0\right) - \frac{1}{r} \right|.$$

Therefore, it is sufficient to prove that

$$\sum_{j=1}^{r'} \sum_{v \in \mathcal{V}} \mathrm{P}\left(h(U) = j, V = v, U \in \mathcal{U}_0\right) \\ \sum_{i=1}^{r} \left| \sum_{\substack{u' \in \mathcal{U}': \\ \phi(u') = i}} \mathrm{P}\left(U' = u' \mid h(U) = j, V = v, U \in \mathcal{U}_0\right) - \frac{1}{r} \right| < 12\lambda, \tag{C.1}$$

with probability greater than $1 - 2rr'|\mathcal{V}| \exp\left(-\frac{c\lambda^3 d}{rr'}\right)$ for a constant $c > 0$.

Let

$$q = \mathrm{P}_V\left(\left\{v \in \mathcal{V} : \mathrm{P}\left(U \in \mathcal{U}_0 | V = v\right) < \frac{1 - \lambda^2}{3}\right\}\right).$$

Then, since

$$\begin{aligned} 1-\lambda^2 &\leq \mathrm{P}\left(U \in \mathcal{U}_0\right) \\ &\leq \sum_{\substack{v\in V: \\ \mathrm{P}(U\in\mathcal{U}_0|V=v)<\frac{1-\lambda^2}{3}}} \mathrm{P}\left(U \in \mathcal{U}_0 | V=v\right)\mathrm{P}_V\left(v\right) + (1-q) \\ &< \frac{1-\lambda^2}{3}q + (1-q), \end{aligned}$$

we get from the extremities above that

$$q < \frac{3\lambda^2}{2}. \tag{C.2}$$

For $u \in \mathcal{U}_0$ and $v \in \mathcal{V}$ satisfying

$$\begin{aligned} \mathrm{P}\left(U \in \mathcal{U}_0 | V=v\right) &\geq \frac{1-\lambda^2}{3}, \\ \mathrm{P}\left(U = u | V=v, U \in \mathcal{U}_0\right) &> \frac{3}{d(1-\lambda^2)}, \end{aligned}$$

we have that

$$\mathrm{P}\left(U=u|V=v\right) > \frac{1}{d}.$$

Therefore, by (C.2) and (2.23), it follows that

$$\sum_{\substack{(u,v): \\ u\in\mathcal{U}_0,\, \mathrm{P}(U=u|V=v,U\in\mathcal{U}_0)>\frac{3}{d(1-\lambda^2)}}} \mathrm{P}\left(U=u, V=v\right) \leq \lambda^2 + q < \frac{5\lambda^2}{2},$$

which is the same as

$$\begin{aligned} &\sum_{j=1}^{r'}\sum_{v\in\mathcal{V}} \mathrm{P}\left(h(U)=j, V=v, U\in\mathcal{U}_0\right) \times \\ &\qquad \sum_{\substack{u\in\mathcal{U}_0: \\ \mathrm{P}(U=u|V=v,U\in\mathcal{U}_0)>\frac{3}{d(1-\lambda^2)}}} \mathrm{P}\left(U=u|h(U)=j, V=v, U\in\mathcal{U}_0\right) < \frac{5\lambda^2}{2}. \end{aligned} \tag{C.3}$$

The bound in (C.3) will now play the role of [20, inequality (50), p. 3059] and the remaining steps of our proof, which are parallel to those in [20, Lemma B.2], are provided here for completeness.

Setting D to be the set of those (j,v) that satisfy

$$\sum_{\substack{u\in\mathcal{U}:\\ \mathrm{P}(U=u|V=v,U\in\mathcal{U}_0)>\frac{3}{d(1-\lambda^2)}}} \mathrm{P}\left(U=u|h(U)=j,V=v,U\in\mathcal{U}_0\right) \leq \frac{5\lambda}{2},$$

we get that

$$\sum_{(j,v)\in D^c} \mathrm{P}\left(h(U)=j,V=v,U\in\mathcal{U}_0\right) < \lambda. \tag{C.4}$$

Next, defining

$$E=\left\{(j,v):\ \mathrm{P}\left(h(U)=j,V=v,U\in\mathcal{U}_0\right) \qquad \geq \frac{\lambda}{r'}\mathrm{P}\left(V=v,U\in\mathcal{U}_0\right)\right\},$$

it holds for $(j,v)\in E$,

$$\mathrm{P}\left(U=u|h(U)=j,V=v,U\in\mathcal{U}_0\right) \leq \frac{r'}{\lambda}\mathrm{P}\left(U=u|V=v,U\in\mathcal{U}_0\right). \tag{C.5}$$

Also,

$$\begin{aligned}\sum_{(j,v)\in E^c} \mathrm{P}\left(h(U)=j,V=v,U\in\mathcal{U}_0\right) &< \frac{\lambda}{r'}\sum_{j=1}^{r'}\sum_{v\in\mathcal{V}}\mathrm{P}\left(V=v,U\in\mathcal{U}_0\right)\\ &\leq \lambda. \end{aligned} \tag{C.6}$$

Further, for $(j,v)\in E$, if

$$\mathrm{P}\left(U=u|h(U)=j,V=v,U\in\mathcal{U}_0\right) > \frac{3r'}{\lambda d(1-\lambda^2)} \tag{C.7}$$

then from (C.5), we have

$$\mathrm{P}\left(U=u|V=v,U\in\mathcal{U}_0\right) > \frac{3}{d(1-\lambda^2)}. \tag{C.8}$$

Therefore, denoting by $\mathcal{I}(j,v)$ the event $\{h(U) = j, V = v, U \in \mathcal{U}_0\}$, and recalling the conditions that define $\mathcal{U}_0$ in (2.22), we have for $(j,v) \in E \cap D$ that

$$\begin{aligned}
\sum_{\substack{u' \in \mathcal{U}': \\ \mathrm{P}(U'=u'|\mathcal{I}(j,v)) > \frac{3r'}{\lambda d(1-\lambda^2)}}} \mathrm{P}\left(U' = u' | \mathcal{I}(j,v)\right) &= \sum_{\substack{u' \in \mathcal{U}': \\ \mathrm{P}(U=u(u')|\mathcal{I}(j,v)) > \frac{3r'}{\lambda d(1-\lambda^2)}}} \mathrm{P}\left(U = u(u') | \mathcal{I}(j,v)\right) \\
&= \sum_{\substack{u \in \mathcal{U}: \\ \mathrm{P}(U=u|\mathcal{I}(j,v)) > \frac{3r'}{\lambda d(1-\lambda^2)}}} \mathrm{P}\left(U = u | \mathcal{I}(j,v)\right) \\
&\leq \frac{5\lambda}{2}, \qquad \text{(C.9)}
\end{aligned}$$

where the first equality is by (2.22), the second equality is due to U' being a function of U, and the previous inequality is by (C.7), (C.8) and the definition of the set D. Also, using (C.4), (C.6), we get

$$\sum_{(j,v) \in E \cap D} \mathrm{P}\left(h(U) = j, V = v, U \in \mathcal{U}_0\right) \geq 1 - 2\lambda. \qquad \text{(C.10)}$$

Now, the left-side of (C.1) is bounded, using (C.10), as

$$\begin{aligned}
&\sum_{j=1}^{r'} \sum_{v \in \mathcal{V}} \mathrm{P}\left(h(U) = j, V = v, U \in \mathcal{U}_0\right) \times \\
&\quad \sum_{i=1}^{r} \left| \sum_{\substack{u' \in \mathcal{U}': \\ \phi(u')=i}} \mathrm{P}\left(U' = u' \mid h(U) = j, V = v, U \in \mathcal{U}_0\right) - \frac{1}{r} \right| \\
&\leq 4\lambda + \sum_{(j,v) \in E \cap D} \mathrm{P}\left(h(U) = j, V = v, U \in \mathcal{U}_0\right) \times \\
&\qquad \sum_{i=1}^{r} \left| \sum_{\substack{u' \in \mathcal{U}': \\ \phi(u')=i}} \mathrm{P}\left(U' = u' \mid h(U) = j, V = v, U \in \mathcal{U}_0\right) - \frac{1}{r} \right|. \qquad \text{(C.11)}
\end{aligned}$$

Using (C.9), the family of pmfs $\{\mathrm{P}\left(U' = (\cdot) | h(U) = j, V = v, U \in \mathcal{U}_0\right), (j,v) \in E \cap D\}$ satisfies the hypothesis (2.20) of Lemma 2.6 with d replaced by $\frac{\lambda(1-\lambda^2)d}{3r'}$ and ϵ replaced by $5\lambda/2$; assume that $0 < \lambda < 2/45$ so as to meet the condition following

(2.20). This family consists of at most $r'|\mathcal{V}|$ pmfs. Therefore, using Lemma 2.6,

$$\sum_{j=1}^{r'}\sum_{v\in\mathcal{V}} \mathrm{P}\left(h(U)=j, V=v, U\in\mathcal{U}_0\right)\times$$
$$\sum_{i=1}^{r}\left|\sum_{\substack{u'\in\mathcal{U}':\\ \phi(u')=i}} \mathrm{P}\left(U'=u' \mid h(U)=j, V=v, U\in\mathcal{U}_0\right)-\frac{1}{r}\right| < \frac{23\lambda}{2}$$

with probability greater than

$$1-2rr'|\mathcal{V}|\exp\left(-\frac{25\lambda^3(1-\lambda^2)d}{36rr'}\right)\geq 1-2rr'|\mathcal{V}|\exp\left(-\frac{c\lambda^3 d}{rr'}\right),$$

for a constant c. This completes the proof of (C.1), and thereby the lemma. □

Appendix D

Sufficiency Proof of Theorem 4.1

From (4.25), we have

$$nR_{\mathcal{M}} + \frac{\delta}{2} < H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right),$$

where $R_1, ..., R_m$ satisfy conditions (1a) and (1b). For each i and $R_i \geq 0$, consider a (map-valued) rv J_i that is uniformly distributed on the family $\mathcal{J}_i$ of all mappings $\mathcal{X}_i^{nk} \to \{1, \ldots, \lceil \exp(knR_i) \rceil\}$, $i \in \mathcal{M}$. The rvs $J_1, ..., J_m, X_{\mathcal{M}}^{nk}$ are taken to be mutually independent.

Fix ϵ, ϵ', with $\epsilon' > m\epsilon$ and $\epsilon + \epsilon' < 1$. It follows from the proof of the general source network coding theorem [18, Lemma 13.13 and Theorem 13.14] that for all sufficiently large k,

$$\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from } \left(X_i^{nk}, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right), Z_i^k\right), i \in \mathcal{M}\right\}\right)$$
$$\geq 1 - \epsilon, \qquad \text{(D.1)}$$

where, for $i \in \mathcal{M}$,

$$Z_i^k = \begin{cases} \mathbf{F}^k, & j \in [1, m_0], \\ \left(\mathbf{F}^k, G_0^{nk}\right), & m_0 < j \leq m. \end{cases}$$

Below we shall establish that

$$\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : \frac{1}{nk} I\left(j_{\mathcal{M}}(X_{\mathcal{M}}^{nk}) \wedge G_0^{nk}, \mathbf{F}^k\right) \geq \epsilon\right\}\right) \leq \epsilon', \tag{D.2}$$

for all k sufficiently large, to which end it suffices to show that

$$\begin{aligned} &\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : \qquad \frac{1}{nk} I\left(j_i(X_i^{nk}) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right) \geq \frac{\epsilon}{m}\right\}\right) \\ &\leq \frac{\epsilon'}{m}, \quad i \in \mathcal{M}, \end{aligned} \tag{D.3}$$

since

$$\begin{aligned} I\left(j_{\mathcal{M}}\left(X_{\mathcal{M}}^{nk}\right) \wedge G_0^{nk}, \mathbf{F}^k\right) &= \sum_{i=1}^m I\left(j_i\left(X_i^{nk}\right) \wedge G_0^{nk}, \mathbf{F}^k \mid j_1\left(X_1^{nk}\right), \ldots, j_{i-1}\left(X_{i-1}^{nk}\right)\right) \\ &\leq \sum_{i=1}^m I\left(j_i\left(X_i^{nk}\right) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right). \end{aligned}$$

Then it would follow from (D.1), (D.2), and definition of $Z_{\mathcal{M}}$ that

$$\begin{aligned} \mathrm{P}\Bigg(\Bigg\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from } \left(X_i^{nk}, Z_i^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right), i \in \mathcal{M}, \\ \text{and } \frac{1}{nk} I\left(j_{\mathcal{M}}(X_{\mathcal{M}}^{nk}) \wedge G_0^{nk}, \mathbf{F}^k\right) < \epsilon\Bigg\}\Bigg) \geq 1 - \epsilon - \epsilon'. \end{aligned}$$

This shows the existence of a particular realization $\mathbf{F}'$ of $J_{\mathcal{M}}$ that satisfies (4.26) and (4.27).

It now remains to prove (D.3). Defining

$$\tilde{\mathcal{J}}_i = \left\{j_{\mathcal{M}\setminus\{i\}} \in \mathcal{J}_{\mathcal{M}\setminus\{i\}} : X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from } \left(X_i^{nk}, Z_i^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right), \right)\right\},$$

we have by (D.1) that $\mathrm{P}\left(J_{\mathcal{M}\setminus\{i\}} \in \tilde{\mathcal{J}}_i\right) \geq 1 - \epsilon$. It follows that

$$\begin{aligned} &\mathrm{P}\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : \frac{1}{nk} I\left(j_i(X_i^{nk}) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right) \geq \frac{\epsilon}{m}\right\}\right) \\ &\leq \epsilon + \sum_{j_{\mathcal{M}\setminus\{i\}} \in \tilde{\mathcal{J}}_i} \mathrm{P}\left(J_{\mathcal{M}\setminus\{i\}} = j_{\mathcal{M}\setminus\{i\}}\right) p\left(j_{\mathcal{M}\setminus\{i\}}\right), \end{aligned}$$

since J_i is independent of $J_{\mathcal{M}\backslash\{i\}}$, where $p\left(j_{\mathcal{M}\backslash\{i\}}\right)$ is defined as

$$\mathrm{P}\left(\left\{j_i \in \mathcal{J}_i : \quad \frac{1}{nk} I\left(j_i(X_i^{nk}) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\backslash\{i\}}\left(X_{\mathcal{M}\backslash\{i\}}^{nk}\right)\right) \geq \frac{\epsilon}{m}\right\}\right).$$

Thus, (D.3) will follow upon showing that

$$p\left(j_{\mathcal{M}\backslash\{i\}}\right) \leq \frac{\epsilon'}{m} - \epsilon, \quad j_{\mathcal{M}\backslash\{i\}} \in \tilde{\mathcal{J}}_i, \tag{D.4}$$

for all k sufficiently large. Fix $j_{\mathcal{M}\backslash\{i\}} \in \tilde{\mathcal{J}}_i$. We take recourse to Lemma 2.7 and set

$$U = X_{\mathcal{M}}^{nk}, \quad U' = X_i^{nk}, \quad V = \left(G_0^{nk}, \quad \mathbf{F}^k\right), \quad h = j_{\mathcal{M}\backslash\{i\}}, \text{ and}$$
$$\mathcal{U}_0 = \left\{x_{\mathcal{M}}^{nk} \in \mathcal{X}_{\mathcal{M}}^{nk} : x_{\mathcal{M}}^{nk} = \psi_i\left(x_i^{nk}, j_{\mathcal{M}\backslash\{i\}}\left(x_{\mathcal{M}\backslash\{i\}}^{nk}\right), \mathbf{F}^k\left(x_{\mathcal{M}}^{nk}\right), g_0^n\left(x_{\mathcal{M}}^n\right)\mathbf{1}\left(m_0 < i \leq m\right)\right)\right\}$$

for some mapping ψ_i. By the definition of $\tilde{\mathcal{J}}_i$,

$$\mathrm{P}\left(U \in \mathcal{U}_0\right) \geq 1 - \epsilon,$$

so that condition (2.22)(i) preceding Lemma 2.7 is met. Condition (2.22)(ii), too, is met from the definition of $\mathcal{U}_0, h$ and V.

Upon choosing

$$d = \exp\left[k\left(H\left(X_{\mathcal{M}}^n | G_0^n, \mathbf{F}\right) - \frac{\delta}{2}\right)\right],$$

in (2.23), the hypotheses of Lemma 2.7 are satisfied, for appropriately chosen λ, and for sufficiently large k. Then, by Lemma 2.7, with

$$r = \lceil \exp\left(knR_i\right) \rceil, \quad r' = \left\lceil \exp\left(knR_{\mathcal{M}\backslash i}\right) \right\rceil,$$

and with J_i in the role of ϕ, (D.4) follows from (2.24) and (2.25). □

Appendix E

Proof of (5.28) and (5.31)

It remains to prove that there exists ϵ-CR J, recoverable from $\mathbf{F}$ such that $J, \mathbf{F}$ satisfy (5.28) and (5.31). We provide a CR generation scheme with r stages. For $1 \leq k \leq r$, denote by $\mathcal{E}_k$ the error event in the kth stage (defined below recursively in terms of $\mathcal{E}_{k-1}$), and by $\mathcal{E}_0$ the negligible probability event corresponding to X_1^n, X_2^n not being $P_{X_1X_2}$-typical.

Consider $1 \leq k \leq r$, k odd. For brevity, denote by V the rvs U^{k-1} and by U the rv U_k; for $k = 1$, V is taken to be a constant. Suppose that conditioned on $\mathcal{E}_{k-1}^c$ terminals 1 and 2 observe, respectively, sequences $\mathbf{x}_1 \in \mathcal{X}_1^n$ and $\mathbf{x}_2 \in \mathcal{X}_2^n$, as well as a common sequence $\mathbf{v} \in \mathcal{V}^n$ such that $(\mathbf{v}, \mathbf{x}_1, \mathbf{x}_2)$ are jointly $P_{VX_1X_2}$-typical. For $\delta > 0$, generate at random $\exp\left[n(I(X_1, X_2 \wedge U \mid V) + \delta)\right]$ sequences $\mathbf{u} \in \mathcal{U}^n$ that are jointly P_{UV}-typical with $\mathbf{v}$, denoted by $\mathbf{u}_{ij}$, $1 \leq i \leq N_1$, $1 \leq j \leq N_2$, where

$$N_1 = \exp\left[n\left(I(X_1 \wedge U \mid X_2, V) + 3\delta\right)\right], \qquad N_2 = \exp\left[n\left(I(X_2 \wedge U \mid V) - 2\delta\right)\right].$$

The sequences $\mathbf{u}_{ij}$ are generated independently for different indices ij. Denote by $L^{(k)}(\mathbf{v}, \mathbf{x}_1)$ a sequence $\mathbf{u}_{ij}$, $1 \leq i \leq N_1$, $1 \leq j \leq N_2$, that is jointly P_{UVX_1}-typical with $(\mathbf{v}, \mathbf{x}_1)$ (if there exist more than one such sequences, choose any of them). The error event when no such sequence is found is denoted by $\mathcal{E}_{k1}$; this happens with probability vanishing to 0 doubly exponentially in n. The communication $F_k(\mathbf{v}, \mathbf{x}_1)$ is defined to equal the first index i of $\mathbf{u}_{ij} = L^{(k)}(\mathbf{v}, \mathbf{x}_1)$. Upon observing $F_k(\mathbf{v}, \mathbf{x}_1) = i$, the terminal 2 computes $L_2^{(k)}(\mathbf{v}, \mathbf{x}_2, i)$ as the unique sequence in $\{\mathbf{u}_{ij}, 1 \leq j \leq N_2\}$, that is jointly typical with $(\mathbf{v}, \mathbf{x}_2)$. If no such sequence is found or if several such sequences are found an error event $\mathcal{E}_{k2}$ occurs. Clearly, the rate of communication F_k is bounded above by

$$\frac{1}{n}\log N_1 = I(X_1 \wedge U \mid X_2, V) + 3\delta = I(X_1 \wedge U_k \mid X_2, U^{k-1}) + 3\delta, \tag{E.1}$$

and also, for large n,

$$\begin{aligned}\frac{1}{n}H(L^{(k)}) &\leq \frac{1}{n}\log(1 + N_1N_2) \leq I(X_1, X_2 \wedge U \mid V) + 2\delta \\ &= I(X_1, X_2 \wedge U_k \mid X_2, U^{k-1}) + 2\delta. \end{aligned} \tag{E.2}$$

Denote by $\mathcal{E}_{k3}$ the event $\left(L^{(k)}(\mathbf{v}, \mathbf{x}_1), \mathbf{v}, \mathbf{x}_1, \mathbf{x}_2\right)$ not being jointly $P_{UVX_1X_2}$-typical. The error event $\mathcal{E}_k$ is defined as $\mathcal{E}_k = \mathcal{E}_{k-1} \cup \mathcal{E}_{k1} \cup \mathcal{E}_{k2} \cup \mathcal{E}_{k3}$. Then, conditioned on $\mathcal{E}_k^c$ the terminals share sequences $(\mathbf{u}_{ij}, \mathbf{v})$ that are jointly typical with $(\mathbf{x}_1, \mathbf{x}_2)$. In the next stage $k + 1$, the sequence $(\mathbf{u}_{ij}, \mathbf{v})$ plays the role of the sequence $\mathbf{v}$. The scheme for stages with even k is defined analogously with roles of $\mathcal{X}_1$ and $\mathcal{X}_2$ interchanged. We claim that $L^{(1)}, ..., L^{(r)}$ constitutes the required CR along with the communication $\mathbf{F} = F_1, ..., F_k$. Then, (5.31) follows from (E.1), and the second inequality in (5.28) follows from (E.2). Moreover, for every realization $\mathbf{u}_1, ..., \mathbf{u}_r$ of

$L^{(1)}, ..., L^{(r)}$, with $E = \mathbf{1}_{\mathcal{E}_r}$ we have,

$$\begin{aligned}
&\mathrm{P}\left(L^{(1)}, ..., L^{(r)} = \mathbf{u}_1, ..., \mathbf{u}_r \mid E = 0\right) \\
&\leq \mathrm{P}\left(\{(\mathbf{x}_1, \mathbf{x}_2) : (\mathbf{u}_1, ..., \mathbf{u}_r, \mathbf{x}_1, \mathbf{x}_2) \text{ are jointly } P_{U^r X_1 X_2} \text{ typical}\}\right) \\
&\leq \exp\left[-n(I(X_1, X_2 \wedge U^r) - \delta)\right],
\end{aligned}$$

for n large, which further yields

$$\frac{1}{n} H(L^{(1)} ... L^{(r)} \mid E = 0) \geq I(X_1, X_2 \wedge U^r) - \delta.$$

Therefore,

$$\begin{aligned}
\frac{1}{n} H(L^{(1)} ... L^{(r)}) &\geq \frac{1}{n} H(L^{(1)} ... L^{(r)} \mid E = 0) - \mathrm{P}(\mathcal{E}_r) \log |\mathcal{X}_1||\mathcal{X}_2| \\
&\geq I(X_1, X_2 \wedge U^r) - \delta - \mathrm{P}(\mathcal{E}_r) \log |\mathcal{X}_1||\mathcal{X}_2|.
\end{aligned}$$

Thus, the claim will follow upon showing that $\mathrm{P}(\mathcal{E}_r) \to 0$ as $n \to \infty$. In particular, it remains to show that $\mathrm{P}(\mathcal{E}_{k2}) \to 0$ and $\mathrm{P}(\mathcal{E}_{k3}) \to 0$, $k = 1, ..., r$, as $n \to \infty$. As before, we show this for odd k and the proof for even k follows *mutatis mutandis.* To that end, note first that for any jointly P_{UVX_1}-typical $(\mathbf{u}, \mathbf{v}, \mathbf{x}_1)$, the set of $\mathbf{x}_2 \in \mathcal{X}_2^n$ such that $(\mathbf{u}, \mathbf{v}, \mathbf{x}_1, \mathbf{x}_2)$ are jointly typical with $(\mathbf{u}, \mathbf{v}, \mathbf{x}_1)$ has conditional probability close to 1 conditioned on $U^n = \mathbf{u}, V^n = \mathbf{v}, X_1^n = \mathbf{x}_1$, and so by the Markov relation $X_2 -\!\!\circ\!\!- V, X_1 -\!\!\circ\!\!- U$, also conditioned on $V^n = \mathbf{v}, X_1^n = \mathbf{x}_1$. Upon choosing $\mathbf{u} = L^{(k)}(\mathbf{v}, \mathbf{x}_1)$ in the argument above, we get $\mathrm{P}(\mathcal{E}_{k2}) \to 0$. Finally, we show that $\mathrm{P}(\mathcal{E}_{k3})$ will be small, for large probability choices of the random codebook $\{\mathbf{u}_{ij}\}$. Specifically, for fixed typical sequences $(\mathbf{v}, \mathbf{x}_1, \mathbf{x}_2)$, the probability $\mathrm{P}(\mathcal{E}_{k3} \mid V^n = \mathbf{v}, X_1^n = \mathbf{x}_1, X_2^n = \mathbf{x}_2)$ is bounded above exactly as in [2, equation (4.15)]:

$$
\begin{aligned}
&\mathrm{P}\left(\mathcal{E}_{k3} \mid V^n = \mathbf{v}, X_1^n = \mathbf{x}_1, X_2^n = \mathbf{x}_2\right) \\
&\leq \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{l=1, l \neq j}^{N_2} \mathbb{P}\Big((\mathbf{u}_{ij}, \mathbf{v}, \mathbf{x}_1) \text{ jointly } P_{UVX_1}\text{-typical}, (\mathbf{u}_{il}, \mathbf{v}, \mathbf{u}_{il}) \text{ jointly } P_{UVX_2}\text{-typical}\Big) \\
&\leq N_1 N_2^2 . \exp[-n(I(X_1 \wedge U \mid V) + I(X_1 \wedge U \mid V) + o(n))] \\
&\leq \exp[-n\delta + o(n)],
\end{aligned}
$$

for all n sufficiently large. Note that the probability distribution in the calculation above comes from codebook generation, and in particular, the second inequality above uses the fact that $\mathbf{u}_{il}$ and $\mathbf{u}_{ij}$ are independently selected for $l \neq j$. Thus, $\mathrm{P}\left(\mathcal{E}_{k3} \mid \mathcal{E}_{k2}\right) \to 0$ for an appropriately chosen codebook, which completes the proof.

□

Appendix F

Properties of Interactive Communication

We prove two basic properties of interactive communication that were used in the general converse proof in Section 6.8. We retain the notation introduced therein.

Lemma F.1. *For every $A_0 \in \sigma^k$ such that*

$$\frac{d\,\mathrm{P}}{d\,\tilde{\mathrm{P}}}(y^k) > 0, \qquad y^k \in A_0, \tag{F.1}$$

it holds that

$$\tilde{\mathrm{P}}_{Y_1,\dots,Y_k|\mathbf{F}}\left(A_0 \mid \mathbf{i}\right) = \frac{d\,\mathrm{P}_{\mathbf{F}}}{d\,\tilde{\mathrm{P}}_{\mathbf{F}}}(\mathbf{i}) \int_{A_0} \frac{d\,Q_{\mathbf{i}}}{d\,\mathrm{P}}\,d\,\tilde{\mathrm{P}}, \qquad \tilde{\mathrm{P}}_{\mathbf{F}} \text{ a.s. in } \mathbf{i} \tag{F.2}$$

Proof. It suffices to show that the right-side of (F.2) constitutes a version of $\mathbb{E}_{\tilde{\mathrm{P}}}\left\{\mathbf{1}_{A_0} \mid \sigma(\mathbf{F})\right\}$, i.e.,

$$\int_{\mathbf{F}^{-1}(B)} \mathbf{1}_{A_0} d\,\tilde{\mathrm{P}} = \int_B \left(\int_{A_0} \frac{d\,\mathrm{P}_{\mathbf{F}}}{d\,\tilde{\mathrm{P}}_{\mathbf{F}}}(z) \frac{d\,Q_z}{d\,\mathrm{P}} d\,\tilde{\mathrm{P}}\right) \tilde{\mathrm{P}}_{\mathbf{F}}\left(d\,z\right), \tag{F.3}$$

for every set B in the range σ-field of $\mathbf{F}$. To show that, we note for every $A \in \sigma^k$ that

$$\begin{aligned}\int_{\mathbf{F}^{-1}(B)} \mathbf{1}_A d\,\mathrm{P} &= \int_B \mathrm{P}_{Y|\mathbf{F}}\left(A \mid z\right) \mathrm{P}_{\mathbf{F}}\left(d\,z\right) \\ &= \int_B \left(\int_A \frac{d\,Q_z}{d\,\mathrm{P}}\,d\,\mathrm{P}\right) \mathrm{P}_{\mathbf{F}}\left(dz\right), \end{aligned} \tag{F.4}$$

where the previous step uses the assumption (6.73). Using Fubini's and Tonelli's theorems to interchange the order of integrals in (F.4), we get

$$\begin{aligned}\int_{\mathbf{F}^{-1}(B)} \mathbf{1}_A \, d\,\mathrm{P} &= \int_A \left(\int_B \frac{d\,Q_z}{d\,\mathrm{P}} \, \mathrm{P}_{\mathbf{F}}\left(dz\right) \right) d\,\mathrm{P}, \\ &= \int_A \mathbf{1}_{\mathbf{F}^{-1}(B)} \, d\,\mathrm{P},\end{aligned}$$

which further implies

$$\mathbf{1}_{\mathbf{F}^{-1}(B)} = \int_B \frac{d\,Q_z}{d\,\mathrm{P}} \mathrm{P}_{\mathbf{F}}\left(dz\right), \qquad \mathrm{P} \text{ a.s.,} \tag{F.5}$$

since the set $A \in \sigma^k$ was arbitrary. Next, for every B in the range σ-field of $\mathbf{F}$, it follows from (F.5) and (F.1) that

$$\begin{aligned}\int_{\mathbf{F}^{-1}(B)} \mathbf{1}_{A_0} d\,\tilde{\mathrm{P}} &= \int_{A_0} \mathbf{1}_{\mathbf{F}^{-1}(B)} d\,\tilde{\mathrm{P}} \\ &= \int_{A_0} \frac{1}{d\,\mathrm{P}/d\,\tilde{\mathrm{P}}} \mathbf{1}_{\mathbf{F}^{-1}(B)} d\,\mathrm{P} \\ &= \int_{A_0} \frac{1}{d\,\mathrm{P}/d\,\tilde{\mathrm{P}}} \int_B \frac{d\,Q_z}{d\,\mathrm{P}} \mathrm{P}_{\mathbf{F}}\left(dz\right) d\,\mathrm{P} \\ &= \int_{A_0} \int_B \frac{d\,Q_z}{d\,\mathrm{P}} \mathrm{P}_{\mathbf{F}}\left(dz\right) d\,\tilde{\mathrm{P}}.\end{aligned} \tag{F.6}$$

The claim (F.3) follows upon interchanging the order of integrals in (F.6). □

Lemma F.2. *Given measurable sets $A_i \in \sigma_i$, $1 \leq i \leq k$, for $\tilde{\mathrm{P}}$ in (6.76),*

$$\tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}\left(A_1 \times \ldots \times A_k \mid \mathbf{i}\right) = \prod_{j=1}^{k} \tilde{\mathrm{P}}_{Y_j|\mathbf{F}}\left(A_j \mid \mathbf{i}\right), \qquad \mathrm{P}_{\mathbf{F}} \text{ a.s. in } \mathbf{i}, \tag{F.7}$$

where $\tilde{\mathrm{P}}_{Y_1,\ldots,Y_k|\mathbf{F}}$ is the regular conditional probability on $(\mathcal{Y}^k, \sigma^k)$ conditioned on $\sigma(\mathbf{F})$.

Proof. For $1 \leq l \leq r$, $1 \leq j \leq k$, denote by Φ_{lj} the interactive communication preceding F_{lj}, by $\mathbf{F}_{lj}$ the rv (F_{lj}, Φ_{lj}), and by $\mathbf{i}_{lj}$ a realization of $\mathbf{F}_{lj}$. Without loss of generality, we choose a version of $\tilde{\mathrm{P}}_{Y^k|\mathbf{F}}$ that satisfies

$$\tilde{\mathrm{P}}_{Y^k|\mathbf{F}_{lj}}\left(\mathbf{F}_{lj}^{-1}(\mathbf{i}_{lj})^c \mid \mathbf{i}_{lj}\right) = 0, \quad \tilde{\mathrm{P}}_{\mathbf{F}_{lj}} \text{ a.s.}, \tag{F.8}$$

for all $1 \leq l \leq r, 1 \leq j \leq k$. The following property of interactive communication is pivotal to our proof: For each i_{lj}^-, $\Phi_{lj}^{-1}(i_{lj}^-)$ is a product set, i.e.,

$$\Phi_{lj}^{-1}(i_{lj}^-) = A_1' \times ... \times A_k', \quad A_j' \in \sigma_j,\ 1 \leq j \leq k.$$

We prove the claim by induction upon observing that $\tilde{\mathrm{P}}_{Y^k|\mathbf{F}_{lj}}$ can be obtained by conditioning $\tilde{\mathrm{P}}_{Y^k|\Phi_{lj}}$ on the rv F_{lj}.

Formally, denote by $\sigma^k(i_{lj}^-) = \sigma_1(i_{lj}^-) \times ... \times \sigma_k(i_{lj}^-)$ the σ-field induced by σ^k on $A_1' \times ... \times A_k'$, and by $\sigma\left(F_{lj}(\cdot, i_{lj}^-)\right)$ the smallest sub-σ-field of $\sigma^k(i_{lj}^-)$ with respect to which F_{lj} is measurable (for i_{lj}^- fixed). Using (F.8), we choose a version of $\tilde{\mathrm{P}}_{Y^k|\mathbf{F}}$ such that for each $1 \leq l \leq r$ and $1 \leq j \leq k$, $\tilde{\mathrm{P}}_{Y^k|\mathbf{F}_{lj}}(\cdot \mid \mathbf{i}_{lj})$ is the regular conditional probability on the probability space

$$\left(A_1' \times ... \times A_k',\ \sigma^k(i_{lj}^-),\ \tilde{\mathrm{P}}_{Y^k|\Phi_{lj}}\left(\cdot \mid i_{lj}^-\right)\right)$$

conditioned on $\sigma\left(F_{lj}(\cdot, i_{lj}^-)\right)$. Specifically,

$$\tilde{\mathrm{P}}_{Y^k|\mathbf{F}_{lj}}(A \mid \mathbf{i}_{lj}) = \mathbb{E}_{\tilde{\mathrm{P}}_{Y^k|\Phi_{lj}}\left(\cdot|i_{lj}^-\right)}\left\{\mathbf{1}_A \mid \sigma\left(F_{lj}(\cdot, i_{lj}^-)\right)\right\}(i_{lj}), \qquad A \in \sigma^k, \tag{F.9}$$

where the underlying σ-field for the conditional expectation is $\sigma^k(i_{lj}^-)$. For this version of $\tilde{\mathrm{P}}_{Y^k|\mathbf{F}}$, we show below that if (F.7) holds with Φ_{lj} in the role of $\mathbf{F}$, then it holds with $\mathbf{F}_{lj}$ in the role of $\mathbf{F}$. Lemma F.2 then follows by induction since (F.7) holds with $\mathbf{F} = \emptyset$.

It remains to prove the assertion above. To that end, for $B \in \mathcal{F}_{lj}$, denote by $F_{lj}^{-1}\left(B, i_{lj}^{-}\right)$ the set

$$\left\{y_j \in \mathcal{Y}_j : F_{lj}\left(y_j, i_{lj}^{-}\right) \in B\right\}.$$

With an abuse of notation, we do not distinguish between the sets $F_{lj}^{-1}\left(B, i_{lj}^{-}\right)$ and its cylindrical extension

$$\mathcal{Y}_1 \times ... \times F_{lj}^{-1}\left(B, i_{lj}^{-}\right) \times ... \times \mathcal{Y}_k.$$

Then, using the notation $\tilde{Q}_{i_{lj}^{-}}$ and $\tilde{Q}^t_{i_{lj}^{-}}, 1 \leq t \leq k$, for the probability measures $\tilde{\mathrm{P}}_{Y^k|\Phi_{lj}}\left(\cdot \mid i_{lj}^{-}\right)$ and $\tilde{\mathrm{P}}_{Y_t|\Phi_{lj}}\left(\cdot \mid i_{lj}^{-}\right)$, $1 \leq t \leq k$, respectively, our induction hypothesis states

$$\tilde{Q}i_{lj}^{-}(A_1 \times .... \times A_k) = \prod_{t=1}^{k} \tilde{Q}i_{lj}^{-}(A_t), \quad A_t \in \sigma_t,\ 1 \leq t \leq k. \tag{F.10}$$

It follows that

$$\begin{aligned}
&\int_{F_{lj}^{-1}(B, i_{lj}^{-})} \mathbf{1}_{A_1 \times ... \times A_k}\, d\,\tilde{Q}_{i_{lj}^{-}} \\
&= \int_{F_{lj}^{-1}(B, i_{lj}^{-})} \mathbf{1}_{A_1 \cap A_1' \times ... \times A_k \cap A_k'}\, d\,\tilde{Q}_{i_{lj}^{-}} \\
&= \left[\prod_{t \neq j} \int \mathbf{1}_{A_t \cap A_t'}\, d\,\tilde{Q}^t_{i_{lj}^{-}}\right] \int_{F_{lj}^{-1}(B, i_{lj}^{-})} \mathbf{1}_{A_j \cap A_j'}\, d\,\tilde{Q}^j_{i_{lj}^{-}},
\end{aligned} \tag{F.11}$$

where the first equality uses (F.8) and the second uses (F.10). Defining

$$P_{lj}^t(A) \triangleq \mathbb{E}_{\tilde{Q}i_{lj}^{-}{}^{t}}\left\{\mathbf{1}_A \mid \sigma\left(F_{lj}(\cdot, i_{lj}^{-})\right)\right\}, \quad A \in \sigma_t(i_{lj}^{-}),\ 1 \leq t \leq k,$$

we have from (F.11) that

$$\begin{aligned}
&\int_{F_{lj}^{-1}(B,i_{lj}^-)} \mathbf{1}_{A_1\times\ldots\times A_k}\,d\tilde{Q}_{i_{lj}^-}\\
&=\left[\prod_{t\neq j}\int P_{lj}^t(A_t\cap A_t')\,d\tilde{Q}^t_{i_{lj}^-}\right]\int_{F_{lj}^{-1}(B,i_{lj}^-)} P_{lj}^j(A_j\cap A_j')\,d\tilde{Q}^j_{i_{lj}^-}\\
&=\int_{F_{lj}^{-1}(B,i_{lj}^-)}\prod_{t=1}^k P_{lj}^t(A_t\cap A_t')\,d\tilde{Q}_{i_{lj}^-},
\end{aligned}$$

where the second equality uses (F.10). Thus, by (F.9),

$$\tilde{\mathrm{P}}_{Y^k|\mathbf{F}_{lj}}\left(A_1\times\ldots\times A_k\mid \mathbf{i}_{lj}\right)=\prod_{t=1}^k P_{lj}^t(A_t\cap A_t'),\quad \tilde{\mathrm{P}}_{\mathbf{F}_{lj}}\text{ a.s. in }\mathbf{i}_{lj}. \tag{F.12}$$

Since by (F.8) $P_{lj}^t(A_t')=1$, $1\leq t\leq k$, it follows from (F.12) that

$$\begin{aligned}
P_{lj}^t(A_t)&=P_{lj}^t(A_t\cap A_t')\\
&=\tilde{\mathrm{P}}_{Y^k|\mathbf{F}_{lj}}\left(A_1'\times\ldots\times A_{t-1}'\times A_t\times A_{t+1}'\times\ldots\times A_k'\mid \mathbf{i}_{lj}\right)\\
&=\tilde{\mathrm{P}}_{Y_t|\mathbf{F}_{lj}}\left(A_t\mid \mathbf{i}_{lj}\right).
\end{aligned}$$

The previous observation, along with (F.12), implies that (F.7) holds with $\mathbf{F}_{lj}$ in the role $\mathbf{F}$. □

Index

Bibliography

[1] R. Ahlswede and I. Csiszár. Common randomness in information theory and cryptography–part i: Secret sharing. *IEEE Trans. Inf. Theory*, 39(4):1121–1132, July 1993. [7, 11, 12, 21, 24, 26, 63, 142]

[2] R. Ahlswede and I. Csiszár. Common randomness in information theory and cryptography–part ii: CR capacity. *IEEE Trans. Inf. Theory*, 44(1):225–240, January 1998. [4, 8, 21, 32, 108, 126, 187]

[3] R. Ahlswede and I. Csiszár. On the oblivious transfer capacity. *IEEE International Symposium on Information Theory*, pages 2061–2064, 2007. [7]

[4] E. Arikan. An inequality on guessing and its application to sequential decoding. *IEEE Trans. Inf. Theory*, 42(1):99–105, January 1996. [16, 131]

[5] E. Arikan and N. Merhav. The Shannon cipher system with a guessing wiretapper. *IEEE Trans. Inf. Theory*, 45(6):1860–1866, September 1999. [16, 131]

[6] R. B. Ash. *Real Analysis and Probability.* Academic Press, 1972. [156]

[7] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer. Generalized privacy amplification. *IEEE Trans. Inf. Theory*, 41(6):1915–1923, November 1995. [8, 32]

[8] S. Bobkov and M. Madiman. Concentration of the information in data with log-concave distributions. *Annals of Probability*, 39(4):1528–1543, 2011. [158]

[9] M. Braverman and A. Rao. Information equals amortized communication. *Annual Symposium on Foundations of Computer Science*, pages 748–757, 2011. [167]

[10] C. Chan. Generating secret in a network. *Ph. D. Dissertation, Massachussetts Institute of Technology*, 2010. [146]

[11] C. Chan. Multiterminal secure source coding for a common secret source. *Proc. Conference on Communication, Control, and Computing (Allerton)*, pages 188–195, Sep 2011. [8]

[12] C. Chan and L. Zheng. Mutual dependence for secret key agreement. *Proc. Annual Conference on Information Sciences and Systems (CISS 2010).* [26]

[13] P.N. Chen and F. Alajaji. Csiszár's cutoff rates for arbitrary discrete sources. *IEEE Trans. Inf. Theory*, 47(1):330–338, 2001. [16]

[14] C. Y. Chong and S. P. Kumar. Sensor networks: Evolution, opportunities, and challenges. *Proc. IEEE*, 91(8):1247 – 1256, August 2003. [1]

[15] T. Cover and S. Pombra. Gaussian feedback capacity. *IEEE Trans. Inf. Theory*, 35(1):37–43, January 1989. [158]

[16] I. Csiszár. Almost independence and secrecy capacity. *Prob. Pered. Inform.*, 32(1):48–57, 1996. [24, 41, 84]

[17] I. Csiszár and J. Körner. *Information theory: Coding theorems for discrete memoryless channels.* Academic Press, 1981. [47, 55, 60, 172]

[18] I. Csiszár and J. Körner. *Information theory: Coding theorems for discrete memoryless channels. 2nd edition.* Cambridge University Press, 2011. [24, 72, 86, 87, 110, 111, 137, 153, 154, 155, 181]

[19] I. Csiszár and P. Narayan. Common randomness and secret key generation with a helper. *IEEE Trans. Inf. Theory*, 46(2):344–366, March 2000. [41]

[20] I. Csiszár and P. Narayan. Secrecy capacities for multiple terminals. *IEEE Trans. Inf. Theory*, 50(12):3047–3061, December 2004. [2, 4, 7, 8, 12, 15, 19, 22, 23, 24, 25, 26, 33, 34, 35, 49, 51, 56, 57, 63, 72, 88, 106, 137, 153, 172, 174, 175, 178]

[21] I. Csiszár and P. Narayan. Secrecy capacities for multiterminal channel models. *IEEE Trans. Inf. Theory*, 54(6):2437–2452, June 2008. [7, 8, 15, 22, 23, 26, 77, 137]

[22] 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*, 38(1):97–139, 2008. [14]

[23] T. Eisenbarth and S. Kumar. A survey of lightweight-cryptography implementations. *IEEE Design and Test of Computers*, 24(6):522–533, December 2007. [12]

[24] K. Eswaran and M. Gastpar. Rate loss in the ceo problem. *Proc. Conference on Information Sciences and Systems*, March 2005. [13]

[25] N. M. Freris, H. Kowshik, and P. R. Kumar. Fundamentals of large sensor networks: Connectivity, capacity, clocks, and computation. *Proc. IEEE*, 98(11):1828 – 1846, November 2010. [1]

[26] P. Gács and J. Körner. Common information is far less than mutual information. *Problems of Control and Information Theory*, 2(2):149–162, 1973. [9, 20, 27, 28, 47, 48, 135, 169, 170, 171]

[27] R. G. Gallager. Finding parity in a simple broadcast nework. *IEEE Trans. Inf. Theory*, 34(2):176–180, March 1988. [7]

[28] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas. Silicon physical random functions. *Proc. ACM Conference on Computer Communications and Security*, pages 148 –160, November 2002. [1]

[29] A. Giridhar and P. R. Kumar. Computing and communicating functions over sensor networks. *IEEE Journ. on Select. Areas in Commun.*, 23(4):755–764, April 2005. [7]

[30] B. Girod, A. Margot, S. Rane, and D. Rebollo-Monedero. Distributed video coding. *Proc. IEEE*, 93(1):71–83, January 2005. [1]

[31] A. A. Gohari and V. Anantharam. Information-theoretic key agreement of multiple terminalspart i. *IEEE Trans. Inf. Theory*, 56(8):3973 – 3996, August 2010. [63]

[32] T. S. Han. *Information-Spectrum Methods in Information Theory [English Translation]*. Series: Stochastic Modelling and Applied Probability, Vol. 50, Springer, 2003. [38]

[33] T. S. Han and S. Verdú. Approximation theory of output statistics. *IEEE Trans. Inf. Theory*, 39(3):752–772, May 1993. [38]

[34] M. K. Hanawal and R. Sundaresan. The Shannon cipher system with a guessing wiretapper: General sources. *IEEE Trans. Inf. Theory*, 57(4):2503–2516, April 2011. [16, 131]

[35] G. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities. 2nd edition.* Cambridge University Press, 1952. [144, 150, 162]

[36] R. Impagliazzo and D. Zuckerman. How to recycle random bits. *Proc. Annual Symposium on Foundations of Computer Science*, pages 248–253, 1989. [25]

[37] A. K. Jain, A. Ross, and S. Pankanti. Biometrics: a tool for information security. *IEEE Transactions on Information Forensics and Security*, 1(2):125–143, 2006. [3]

[38] S. Kamath and V. Ananthram. A new dual to the Gács-Körner common information defined via the Gray-Wyner system. *Proc. Conference on Communication, Control, and Computing (Allerton)*, pages 1340–1346, Oct 2010. [113]

[39] A. H. Kaspi. Two-way source coding with a fidelity criterion. *IEEE Trans. Inf. Theory*, 31(6):735–740, November 1985. [109]

[40] J. Korner and K. Marton. How to encode the modulo-two sum of binary sources. *IEEE Trans. Inf. Theory*, 25(2):219–221, March 1979. [7]

[41] M. Loeve. *Probability theory. 2nd Edition.* Van Nostrand New York, 1960. [135, 169]

[42] N. Ma and P. Ishwar. Some results on distributed source coding for interactive function computation. *IEEE Trans. Inf. Theory*, 57(9):6180–6195, September 2011. [110]

[43] N. Ma, P. Ishwar, and P. Gupta. Information-theoretic bounds for multiround function computation in collocated networks. *IEEE International Symposium on Information Theory*, pages 2306–2310, 2009. [7]

[44] M. Madiman and A. Barron. Generalized entropy power inequalities and monotonicity properties of information. *IEEE Trans. Inf. Theory*, 53(7):2317–2329, July 2007. [26]

[45] M. Madiman and A. Barron. Entropy and set cardinality inequalities for partition-determined functions. *Random Structures and Algorithms*, 40:399–424, 2012. [26, 152]

[46] M. Madiman and P. Tetali. Information inequalities for joint distributions, with interpretations and applications. *IEEE Trans. Inf. Theory*, 56(6):2699–2713, June 2010. [26, 77, 146]

[47] J. L. Massey. Guessing and entropy. *Proc. IEEE International Symposium on Information Theory*, 1994. [16, 131]

[48] U. M. Maurer. Secret key agreement by public discussion from common information. *IEEE Trans. Inf. Theory*, 39(3):733–742, May 1993. [7, 11, 12, 21, 24, 26, 63, 126]

[49] U. M. Maurer. *Communications and Cryptography: Two sides of One Tapestry*, chapter 26, pages 271–285. Norwell, MA: Kluwer, R.E. Blahut et al., Eds. edition, 1994. [24, 41]

[50] A.C.A. Nascimento and A. Winter. On the oblivious transfer capacity of noisy correlations. *Proc. IEEE International Symposium on Information Theory*, pages 1871–1875, 2009. [7]

[51] S. Nitinawarat and P. Narayan. Secret key generation for correlated Gaussian sources. *IEEE Trans. Inf. Theory*, 58(6):3373–3391, June 2012. [155, 159]

[52] A. Orlitsky and A. El Gamal. Communication with secrecy constraints. *STOC*, pages 217–224, 1984. [7]

[53] A. Orlitsky and J. R. Roche. Coding for computing. *IEEE Trans. Inf. Theory*, 47(3):903–917, March 2001. [7]

[54] R. S. Pappu. Physical one-way functions. *Ph. D. Dissertation, Massachussetts Institute of Technology*, 2001. [3]

[55] R. Puri, A. Majumdar, P. Ishwar, and K. Ramchandran. Distributed video coding in wireless sensor networks. *IEEE Signal Processing Magazine*, pages 94–106, July 2006. [1]

[56] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst. J.*, 40(3):614–634, 2001. [1]

[57] A. Rényi. On measures of entropy and information. *Proc. Fourth Berkeley Symposium on Mathematics Statistics and Probability, Vol. 1 (Univ. of Calif. Press)*, pages 547–561, 1961. [35]

[58] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948. [2, 16]

[59] H. Shimokawa. Rényi's entropy and error exponent of source coding with countably infinite alphabet. *Proc. IEEE International Symposium on Information Theory*, 2006. [16]

[60] D. Slepian and J. Wolf. Noiseless coding of correlated information source. *IEEE Trans. Inf. Theory*, 19(4):471–480, July 1973. [83, 100, 105]

[61] H. Tyagi. Minimal public communication for maximum rate secret key generation. *Proc. IEEE International Symposium on Information Theory*, pages 578–582, 2011. [95]

[62] H. Tyagi. Distributed computing with privacy. *Proc. IEEE International Symposium on Information Theory*, pages 1157–1161, 2012. [64]

[63] H. Tyagi. Common information and secret key capacity. *To appear, IEEE Trans. Inf. Theory*, 2013. [19, 95]

[64] H. Tyagi. Distributed function computation with confidentiality. *IEEE Journal on Selected Areas in Communications*, 31(4):691–701, April 2013. [64]

[65] H. Tyagi and P. Narayan. How many queries will resolve common randomness? *Proc. IEEE International Symposium on Information Theory*, 2013. [19, 131]

[66] H. Tyagi and P. Narayan. How many queries will resolve common randomness? *To appear, IEEE Trans. Inf. Theory*, 2013. [131]

[67] H. Tyagi, P. Narayan, and P. Gupta. Secure computing. *Proc. IEEE International Symposium on Information Theory*, pages 2612 – 2616, 2010. [39]

[68] H. Tyagi, P. Narayan, and P. Gupta. When is a function securely computable? *IEEE Trans. Inf. Theory*, 57(10):6337–6350, October 2011. [19, 39, 92]

[69] H. Tyagi, P. Narayan, and P. Gupta. When is a function securely computable? *Proc. IEEE International Symposium on Information Theory*, pages 2876–2880, 2011. [39]

[70] H. Tyagi and S. Watanabe. Work in progress. 2013. [6]

[71] S. Venkatesan and V. Anantharam. The common randomness capacity of a pair of independent discrete memoryless channels. *IEEE Trans. Inf. Theory*, 44(1):215–224, January 1998. [144]

[72] S. Venugopal, R. Buyya, and K. Ramamohanarao. A taxonomy of data grids for distributed data sharing, management and processing. *ACM Computing Surveys*, 38(3), 2006. [1]

[73] S. Watanabe. Personal communication. 2013. [25]

[74] H. S. Witsenhausen. On sequences of pairs of dependent random variables. *Siam J. Appl. Math.*, 28(1):100–113, January 1975. [20]

[75] Y. Wu. Personal communicattion. 2013. [16]

[76] A. D. Wyner. Recent results in the Shannon theory. *IEEE Trans. Inf. Theory*, 20(1):2–10, January 1974. [43]

[77] A. D. Wyner. The common information of two dependent random variables. *IEEE Trans. Inf. Theory*, 21(2):163–179, March 1975. [12, 28, 118]

[78] G. Xu and B. Chen. The sufficiency principle for decentralized data reduction. *Proc. IEEE International Symposium on Information Theory*, pages 319–323, 2012. [13]

[79] A. C. Yao. Some complexity questions related to distributive computing. *Proc. Annual Symposium on Theory of Computing*, pages 209–213, 1979. [7]

[80] A. C. Yao. Protocols for secure computations. *Proc. Annual Symposium on Foundations of Computer Science*, pages 160–164, 1982. [6, 9]

[81] C. Ye. Information theoretic generation of multiple secret keys. *Ph. D. Dissertation, University of Maryland, College Park*, 2005. [5, 43, 44]

[82] C. Ye and P. Narayan. Secret key and private key constructions for simple multiterminal source models. *Proc. IEEE International Symposium on Information Theory*, pages 2133–2137, 2005. [43, 44]

[83] C. Ye and P. Narayan. Secret key and private key constructions for simple multiterminal source models. *IEEE Trans. Inf. Theory*, 58(2):639–651, February 2012. [5]