Spectral approach to the communication complexity of multi-party key agreement

Geoffroy Caillat-Grenier, Andrei Romashchenko

Spectral approach to the communication complexity of multi-party key agreement

Geoffroy Caillat-Grenier and Andrei Romashchenko

September 29, 2023

Abstract

We propose a linear algebraic method, rooted in the spectral properties of graphs, that can be used to prove lower bounds in communication complexity. Our proof technique effectively marries spectral bounds with information-theoretic inequalities. The key insight is the observation that, in specific settings, even when data sets X and Y are closely correlated and have high mutual information, the owner of X cannot convey a reasonably short message that maintains substantial mutual information with Y. In essence, from the perspective of the owner of Y, any sufficiently brief message $m = m(X)$ would appear nearly indistinguishable from a random bit sequence.

We employ this argument in several problems of communication complexity. Our main result concerns cryptographic protocols. We establish a lower bound for communication complexity of multi-party secret key agreement with unconditional, i.e., information-theoretic security. Specifically, for one-round protocols (simultaneous messages model) of secret key agreement with three participants we obtain an asymptotically tight lower bound. This bound implies optimality of the previously known *omniscience* communication protocol (this result applies to a non-interactive secret key agreement with three parties and input data sets with an arbitrary symmetric information profile).

We consider communication problems in one-shot scenarios when the parties' inputs are not produced by any i.i.d. sources, and there are no ergodicity assumptions on the input data. In this setting, we found it natural to present our results using the framework of Kolmogorov complexity.

Keywords: communication complexity, Kolmogorov complexity, information-theoretic cryptography, multi-party secret key agreement, expander mixing lemma, information inequalities

1 Introduction

Within computer science, a broad range of communication complexity problems has been studied in recent decades. In these problems several (two or more) agents solve together some task (compute a function, search an elements in a set, sample a distribution, and so on) when the input data are distributed among the agents. In different context we may impose different constraints on the class of admissible protocols (protocols can be deterministic or randomized, one-way or interactive, with a one shot of simultaneous messages or with several rounds, etc.). The cost of a communication protocol is the total number of bits that must be exchanged between participants, typically in the worst-case situation.

In this paper we focus on communication problems with three parties (Alice, Bob, and Charlie), though our techniques can be extended to bigger number of participants. We deal with the situation when the input data accessible to Alice, Bob, and Charlie are correlated. In a popular model *number-on-forehead*, the datasets given to Alice, Bob, and Charlie have large intersections, which is a very particular form of correlation between the data. We study a more general setting (more usual in cryptography and information theory) where the input data sets given to the parties have large mutual information, but it might be impossible to materialize this mutual information as common chunks of bits shared by several parties.

The principal communication problem under consideration is *secret key agreement*: Alice, Bob, and Charlie use the correlation between their input data sets to produce a common secret key. A special feature of this setting is the implicit presence of another participant in the game, Eve (eavesdropper/adversary). The eavesdropper can intercept all messages between Alice, Bob, and Charlie, but this should not give Eve any information about the final result of the protocol — the produced secret key. A secret key agreement (for two or many participants) is

one of the basic primitives in cryptography; it can serve as a part of more sophisticated protocols (the produced secret key can be used in a one-time pad encryption or in more complicated cryptographic schemes).

In practice, the most standard and well known method of secret key agreement is the Diffie-Hellman key exchange [1, 2] and its generalizations, see [3]. The security of this protocol is based on assumptions of computational complexity. In particular, the Diffie-Hellman scheme is secure only if the eavesdropper cannot solve efficiently the problem of discrete logarithms. Such an assumption looks plausible for most practical applications. However, theoretical cryptography studies also secret key agreement in information-theoretic settings, where we impose no restrictions on the computational power of the eavesdropper. Besides a natural theoretical interest, such a scheme can be useful as a building block in more complex protocols. In particular, a protocol of information-theoretic secret key agreement (pretty conventional, involving communication and computational tools conceivable in the framework of the classical physics) is an indispensable component of the protocol of quantum key distribution ([4, 5, 6]).

Example 1. Let us recall that the standard protocols of quantum key distribution (see, e.g., [4]) can be subdivided into two phases: in the first one, two parties use a quantum communication channel and quantum measurements to produce on both ends a pair of preliminary results that look like two strongly correlated but not identical sequences of random bits; in the second phase, the parties use a classical communication channel and purely classical computations to perform some sanity check and make sure that the quantum communication was not compromised, and then extract a common secret key from the pair of correlated sequences of bits produced in the quantum phase. The last part of this scheme is exactly an information-theoretic secret key agreement used in the setting when the two parties are preliminary given a pair of highly correlated inputs. The size of the shared secret key produced in the classical phase of the protocol depends on the rate of correlation between the sequences generated by the parties in the quantum phase.

Besides quantum cryptography, secret-key agreement based on correlated information appears in various cryptographic schemes connected with noisy data (biometric information, observations of an inherently noisy communication channel or other physical phenomenon, see the discussions in [7, 9]), in the bounded-storage model ([10, 11]), and so on. We refer the reader to the survey [8] for a more detailed discussion.

In the Diffie-Hellman scheme, the parties may start the protocol from zero, holding initially no secret information. In contrast, a secret key agreement with information-theoretic secrecy is impossible if the parties start from scratch. To produce a key that is secret in information-theoretic sense, the participants of the protocol need to be given some input data (inaccessible to the eavesdropper). The pieces of input data provided to the parties must be correlated with each other, and the measure of this correlation determines the optimal size of the common secret key that can be produced.

So far we were very informal and did not specify the mathematical definitions behind the words *secrecy* (of the key) and *correlation* (between parties' inputs). Let us describe the settings of information-theoretic secret key agreement more precisely. This can be done in different mathematical frameworks.

Historically, information-theoretically secure protocols of secret key agreement were introduced in classical information theory, [24, 25]. In this setting, the input data of the parties are produced by correlated random variables. In the settings with two parties it is usually assumed that there is a sequence of i.i.d. pairs of random variables with finite range, (X_i, Y_i), $i = 1, \dots, n$, and Alice and Bob receive the values of $(X_1 \dots X_n)$ and $(Y_1 \dots Y_n)$ respectively,

$$\begin{array}{ll} \text{Alice} & \leftarrow \quad (X_1 X_2 \dots X_n), \\ \text{Bob} & \leftarrow \quad (\,Y_1\, Y_2\, \dots\, Y_n\,). \end{array}$$

Then Alice and Bob run a communication protocol and try to produce a common value (secret key) W asymptotically independent of the *transcript* (the transcript consist of the messages sent by Alice and Bob to each other). Ahlswede–Csiszar [24] and Maurer [25] found a characterization of the optimal size of W in terms of Shannon's entropy of the input data. They showed that the optimal size of the secret key is asymptotically equal to the mutual information between Alice's and Bob's inputs. A similar characterization of the optimal secret key is known for multi-party protocols, with $k \geq 3$ parties, [22]. The problem of secret key agreement and a related problem of *common randomness generation* were extensively studied in the information theory community and also (in somewhat different settings) in theoretical computer science, see, e.g., [34, 35] and the survey [36].

In this paper we follow the paradigm of building the foundations of cryptography in the framework of algorithmic information theory, as suggested in a general form in [19] and more specifically for secret key agreement in [14, 15]. In this approach, the information-theoretic characteristics of the data are defined not in terms of Shannon's entropy but in terms of Kolmogorov complexity. In this setting, we can talk about properties of *individual*

inputs, keys, transcripts, and not about *probability distributions*. We assume that the parties (Alice, Bob, Charlie) are given as inputs binary strings x, y, z respectively,

$$\begin{array}{lll} \text{Alice} & \leftarrow & x, \\ \text{Bob} & \leftarrow & y, \\ \text{Charlie} & \leftarrow & z, \end{array}$$

and that the parties know the complexity profile of these strings, i.e., the optimal compression rate of these inputs (precisely or at least approximately, see below). The secrecy of the produced key means that this key must be incompressible, even conditional on the public data including the transcript of the communication protocol. In other words, the mutual information (in the sense of Kolmogorov complexity) between the key and the messages sent via the communication channel (the transcript) must be negligibly small. Practically, this property guarantees that the adversary can crack an encryption scheme based on this key only by the brute-force search, see the discussion in [15].

Remark 1. The approach based on Kolmogorov complexity seems more general since we do not need to assume that inputs have any property of stationarity or ergodicity, we do not fix in advance the probability distribution of the pairs of inputs, we do not even assume the existence of such a distribution. However, the frameworks of Shannon and Kolmogorov for the definition of secrecy have similar practical interpretations. Indeed, a distribution W on $\{0,1\}^n$ has a high entropy, i.e., $H(W) \approx n$, if and only if with a high probability W returns an n-bit string with Kolmogorov complexity close to n (a random source with a high entropy typically produces incompressible values). For a more detailed discussion of the connection between Shannon's and Kolmogorov's formalism see [12]. The formal statements in Kolmogorov's framework are usually stronger than their homologues in Shannon's framework, and theorems from the former theory in most cases formally imply the corresponding results from the latter theory, see [14].

A characterization of the optimal size of the secret key in term of Kolmogorov complexity was suggested in [14]. We begin with the case of two parties, see Theorem 1 below. In this theorem, a communication protocol is randomized (we assume that the parties may use a public source of random bits, which is also accessible to the eavesdropper). Let x and y stand for inputs of Alice and Bob, r denote the string of bits produced by a public source of randomness (used by the parties and accessible to the eavesdropper), and t denote the transcript of the protocol.

Theorem 1 ([14]). *(i) For any numbers $k, \ell \in \mathbb{N}$ and $\epsilon, \delta > 0$ there exist a randomized communication protocols $\pi_{k,\ell,\epsilon,\delta}$ such that on every pair of input strings (x, y) (of length at most n) satisfying*[1] $\mathrm{C}(x) \stackrel{\delta}{=} k$ *and* $\mathrm{C}(x \mid y) \stackrel{\delta}{=} \ell$, *Alice and Bob with probability $1 - \epsilon$ both obtain a result $w = w(x, y, r)$ such that*

$$[\textit{length of } w \textit{ in bits}] = \mathrm{C}(x) - \mathrm{C}(x \mid y) - O(\delta) - o(n) \textit{ and } \mathrm{C}(w \mid \langle t, r\rangle) \geq |w| - o(n) \tag{1}$$

(for $n = |x| + |y|$), which means that the size of the produced secret key is asymptotically equal to the mutual information between Alice's and Bob's inputs, and the leakage of information on the key to the eavesdropper (who can access the transcript of the protocol t and public randomness r) is negligibly small.

(ii) The size of the key in (i) is pretty much optimal: no communication protocol can produce a key w longer than $\mathrm{C}(x) - \mathrm{C}(x \mid y) + O(\delta) + o(n)$ without loosing the property of secrecy $\mathrm{C}(w \mid \langle t, r\rangle) \geq [\textit{length of } w \textit{ in bits}] - o(n)$ (the size of a secret key cannot be made asymptotically greater than the mutual information between Alice's and Bob's inputs).

Remark 2. In Theorem 1, the values of k and ℓ are embedded in the communication protocol $\pi_{k,\ell,\epsilon,\delta}$. This means that the parties in some sense "know" (at least approximately) the values of $\mathrm{C}(x)$ and $\mathrm{C}(x \mid y)$. This is similar to the settings of the classical information theory, where the parties "know" the probability distribution on random inputs and can use a suitable protocol. The theorem is nontrivial if the approximation rate $\delta = o(n)$ as $n \to \infty$.

Remark 3. The precision in Eq. (1) in Theorem 1 can be made tighter: there exists a communication protocol which guarantees

$$[\text{length of } w \text{ in bits}] = \mathrm{C}(x) - \mathrm{C}(x \mid y) - O(\delta) - O(\log n) \text{ and } \mathrm{C}(w \mid \langle t, r\rangle) \geq |w| - O(1). \tag{2}$$

[1] Here the term $\mathrm{C}(x)$ stands for the plain Kolmogorov complexity of x (optimal compression of x), the term $\mathrm{C}(x \mid y)$ stands for conditional Kolmogorov complexity of x conditional on y (optimal compression of x given advice y), and the notation $\mathrm{C}(x) \stackrel{\delta}{=} k$ and $\mathrm{C}(x \mid y) \stackrel{\delta}{=} \ell$ means that $|\mathrm{C}(x) - k| \leq \delta$ and $|\mathrm{C}(x \mid y) - \ell| \leq \delta$.

Theorem 1 can be extended to the multi-party setting, where $k > 2$ parties are given correlated data and need to agree on common secret key communicating via a public channel. Let us discuss in more detail the version with $k = 3$ participants. We assume now that three parties (Alice, Bob, and Charlie) are involved in the protocol. They are given inputs x, y, z respectively. We assume that all parties have an access to a common source of random bits (we denote by r the bits produced by this source) and exchange messages via a public channel (we use the conventional definition of a multi-party communication protocol with a public source of random bits, see [21]). It is assumed that every message sent by any party reaches every other party (and the eavesdropper). In what follows we consider only triples of inputs (x, y, z) with a "symmetric" complexity profile such that $\mathrm{C}(x) \approx \mathrm{C}(y) \approx \mathrm{C}(z)$ and $\mathrm{C}(x, y) \approx \mathrm{C}(x, z) \approx \mathrm{C}(y, z)$.

Theorem 2 (symmetric version of [14, Theorem 5.11])**.** *(i) For any profile $(k_1, k_2, k_3) \in \mathbb{N}^3$ and $\epsilon, \delta > 0$ there exist a randomized communication protocols $\pi_{k_1,k_2,k_3,\epsilon,\delta}$ for three parties such that on every triple of binary input strings (x, y, z) (of length at most n) satisfying*

$$\mathrm{C}(x) \stackrel{\delta}{=} \mathrm{C}(y) \stackrel{\delta}{=} \mathrm{C}(z) \stackrel{\delta}{=} k_1,\ \mathrm{C}(x, y) \stackrel{\delta}{=} \mathrm{C}(x, z) \stackrel{\delta}{=} \mathrm{C}(y, z) \stackrel{\delta}{=} k_2,\ \mathrm{C}(x, y, z) \stackrel{\delta}{=} k_3 \tag{3}$$

Alice, Bob, and Charlie can agree with probability $1 - \epsilon$ on a key $w = w(x, y, z, r)$ such that

$$[\textit{length of } w \textit{ in bits}] = \ \tfrac{I(x:y|z)+I(x:z|y)+I(y:z|z)}{2} + I(x : y : z)\ - O(\delta) - o(n) \tag{4}$$

(for $n = |x| + |y| + |z|$) and

$$\mathrm{C}(w \mid \langle t, r\rangle) \geq |w| - o(n). \tag{5}$$

(ii) The size of the key in (i) is asymptotically optimal, i.e., no communication protocol can give a key w asymptotically longer than

$$\frac{1}{2}\left(I(x : y \mid z) + I(x : z \mid y) + I(y : z \mid z)\right) + I(x : y : z) + O(\delta) + o(n) \tag{6}$$

without loosing the property of secrecy (5)*.*

Remark 4. The general version of [14, Theorem 5.11] applies to a triple of inputs with arbitrary (possibly non-symmetric) complexity profile. In the general case, the characterization of the optimal size of the secret key is more involved then (4) and involves piece-wise linear expression involving the terms of the mutual information for x, y, and z, see [14]. We discuss only symmetric complexity profiles in order to avoid cumbersome formulas and focus on the most essential combinatorial ideas behind the proofs.

The known proofs of the positive parts of Theorem 1 and Theorem 2 (the existence of protocols) are quite explicit and constructive: we know specific communication protocols that allow to produce a secret key of the optimal size. More specifically, the proofs suggested in [14] provide a protocol for Theorem 1(i) with communication complexity

$$\min\{\mathrm{C}(x \mid y), \mathrm{C}(y \mid x)\} + O(\delta) + O(\log n) \tag{7}$$

and a protocol[2] for Theorem 2(i) with communication complexity

$$\mathrm{C}(x, y, z) - \frac{1}{2}\big(I(x : y \mid z) + I(x : z \mid y) + I(y : z \mid z)\big) - I(x : y : z) + O(\delta) + O(\log n). \tag{8}$$

The communication complexity (7) from Theorem 1(i) is known to by asymptotically optimal, see [15]. In this paper we study the communication complexity of the problem from Theorem 2. In fact, (8) is *not* optimal for general communication protocols; however, we show that this communication complexity is asymptotically optimal in the class of protocols with *simultaneous messages*, i.e., in the model where Alice, Bob, and Charlie send their messages in parallel, receive the messages sent by their vis-a-vis, and compute the result (secret key) without any further interaction.

Theorem 3 (main result)**.** *In the setting of Theorem 2, communication complexity of a protocol with simultaneous messages (the total number of bits sent by Alice, Bob, and Charlie) for triples of inputs (x, y, z) with a symmetric complexity profile* (3)*) cannot be smaller than*

$$\mathrm{C}(x, y, z) - \frac{1}{2}\big(I(x : y \mid z) + I(x : z \mid y) + I(y : z \mid z)\big) - I(x : y : z) - O(\delta) - O(\log n). \tag{9}$$

[2] The scheme proposed in [14] is the so called *omniscience* protocol. In this protocol, all parties send simultaneously their messages (random hash-values of the inputs) so that each of them learns completely the entire triple of inputs (x, y, z) (this explains the term *omniscience*). The total length of the sent messages is less than $\mathrm{C}(x, y, z)$, so an eavesdropper can learn only a partial information on the inputs. The gap between the total complexity of $\mathrm{C}(x, y, z)$ and the divulged information is used to produce a secret key.

Communication complexity (9) is not optimal for general (multi-round) communication protocols of secret key agreement, see Proposition 2.

The proof of our main result combines information-theoretic techniques and spectral bounds for graphs (the expander mixing lemma). Spectral bounds *per se* are not new in communication complexity (see, e.g., the usage of Lindsey's lemma in [27]). Information-theoretic methods are also pretty common in this area. But the combination of these two techniques seems to be less standard. The key step of the proof is the observation that in some setting, when parties hold correlated data sets, for each of them it is hard to send a message that has non-negligible mutual information with the partners' data. In other words, a "too short" message sent by Alice would have zero mutual information with the data (y, z) given to Bob and Charlie. For secret key agreement protocols, this observation implies that the messages of every party inevitably have to be quite long. A similar argument can be used in problems that are not connected with cryptography, see Theorem 6.

The rest of the paper is organized as follows. In Section 2 we recall several standard definitions and introduce the notation. In Section 3 we explain informally the scheme of our argument. In Section 4 we prove the main technical tool of this paper, Theorem 5 (which claims that in some setting, it is hard to send a message that has non-negligible mutual information with the partners' data). In Section 5 we illustrate the application of our technique with a simple example that is not related to cryptography. In Section 6 we prove Theorem 3 for a restricted ("the most important") class of complexity profiles; this is the main technical contribution of the paper. In Section 7 we extend this result and prove Theorem 3 for all (symmetric) complexity profiles. We conclude with a discussion of limitations of our technique and open problems. Several technical lemmas are deferred to Appendix.

2 Preliminaries and Notation

2.1 General notation.

For a binary string x we denote its length $|x|$. For a finite set S we denote its cardinality $\#S$.

In what follows we manipulate with equalities and inequalities for Kolmogorov complexity. Since many of them hold up to a logarithmic term, we use the notation $A \stackrel{\lg}{=} B$, $A \stackrel{\lg}{\leq} B$, and $A \stackrel{\lg}{\geq} B$ for $|A - B| = O(\log n)$, $A \leq B + O(\log n)$, and $B \leq A + O(\log n)$ respectively, where n is clear from the context (n is usually the length of the strings involved in the inequality).

$\mathbb{F}_q$ denotes the field of q elements (usually $q = 2^n$). A k-dimensional vector over $\mathbb{F}_q$ is a k-tuple $(x_1, \ldots, x_k) \in \mathbb{F}_q^k$. We say that two vectors $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ in $\mathbb{F}_q^k$ are orthogonal to each other if $x_1y_1 + \ldots + x_ky_k = 0$ (the addition and multiplication are computed in the field $\mathbb{F}_q$). A vector is called self-orthogonal if it is orthogonal to itself. In a k-dimensional space over the field of characteristic 2 there are 2^{k-1} self-orthogonal vectors $(x_1, \ldots, x_k)$ and they form a linear subspace of co-dimension 1 (a vector is self-orthogonal iff $x_1 + \ldots + x_k = 0$). A *direction* in $\mathbb{F}_q^k$ is an equivalence class of non-zero vectors over $\mathbb{F}_q$ that are proportional to each other (a direction can be understood as a point in the projective space of dimension $k-1$). Two directions are orthogonal to each other if every vector in the first one is orthogonal to every vector in the second one.

$\mathrm{C}(x)$ stands for Kolmogorov complexity of x (the length of the shortest program[3] producing x) and $\mathrm{C}(x \mid y)$ (the length of the shortest program producing x given input y) stands for Kolmogorov complexity of x given y. Respectively, $\mathrm{I}(x : y)$ and $\mathrm{I}(x : y \mid z)$ denote the mutual information between x and y and the conditional information between x and y given z. We use the notation $\mathrm{I}(x : y : z) := \mathrm{I}(x : y) - \mathrm{I}(x : y \mid z)$. For a tuple of strings $(x_1, \ldots, x_n)$ its *complexity profile* is the vector consisting of the complexity values $\mathrm{C}(x_{i_1}, \ldots, x_{i_s})$ (for all $2^n - 1$ sub-tuples $1 \leq i_1 < \ldots < i_s \leq n$).

Kolmogorov complexity can be relativized: $\mathrm{C}^{\mathcal{O}}(x)$ and $\mathrm{C}^{\mathcal{O}}(x \mid y)$ stand for Kolmogorov complexity of x (conditional on y) assuming that the universal decompressor can access oracle $\mathcal{O}$. If the oracle is a finite string s, then $\mathrm{C}^{\mathcal{O}}(x) = \mathrm{C}(x \mid s) + O(1)$.

For more detail on the basic facts about Kolmogorov complexity see Appendix. A comprehensive introduction in the theory of Kolmogorov complexity can be found in [13] and [16].

2.2 Communication complexity.

We use the conventional notion of a communication protocol for two or three parties, see for detailed definitions [21]. We discuss *deterministic protocols* and *randomized protocols with a public source of random bits* (see

[3] In an optimal programming language, see Appendix for more detail.

Appendix for more detail).

In general, a communication protocol may consist of several rounds, when each next message of every party depends on the previously sent messages. In the *simultaneously messages* model there is no interaction: all parties send in parallel their messages that depend only on their own input data (and the random bits), and then compute the final result.

We usually denote the inputs of Alice, Bob, and Charlie as x, y, and z respectively (*number-in-hand* model). A deterministic communication protocol for inputs $x, y, z \in \{0,1\}^n$ returns a result $w = w(x, y, z)$. In a randomized protocol the result depends also on the public source of random bits r, and $w = w(x, y, z, r)$. The sequence of messages sent by the parties to each other while following the steps of the protocol is called a *transcript* $t = t(x, y, z)$ of the communication ($t = t(x, y, z, r)$ for randomized protocols). Communication complexity of a protocol is the maximal length of its transcript (measured in bits), i.e., $\max\limits_{x,y,z,r} |t(x, y, z, r)|$.

A communication protocol *computing a function* $F(x, y, z)$ returns a correct result if $w(x, y, z, r) = F(x, y, z)$. For a *secret key agreement* protocol, the definition of a *correct result* w is subtler: we need that (i) w is of the required size and (ii) it is almost incompressible even given the transcript of the communication t and the public random bits r. For a more detailed discussion of this setting we refer the reader to [14].

We will assume that the communication protocol has a "uniform" description. More technically, we assume that for n-bit inputs (the full description of such a protocol) has an efficient description of size $O(\log n)$. For such a protocol we do not loose much security even if the description of the protocol is available to the eavesdropper. Thus, we cannot "cheat" by embedding in the structure of the protocol any secret information hidden from the adversary.

If the length of inputs is equal to n, we may assume w.l.o.g. that the used string of public random bits r is of length $O(n)$. (Using longer sources of random bits may only slightly affect the probability of an erroneous result. For protocols computing a function, $O(\log n)$ bits is enough due to the Newman's theorem, [26]; in secret key agreement protocols we may need $O(n)$ random bits, see [15]). This is crucial for our setting: we may assume that the terms $O(\log \mathrm{C}(r))$ involved in inequalities for Kolmogorov complexity match in order of magnitude the terms $O(\log(\mathrm{C}(x) + \mathrm{C}(y) + \mathrm{C}(z)))$, where x, y, z are the input data.

2.3 Reminder of the spectral graph technique.

Let $G = (L \cup R, E)$ be a bi-regular bipartite graph where each vertex in L has degree D_L, each vertex in R has degree D_R, and each edge $e \in E$ connects a vertex from L with a vertex from R (observe that $\#E = \#L \cdot D_L = \#R \cdot D_R$). The adjacency matrix of such a graph is a zero-one matrix $M = \begin{pmatrix} 0 & A \\ A^\top & 0 \end{pmatrix}$ where A is a matrix of dimension $(\#L) \times (\#R)$ ($A_{xy} = 1$ if and only if there is an edge between the x-th vertex in L and the y-th vertex in R). Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ be the eigenvalues of M, where $N = \#L + \#R$ is the total number of vertices. Since M is symmetric, all λ_i are real numbers. It is well known that for a bipartite graph the spectrum is symmetric, i.e., $\lambda_i = -\lambda_{N-i+1}$ for each i, and $\lambda_1 = -\lambda_N = \sqrt{D_L D_R}$ (see, e.g., [17]). The graphs with a large *spectral gap* (the gap between the first and the second eigenvalues) have the property of good *mixing*, see [20].

Lemma 1 (Expander Mixing Lemma for bipartite graphs, see [17])**.** *Let $G = (L \cup R, E)$ be a regular bipartite graph where each vertex in L has degree D_L and each vertex in R has degree D_R. Then for each $A \subseteq L$ and $B \subseteq R$ we have $\left| E(A,B) - \frac{D_L \cdot \#A \cdot \#B}{\#R} \right| \leq \lambda_2 \sqrt{\#A \cdot \#B}$, where λ_2 is the second largest eigenvalue of the adjacency matrix of G and $E(A,B)$ is the number of edges between A and B.*

We apply Lemma 1 for the case $\frac{D_L \cdot \#A \cdot \#B}{\#R} \geq \lambda_2 \sqrt{\#A \cdot \#B}$, as shown in the corollary below.

Corollary 1. *Let $G = (L \cup R, E)$ be a graph from Lemma 1 with the second eigenvalue λ_2. Then for $A \subseteq L$ and $B \subseteq R$ such that $\#A \cdot \#B \geq \left(\frac{\lambda_2 \#R}{D_L}\right)^2$ we have*

$$E(A,B) = O\left(\frac{D_L \cdot \#A \cdot \#B}{\#R}\right). \tag{10}$$

To apply the expander mixing lemma, we need a graph with a large spectral gap. In particular, we use the following well know lemma.

Lemma 2 (see [37]). *Let $G = (L \cup R, E)$ be a graph where L consists of all lines in a projective plane over a finite field $\mathbb{F}_q$, R consists of all points in this plane, and E consists of all incident pairs $(line, point)$. In this graph $\lambda_1 = \Theta(q)$ and $\lambda_2 = O(\sqrt{q})$.*

In the next section we discuss spectral properties of graphs with a more complicated structure.

3 Main technical tools and the scheme of the proof

In this section we sketch the proof of our main result (Theorem 3). In this sketch we ignore technical difficulties that can be resolved with standard techniques or *ad hoc* tricks, and focus on the main ideas used in the proof.

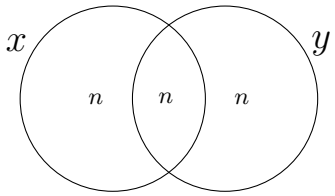
3.1. Setting the parameters. Let us assume that $\delta = O(\log n)$, i.e., all parties of the protocol "know" the complexity profile of the triple of inputs (x, y, z) up to an additive logarithmic term[4]. This assumption does not affects significantly the argument, but it helps to avoid minor technical details and makes the explanation more transparent. To simplify the notation, in this section we discuss only triples of inputs with the profile

$$\begin{aligned} &\mathrm{C}(x) \stackrel{\lg}{=} \mathrm{C}(y) \stackrel{\lg}{=} \mathrm{C}(z) \stackrel{\lg}{=} kn,\ \mathrm{C}(x,y) \stackrel{\lg}{=} \mathrm{C}(x,z) \stackrel{\lg}{=} \mathrm{C}(y,z) \stackrel{\lg}{=} (2k-1)n, \\ &\mathrm{C}(x,y,z) \stackrel{\lg}{=} (3k-3)n \end{aligned} \tag{11}$$

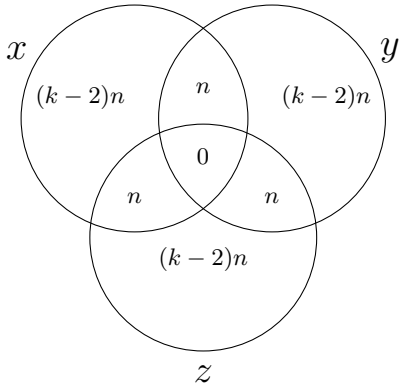
which is equivalent to

$$\begin{aligned} &\mathrm{C}(x \mid y,z) \stackrel{\lg}{=} \mathrm{C}(y \mid x,z) \stackrel{\lg}{=} \mathrm{C}(z \mid x,y) \stackrel{\lg}{=} (k-2)n, \\ &I(x:y \mid z) \stackrel{\lg}{=} I(x:z \mid y) \stackrel{\lg}{=} I(y:z \mid x) \stackrel{\lg}{=} n,\ I(x:y:z) \stackrel{\lg}{=} 0 \end{aligned}$$

see Fig. 1 (b). In this setting, Theorem 2 gives the optimal size of a secret key



(a) Complexity profile for Examples 2 and 3: $\mathrm{C}(x \mid y) \stackrel{\lg}{=} n$, $\mathrm{C}(y \mid x) \stackrel{\lg}{=} n$, $\mathrm{I}(x:y) \stackrel{\lg}{=} n$.



(b) Complexity profile for Proposition 1: $\mathrm{C}(x \mid y,z) \stackrel{\lg}{=} (k-2)n$, $\mathrm{I}(x:y) \stackrel{\lg}{=} n$, $\mathrm{I}(x:yz) \stackrel{\lg}{=} 2n$, $\mathrm{I}(x:y:z) \stackrel{\lg}{=} 0$.

Figure 1: Diagrams with complexity profiles for Examples 2-3 and Proposition 1.

$$\frac{1}{2}\big(\mathrm{I}(x:y \mid z) + I(x:z \mid y) + I(y:z \mid x)\big) + \mathrm{I}(x:y:z) \stackrel{\lg}{=} 1.5n. \tag{12}$$

Our aim is to bound communication complexity for inputs with this complexity profile:

Theorem 4 (special case of Theorem 3). *In the setting of Theorem 2, communication complexity of a protocol with simultaneous messages (the total number of bits sent by Alice, Bob, and Charlie) for some triples of inputs (x, y, z) with complexity profile* (11) *cannot be smaller than $(3k - 4.5)n$, which matches Eq.* (9).

[4] A logarithmic error term is, in some sense, the finest meaningful precision for Kolmogorov complexity. All our arguments can be repeated *mutatis mutandis* for any coarser precision δ such that $\log n \ll \delta(n) \ll n$.

3.2. Preliminary consideration: the need for hard inputs. The optimal size of the secret key in Theorem 1 and Theorem 2 depends only on the complexity profile of (x, y, z) and not on the combinatorial structure of the input. The situation with communication complexity (the number of bits sent by the parties) is different: it may vary significantly for different tuples of inputs with the same complexity profile. When we talk about the communication complexity of a protocol, we mean the worst-case complexity, i.e., the maximal number of sent bits among all admissible inputs. To prove a lower bound for the worst-case communication complexity, we need to provide a triple of inputs for which the parties have to send long messages. We provide a class of inputs that are guaranteed to be "hard" (for all valid protocol, for most triples of inputs from this class, communication complexity is high).

3.3. First step of the argument: conditional on Charlie's message, the mutual information between Alice's and Bob's inputs must increase. We begin with an observation that might seem to have nothing to do with communication complexity. We recall the lower bound for the size of the secret key (that applies to protocols with any communication complexity). In [14] (see Theorem 1(ii)) it is shown that two parties, Alice and Bob, can agree on secret key of complexity k *only if* the mutual information between Alice's input x and Bob's input y is greater than k. The proof of this statement can be easily adapted to the following slightly more general setting:

Lemma 3. *Assume that there is a publicly available information s (accessible to Alice, Bob, and the eavesdropper), and besides this information Alice is given a private input x and Bob is given a private input y. Then, by communication via a public channel accessible to the eavesdropper, Alice and Bob cannot agree on a secret key of complexity greater than $I(x : y \mid s)$.*

We apply this proposition to a protocol with three parties. Let t_C denote the concatenation of the messages sent by Charlie. This is a piece of publicly available information (accessible to Alice, Bob, and the eavesdropper). Due to Lemma 3, Alice and Bob cannot agree on a secret key with Kolmogorov complexity greater than $I(x : y \mid t_C)$ (at this point we ignore whether Charlie can learn the same key or not). Hence, in the settings (11), a secret key of size (12) can be produced only if $I(x : y \mid t_C) \geq^{\lg} 1.5n$. Observe that in the setting (11) the mutual information between x and y is equal to n. This means that the mutual information between Alice's and Bob's inputs *conditional on Charlie's message*, i.e., $I(x : y \mid t_C)$, is bigger than the unconditional mutual information between Alice's and Bob's inputs, i.e., $I(x : y)$. A pretty standard information-theoretic argument implies that the gap between $I(x : y)$ and $I(x : y \mid t_C)$ is not greater than the mutual information between $\langle x, y\rangle$ and t_C, and we conclude that $I(x, y : t_C) \geq^{\lg} n/2$. In other words, Charlie must send a message t_C that has $\geq n/2$ bits of mutual information with the pair of inputs of Alice and Bob. A similar argument implies that Alice must send a message t_A such that $I(y, z : t_A) \geq^{\lg} n/2$ and Bob must send a message t_B such that $I(x, z : t_B) \geq^{\lg} n/2$.

This part of the argument is based on Lemma 3, which re-employs an argument from [14] in a pretty direct way. So at this stage we need no substantially new ideas.

3.4. Second step of the argument: it may be difficult for Alice to send a message increasing the mutual information between Bob's and Charlie's inputs. We have shown above that in the setting (11) Alice, Bob, and Charlie can agree on a secret key of optimal size only if each of them sends a messages that contains $\geq^{\lg} n/2$ bits of mutual information with the inputs of two other parties

We are going to show that this may require sending *very long* messages (much longer than $n/2$ bits). This part of the argument is the main technical contribution of our paper. To explain this idea, we make a digression and discuss a similar problem in simpler settings.

Digression: how to say something that the interlocutor already knows. Let us consider randomized communication protocols with two participants playing non-symmetric roles. We call the participants Speaker and Listener and assume that Speaker holds an input string a and Listener holds another input string b. This is a one-way protocol: Speaker sends a message to Listener in one round, without any feedback. The aim of Speaker is to send to Listener a message that is *not completely unpredictable* from the point of view of Listener. More precisely, Speaker's message must have positive (and non-negligible) mutual information with Listener's input b. We start with a simple example when the task of Speaker is trivial.

Example 2. Let Speaker is given a string $a = uv$ and Listener is given a string $b = uw$, where u, v, and w are independent incompressible strings of length n, i.e., $\mathrm{C}(uvw) \stackrel{\lg}{=} \mathrm{C}(u) + \mathrm{C}(v) + \mathrm{C}(w) \stackrel{\lg}{=} 3n$. Observe that

$$\mathrm{C}(a) \stackrel{\lg}{=} 2n,\ \mathrm{C}(b) \stackrel{\lg}{=} 2n,\ \mathrm{I}(a : b) \stackrel{\lg}{=} n \tag{13}$$

(see the diagram in Fig. 1 (a)). In this setting, if Speaker wants to communicate a message of length n with a *high* mutual information with Listener's y, she may send a part of u, which is know to both participants of the protocol.

On the other hand, if Speaker wants to communicate a message with a *low* mutual information with Listener's b, this is also possible: Speaker may send a part of v, which is know to Speaker but not to the Listener.

Let us proceed with a less trivial example.

Example 3. Now we consider a pair (a, b) with the same complexity profile as in Example 2 but with a different combinatorial structure. Let a be a line in the projective plane over the finite field $\mathbb{F}_{2^n}$ and b be a point in the same projective plane incident to a, and the pair (a, b) have the maximal possible complexity (among all incident pairs (line, point) in the plane). For these a and b we have the same complexity profile (13). Indeed, we need two elements of the field ($2n$ bits of information) to specify a line or a point, but we need only one element of the field (n bits of information) to specify a point when a line is known. However, the combinatorial properties of this pair are very different from the properties of the pair in Example 2.

If Speaker is given a and Listener is given b as above, then Speaker cannot send a *reasonably short* message having non-negligible mutual information with Listener's input b. In fact, if Speaker wants to send to Listener a message $m = m(a)$ having δ bits of mutual information with b, then the size of m must be at least $n + \delta$. In particular, if the message m is shorter than n, then it cannot contain any information on b. We prove this statement in Section 4.

Example 3 is an instance of a much more general phenomenon. Let us have a bipartite graph $G = (V_L, V_R, E)$, where the set of vertices is $V_L \cup V_R$ and the set of edges is $E \subset V_L \times V_R$. We assume that the graph is bi-regular, i.e., all vertices in V_L have the same degree D_L and all vertices in V_R have the same degree D_R (we always assume that $D_L \geq D_R$). We say that G is a *spectral expander*[5] if the second eigenvalue of its adjacency matrix $\lambda_2 = O(\sqrt{D_L})$. Let $(x, y) \in E$ be a "typical" edge of this graph (in the sense that its Kolmogorov complexity is close to the maximum possible value), and let x and y be the inputs given to Alice and Bob respectively. Then we have a property similar to Example 2: if Alice wants to send a message having δ bits of mutual information with Bob's data y, she must send a message of size at least $\log D_R + \delta$. We prove this fact using the Expander Mixing Lemma. (Example 3 corresponds to the graph $G = (V_L, V_R, E)$ where V_L consists of all lines in the plane, V_R consists of all points in the plane, and E is the set of all pairs of incident lines and points; it is known that this graph is a spectral expander.) [End of *Digression*.]

Now we generalize the observations from the *Digression* above and explain the main idea of the proof of Theorem 4. To explain the principal construction, we introduce the notion of a tri-expander hypergraph, which extends the conventional definition of a bipartite expander.

Definition 1. Let $G = (V_1, V_2, V_3, H)$ be a hypergraph where

- the set of vertices consists of three disjoint parts V_1, V_2, V_3 of the same cardinality
- the set of hyperedges is a set $H \subset V_1 \times V_2 \times V_3$.

We consider three bipartite graphs G_1, G_2, G_3 associated with hypergraph G: each G_i is a bipartite graph $(V_i, V_j \times V_\ell, E_i)$ (here $j = i + 1 \mod 3$ and $\ell = i + 2 \mod 3$), where $(x, \langle y, z\rangle) \in E_i$ if and only if the triple $\{x, y, z\}$ corresponds to a hyperedge in H. The hypergraph is called *tri-expander* if the graphs G_1, G_2, G_3 are bi-regular spectral expanders.

Remark 5. The definition of a tri-expander and an application of the expander mixing lemma to the associated bipartite graphs (see below) seems to be similar but not literally equivalent to the definition of the second eigenvalue for 3-uniform hypergraph and the hypergraph generalization of the expander mixing lemma in [18].

We show that the communication is costly for a triple of inputs (x, y, z) that is a hyperedge in a tri-expander. To this end, we combine the idea from paragraph 3.3 with an argument similar to the observation sketched in the *Digression*: each party must send a message having non-negligible mutual information with two other inputs (an information-theoretic argument) but this is only possible when each of the messages is very long (due to the spectral bound and the expander mixing lemma).

3.5. Construction of a tri-expander. To conclude the proof of the main result it remains to show that there exists a tri-expander with suitable parameters:

Proposition 1. *For all integer numbers $k \geq 0$ and $n \geq 1$ there exists a tri-expander $G = (V_1, V_2, V_3, H)$ such that*

[5] We use the term *expander* without assuming that the degree of a graph is constant.

- $\#V_1 = \#V_2 = \#V_3 = \Theta(2^{kn})$,
- *for all $i \neq j$, for every $x \in V_i$ there exists $\Theta(2^{(k-1)n})$ vertices $y \in V_j$ such that x and y are adjacent in the hypergraph,*
- $\#H = \Theta(2^{kn} \cdot 2^{(k-1)n} \cdot 2^{(k-2)n})$.

Proof. We construct such a tri-expander explicitly. We fix the finite field $\mathbb{F}_{2^n}$ with $q = 2^n$ elements, the $(k+2)$-dimensional space $\mathcal{L}$ over this field, and the subspace $\mathcal{L}_{so} \subset \mathcal{L}$ that consists of self-orthogonal vectors. Observe that $\#\mathcal{L}_{so} = \#\mathcal{L}/q = q^{k+1}$ (a subspace of co-dimension 1 in $\mathcal{L}$). Let V denote the space of all *directions* in $\mathcal{L}_{so}$ except for the direction $(1, \ldots, 1)$ (which is self-orthogonal for even k). Observe that $\#V = \Theta(q^k)$.

We let $V_1 = V_2 = V_3 = V$ and define H as the set of all triple $(x, y, z) \in V^3$ such that x, y, z are *distinct and pairwise orthogonal* directions in $\mathcal{L}_{so}$.

For every vector $x \in \mathcal{L}_{so}$, the condition of being orthogonal to x determines in $\mathcal{L}_{so}$ a subspace of co-dimension 1; this subspace consists of q^k vectors (including x itself as it is self-orthogonal) and, respectively, $(q^k-1)/(q-1)$ directions (again, including the direction collinear with x). If we have two non-collinear vectors $x, y \in \mathcal{L}_{so}$, then the condition of being orthogonal to x and y determines in $\mathcal{L}_{so}$ a subspace of co-dimension 2; this subspace consists of q^{k-1} vectors (including x and y), which corresponds to $(q^{k-1} - 1)/(q - 1) = \Theta(q^{k-2})$ directions (once again, including the directions collinear with x and with y).

Thus, we have $\Theta(q^k)$ individual vertices, $\Theta(q^k \cdot q^{k-1})$ pairs of adjacent vertices, and $\Theta(q^k \cdot q^{k-1} \cdot q^{k-2})$ adjacent triples (hyperedges). It remains to compute the eigenvalues of the associated bipartite graphs.

Lemma 4. *The hypergraph $G = (V_1, V_2, V_3, H)$ defined above is a tri-expander.*

(This fact might be known, but for lack of a reference we give a proof in Appendix B.) In the proof we use rich symmetries of this hypergraph. To guarantee these symmetries, we have imposed the restrictions that may seem artificial: the characteristic of the field is 2, we take into consideration only self-orthogonal vectors, the direction $(1, \ldots, 1)$ is excluded from V. □

Remark 6. A standard counting shows that for most hyperedges (x, y, z) in the graph from Proposition 1 we have
$\mathrm{C}(x) \stackrel{\lg}{=} \log \Theta(q^k) \stackrel{\lg}{=} kn$, $\mathrm{C}(x, y) \stackrel{\lg}{=} \log \Theta(q^k \cdot q^{k-1}) \stackrel{\lg}{=} (2k-1)n$,
$\mathrm{C}(x, y, z) \stackrel{\lg}{=} \log \Theta(q^k \cdot q^{k-1} \cdot q^{k-2}) \stackrel{\lg}{=} (3k-2)n$, and we get the profile (11).

4 When it is hard to say anything that the interlocutor already knows

In this section we explain our main technical tool. We consider randomized communication protocols with two participants, Speaker and Listener. We assume that Speaker holds an input string a and Listener holds another input string b; we assume also that the complexity profile of the pairs (a, b) is known to all parties. The aim of Speaker in this protocol is to send to Listener a message that has non-negligible mutual information with Listener's input b, as we discussed in Section 3.

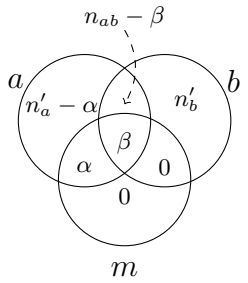


Figure 2: The profile in Theorem 5.

Theorem 5. *Let $G = (V_L, V_R, E)$ be a bipartite spectral expander such that $N = \#V_L$, $M = \#V_R$, and (D_L, D_R) are the degrees of the edges in V_L and V_R respectively. Let $(a, b) \in E$ be a "typical" edge in the graph, i.e., $\mathrm{C}(a, b) \stackrel{\lg}{=} \log \#E$, and $\mathrm{C}(m \mid a) \stackrel{\lg}{=} 0$. Then $\mathrm{I}(m : b) \leq^{\lg} \max\{0, \mathrm{C}(m) - \mathrm{C}(a \mid b)\}$. In particular, if the length of m is less than $\mathrm{C}(a \mid b)$, then $\mathrm{I}(m : b) \stackrel{\lg}{=} 0$.*

Remark 7. The statement of Theorem 5 remain valid if we relativize all terms of Kolmogorov complexity in this statement conditional on a string r such that $\mathrm{I}(r : (a, b)) \stackrel{\lg}{=} 0$. In what follows we present the proof without r. But every step of this argument trivially relativizes conditional on r assuming that $\mathrm{C}(a, b \mid r) \stackrel{\lg}{=} \mathrm{C}(a, b) \stackrel{\lg}{=} \log \#E$, we only need to add routinely the random bit string r to the condition of all terms with Kolmogorov complexity appearing in the proof.

Proof of Theorem 5. We denote $n_a := \log N$, $n_b = \log M$, $n'_a = \log D_R$, $n'_b = \log D_L$, and $n_{ab} := n_a - n'_a$. Using this notation, we have

$$\mathrm{C}(a) \stackrel{\lg}{=} n_a,\ \mathrm{C}(b) \stackrel{\lg}{=} n_b, \mathrm{C}(a \mid b) \stackrel{\lg}{=} n'_a,\ \mathrm{C}(b) \stackrel{\lg}{=} n_b,\ \mathrm{C}(b \mid a) \stackrel{\lg}{=} n'_b,\ \mathrm{I}(a : b) \stackrel{\lg}{=} n_{ab}.$$

Since Speaker computes the message m given the input data a, we have $\mathrm{C}(m \mid a) \stackrel{\lg}{=} 0$. We denote $\alpha := \mathrm{I}(m : a \mid b)$ and $\beta := \mathrm{I}(m : a : b)$. It is easy to verify that $\mathrm{C}(m) = \alpha + \beta$. The complexity profile for the triple (a, b, m) is shown in Fig. 2.

Case 1. Assume that $\mathrm{C}(m) \le n'_a - 2 \cdot \mathsf{const} \cdot \log n$ for some $\mathsf{const} > 0$ (a constant to be specified later). In this case, to prove the theorem, we need to show that $\mathrm{I}(m : y) \stackrel{\lg}{=} 0$. In our notation this is equivalent to $\beta \stackrel{\lg}{=} 0$. More technically, we are going to show that

$$\beta \le \mathsf{const} \cdot \log n. \tag{14}$$

For the sake of contradiction we assume that (14) is false. It is enough to consider the case when β is *somewhat large* but not *too large*, i.e., just slightly above the threshold (14). Indeed, any communication protocol violating (14) can be converted in a different protocols with the same or a smaller value of α and with $\beta = \mathsf{const} \cdot \log n + O(1)$. To this end, we observe that by discarding a few last bits of Speaker's message m we make the protocol only simpler. So, we may replace the initial message m with the shortest prefix of the initial message that still violates (14). Thus, in what follows, we assume w.l.o.g. that

$$\mathsf{const} \cdot \log n < \beta \le \mathsf{const} \cdot \log n + O(1).$$

Let us define $A := \{a' \ :\ \mathrm{C}(a' \mid m) \le \mathrm{C}(a \mid m)\}$ and $B := \{b' \ :\ \mathrm{C}(b' \mid m) \le \mathrm{C}(b \mid m)\}$. We use the following standard claim:

Claim. $\#A = 2^{\mathrm{C}(a|m) \pm O(\log n)} = 2^{n_a - \alpha - \beta \pm O(\log n)}$ *and* $\#B = 2^{\mathrm{C}(b|m) \pm O(\log n)} = 2^{n_b - \beta \pm O(\log n)}$ (see, e.g. [14, Claim 4.7]).

From the claim we obtain $\#A \cdot \#B = 2^{n_a - \alpha - \beta + n_b - \beta \pm O(\log n)} = 2^{n_a + n_b - \mathrm{C}(m) - \beta \pm O(\log n)}$. Since $\mathrm{C}(m) \le n'_a - 2 \cdot \mathsf{const} \log n$ and $\beta < \mathsf{const} \log n + O(1)$, we conclude

$$\begin{aligned} n_a + n_b - \mathrm{C}(m) - \beta n \pm O(\log n) &\ge n_a + n_b - (n'_a - 2 \cdot \mathsf{const} \cdot \log n) - \mathsf{const} \cdot \log n - O(\log n) \\ &\ge n_{ab} + n_b + \mathsf{const} \cdot \log n - O(\log n) \ge n_{ab} + n_b. \end{aligned}$$

(To get the last inequality, we should choose the value of const in (14) so that $\mathsf{const} \cdot \log n$ majorizes the term $O(\log n)$ in the inequality above.) Thus, $\#A \cdot \#B \ge 2^{n_{ab} + n_b} = \frac{M^2}{D_L}$.

With the Expander Mixing Lemma (Corollary 1) we obtain

$$E(A, B) = O\left(\frac{D_L \cdot \#A \cdot \#B}{M}\right) = O\left(\frac{\#A \cdot \#B}{M / D_L}\right).$$

Now observe that given m and the numbers $\mathrm{C}(a \mid m)$ and $\mathrm{C}(b \mid m)$ we can enumerate the sets A and B and, therefore, we can describe (a, b) by the index of this edge in the list of all edges between A and B. The size of such an index is $\log E(A, B)$. Hence,

$$\begin{aligned} \mathrm{C}(a, b \mid m) \le^{\lg} \log E(A, B) &\le^{\lg} (n_a + n_b - \mathrm{C}(m) - \beta) - (n_b - n'_b) \\ &= n_a + n'_b - \mathrm{C}(m) - \beta = \mathrm{C}(a, b) - \mathrm{C}(m) - \beta, \end{aligned}$$

and $\mathrm{C}(a, b) \le^{\lg} \mathrm{C}(m) + \mathrm{C}(a, b \mid m) \le^{\lg} \mathrm{C}(a, b) - \beta$. The terms $O(\log n)$ hidden in the notation $\le^{\lg}$ and $\stackrel{\lg}{=}$ in this inequality do not depend on β. Thus, we get a contradiction if the constant in (14) is chosen large enough.

Case 2. Now we assume that $\mathrm{C}(m) = n'_a + \delta$ for an arbitrary δ. Denote by m' the prefix of m of length $(n'_a - \mathsf{const} \log n)$ and by m'' the suffix of m of length $(\delta + \mathsf{const} \log n)$. We know from Case 1 that $\mathrm{I}(m' : b) \stackrel{\lg}{=} 0$. It remains to apply the chain rule,

$$\mathrm{I}(m : b) \stackrel{\lg}{=} \mathrm{I}(m' : b) + \mathrm{I}(m'' : b \mid m') \stackrel{\lg}{=} \mathrm{I}(m'' : b \mid m') \le^{\lg} |m''| \stackrel{\lg}{=} \delta.$$

and the theorem is proven. □

From this theorem we obtain immediately the following corollary.

Corollary 2. *Let $G = (V_L, V_R, E)$ be a bipartite spectral expander such that $N = \#V_L$, $M = \#V_R$, and (D_L, D_R) are the degrees of the edges in V_L and V_R respectively.*

(a) We assume that Speaker and Listener are given, respectively, a and b that are ends of a typical edge $(a, b) \in E$ in the graph. We consider a one-round communication protocol where Speaker sends to Listener a message $m = m(a)$. Then $\mathrm{I}(m : b) \le^{\lg} \max\{0, \mathrm{C}(m) - \mathrm{C}(a \mid b)\}$. In particular, if the length of m is less than $\mathrm{C}(a \mid b)$, then $\mathrm{I}(m : b) \stackrel{\lg}{=} 0$.

(b) A similar statement is true if Speaker and Listener are given instead of a and b some inputs a' and b' such that $\mathrm{C}(a' \mid a) \stackrel{\lg}{=} 0$ and $\mathrm{C}(b' \mid b) \stackrel{\lg}{=} 0$ (e.g., if Speaker is given a function of a vertex $a \in V_L$ and Listener is given a function of a vertex $b \in V_R$).

5 Protocols with simultaneous messages : a warm-up example

In this section we use Theorem 5 from the previous section to prove a lower bound for communication complexity of the following problem. Alice and Bob hold, respectively, lines a and b in a plane (intersecting at one point c). They send to Charlie (in parallel, without interacting with each other) some messages so that Charlie can reconstruct the intersection point. We argue that the trivial protocol (where Alice and Bob send the full information on their lines) is essentially optimal.

Theorem 6. *Let Alice and Bob be given lines in the projective plane over the finite field $\mathbb{F}_{2^n}$ (we denote them a and b respectively), and it is known that the lines intersect at point c. Another participant of the protocol Charlie has no input information. Alice and Bob (without a communication with each other) send to Charlie messages m_A and m_B so that Charlie can find c, see Fig. 3. For every communication protocol for this problem, for some a, b we have $|m_A| + |m_B| \ge^{\lg} 4n$, which means essentially that in the worst case Alice and Bob must send to Charlie all their data (for a typical pair of lines we have $\mathrm{C}(a) + \mathrm{C}(b) \stackrel{\lg}{=} 4n$).*

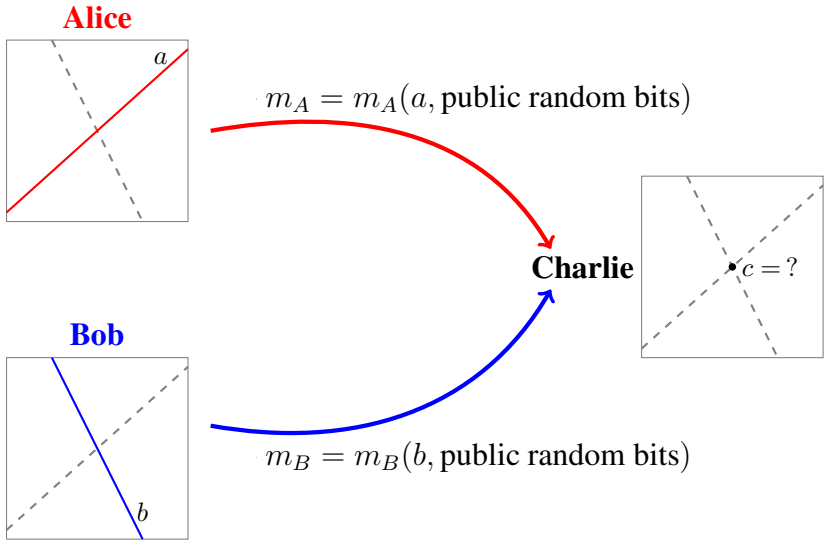


Figure 3: Alice holding a and Bob holding b send simultaneous messages to Charlie, who computes c.

In the setting of Theorem 6, the inputs of Alice and Bob contain n bits of the mutual information with c, so an easy lower bound for the communication complexity is $n + n = 2n$, see Fig. 4. However, due to the spectral properties of graphs implicitly present in this construction, the true communication complexity of this problem is twice bigger.

Sketch of the proof. In this sketch we ignore the public randomness and explain the argument for deterministic protocols. A generalization for protocols with public randomness is pretty straightforward, see the full proof below.

Let (a, b) be a pair of lines in a projective plane over $\mathbb{F}_{2^n}$ intersecting at a point c, such that $\mathrm{C}(a, b) \stackrel{\lg}{=} \mathrm{C}(a) + \mathrm{C}(b) \stackrel{\lg}{=} 4n$ (which is the case for most pairs of lines in the plane). Observe that $\mathrm{I}(a : c) \stackrel{\lg}{=} n$ and $\mathrm{I}(b : c) \stackrel{\lg}{=} n$. It

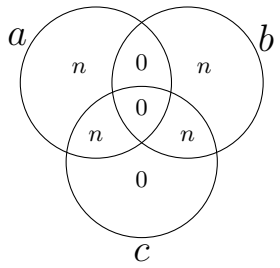


Figure 4: Complexity profile for two lines (a and b) and their intersection point c in the plane over $\mathbb{F}_{2^n}$.

follows that for the messages $m_A = m_A(a)$ and $m_B = m_B(b)$ we have $\mathrm{I}(m_A : c) \leq^{\lg} n$ and $\mathrm{I}(m_B : c) \leq^{\lg} n$. Using standard information theoretic inequalities, one can show that Alice's message m_A and Bob's message m_B determine the point c uniquely only if $\mathrm{I}(m_A : c) \stackrel{\lg}{=} n$ and $\mathrm{I}(m_B : c) \stackrel{\lg}{=} n$. Thus, Alice and Bob must send messages with large enough information on c.

The graph of possible pairs (a, c) and the graph of possible pairs (b, c) (the configurations (line, point)) is the same as in Example 3. Hence, we can apply Theorem 5 (Alice and Bob play the roles of Speaker, and Charlie plays the role of Listener) and conclude that $\mathrm{I}(m_A : c) \leq^{\lg} \max\{0, \mathrm{C}(m_A) - n\}$ and $\mathrm{I}(m_B : c) \leq^{\lg} \max\{0, \mathrm{C}(m_B) - n\}$. In particular, $\mathrm{I}(m_A : c) \stackrel{\lg}{=} n$ and $\mathrm{I}(m_B : c) \stackrel{\lg}{=} n$ only if Kolmogorov complexities of m_A and m_B are both at least $2n$. Thus, the total communication complexity is $\geq^{\lg} 2n + 2n = 4n$. □

In what follows we present the full proof of Theorem 6. The reader can skip it and proceed to the proof of the main result in the next section.

Full proof of Proof of Theorem 6. Let (a, b) be a pair of lines in a projective plane over $\mathbb{F}_{2^n}$ such that

$$\mathrm{C}(a, b) \stackrel{\lg}{=} \mathrm{C}(a) + \mathrm{C}(b) \stackrel{\lg}{=} 4n$$

(which is the case for most pairs of lines in the plane), and let c be the point of intersection of these lines, see Fig. 4.

Denote by r the string of random bits from the public source of randomness (accessible to Alice, Bob, Charlie, and to the eavesdropper). We assume r and the inputs (a, b) are independent, i.e., $\mathrm{I}(r : a, b) \stackrel{\lg}{=} 0$ (this is the case with an overwhelming probability). For such a string r all terms with Kolmogorov complexities involving a, b, c do not change if we add r in the condition:

$$\mathrm{C}(a, b, c \mid r) \stackrel{\lg}{=} \mathrm{C}(a, b, c) \stackrel{\lg}{=} \mathrm{C}(a, b),\ \mathrm{C}(a, c \mid r) \stackrel{\lg}{=} \mathrm{C}(a, c),\ \mathrm{C}(b, c \mid r) = \mathrm{C}(b, c),$$
$$\mathrm{C}(a \mid r) = \mathrm{C}(a),\ \mathrm{C}(b \mid r) = \mathrm{C}(b),\ \mathrm{C}(c \mid r) = \mathrm{C}(c).$$

This implies

$$\mathrm{C}(a \mid c, r) \stackrel{\lg}{=} \mathrm{C}(a, c \mid r) - \mathrm{C}(c \mid r) \stackrel{\lg}{=} \mathrm{C}(a, c) - \mathrm{C}(c) \stackrel{\lg}{=} \mathrm{C}(a \mid c),$$

and, similarly, $\mathrm{C}(b \mid c, r) \stackrel{\lg}{=} \mathrm{C}(b \mid c)$ and $\mathrm{C}(a, b \mid c, r) \stackrel{\lg}{=} \mathrm{C}(a, b \mid c)$.

Observe that the graph of possible pairs (a, c) and the graph of possible pairs (b, c) (the configurations (line, point) on the projective plane) is the same as in Example 3. Hence, we can apply Theorem 5 (Alice and Bob play the roles of Speaker, and Charlie plays the role of Listener; all terms are relativized conditional on r, see Remark 7) and conclude that

$$\mathrm{I}(m_A : c \mid r) \leq^{\lg} \max\{0, \mathrm{C}(m_A \mid r) - n\} \text{ and } \mathrm{I}(m_B : c \mid r) \leq^{\lg} \max\{0, \mathrm{C}(m_B \mid r) - n\}. \tag{15}$$

In particular, $\mathrm{I}(m_A : c \mid r) \geq^{\lg} n$ and $\mathrm{I}(m_B : c \mid r) \geq^{\lg} n$ only if Kolmogorov complexities of m_A and m_B are both at least $2n$.

It is easy to verify that $\mathrm{I}(a : c \mid r) \stackrel{\lg}{=} \mathrm{I}(a : c) \stackrel{\lg}{=} n$ and $\mathrm{I}(b : c \mid r) \stackrel{\lg}{=} \mathrm{I}(b : c) \stackrel{\lg}{=} n$. Since m_A and m_B are computed from (a, r) and (b, r) respectively, we conclude that

$$\mathrm{I}(m_A : c \mid r) \leq^{\lg} n \text{ and } \mathrm{I}(m_B : c \mid r) \leq^{\lg} n. \tag{16}$$

We need to show that these two inequalities turn into equalities. Indeed, by construction,

$$\begin{aligned}
\mathrm{I}(a:b\mid c,r) &\stackrel{\lg}{=} \mathrm{C}(a\mid c,r)+\mathrm{C}(b\mid c,r)-\mathrm{C}(a,b\mid c,r)\\
&\stackrel{\lg}{=} \mathrm{C}(a\mid c)+\mathrm{C}(b\mid c)-\mathrm{C}(a,b\mid c)\\
&\leq^{\lg} n+n-\mathrm{C}(a,b\mid c) \text{ [we need } n+O(1) \text{ bits to specify a line given a point]}\\
&\leq^{\lg} 2n-(\mathrm{C}(a,b)-\mathrm{C}(c))\\
&\leq^{\lg} 2n-4n+2n\stackrel{\lg}{=}0. \text{ [we need } 2n+O(1) \text{ bits to specify a point in the plane]}
\end{aligned}$$

As $\mathrm{C}(m_A\mid a,r)\stackrel{\lg}{=}0$ and $\mathrm{C}(m_B\mid b,r)\stackrel{\lg}{=}0$, we have (see Lemma 7(iii))

$$\mathrm{I}(m_A:m_B\mid c,r)\leq^{\lg}\mathrm{I}(a:b\mid c,r)\stackrel{\lg}{=}0.$$

Therefore, $\mathrm{I}(m_A:m_B:c\mid r)\stackrel{\lg}{=}\mathrm{I}(m_A:m_B\mid r)-\mathrm{I}(m_A:m_B\mid c,r)\geq^{\lg}0$, and

$$\mathrm{I}(m_Am_B:c\mid r)\stackrel{\lg}{=}\mathrm{I}(m_A:c\mid r)+\mathrm{I}(m_B:c\mid r)-\mathrm{I}(m_A:m_B:c\mid r)\leq^{\lg}\mathrm{I}(m_A:c\mid r)+\mathrm{I}(m_B:c\mid r).$$

On the other hand, since Charlie can compute c given the messages (m_A,m_B) and the string of random bits r, we have

$$\mathrm{I}(m_Am_B:c\mid r)\stackrel{\lg}{=}\mathrm{C}(c\mid r)-\mathrm{C}(c\mid m_Am_B,r)\stackrel{\lg}{=}2n-0\geq^{\lg}2n,$$

and, therefore,

$$\mathrm{I}(m_A:c\mid r)+\mathrm{I}(m_B:c\mid r)\stackrel{\lg}{=}2n.$$

Keeping in mind (16), we conclude that $\mathrm{I}(m_A:c\mid r)\stackrel{\lg}{=}n$ and $\mathrm{I}(m_B:c\mid r)\stackrel{\lg}{=}n$. Due to (15), this is possible only of $\mathrm{C}(m_A)\geq^{\lg}\mathrm{C}(m_A\mid r)\geq^{\lg}2n$ and $\mathrm{C}(m_B)\geq^{\lg}\mathrm{C}(m_B\mid r)\geq^{\lg}2n$. This means that the total length of the sent messages is at least $2n+2n=4n$ bits. □

6 Secret key agreement: a lower bound for the most crucial profile

In this section we prove a lower bound for communication complexity of secret key agreement with three parties. Let us recall the setting. We assume that Alice, Bob, and Charlie are given inputs x, y, z respectively with the complexity profile (11). as shown in Fig. 1(b). This is a pretty "generic" complexity profile; by choosing k, we control the gap between the complexities of x,y,z and the mutual informations shared by the inputs.

We consider communication protocols with public randomness. Denote by r the string of random bits accessible for all the parties (including the eavesdropper). We assume that Alice, Bob, and Charlie broadcast simultaneously messages $m_A=m_A(x,r),\ m_B=m_A(y,r),\ m_C=m_A(z,r)$ over a public communication channel. Then each of them computes the final result

$$\mathrm{key}_{\mathrm{Alice}}(x,r,m_B,m_C),\ \mathrm{key}_{\mathrm{Bob}}(y,r,m_A,m_C),\ \mathrm{key}_{\mathrm{Charlie}}(z,r,m_A,m_B).$$

We say that a protocol is successful if $\mathrm{key}_{\mathrm{Alice}}=\mathrm{key}_{\mathrm{Bob}}=\mathrm{key}_{\mathrm{Charlie}}=w$ (i.e., the parties agree on a common key w) and $\mathrm{C}(w\mid\langle m_A,m_B,m_C,r\rangle)\stackrel{\lg}{=}|w|$ (i.e., the eavesdropper gets no information on this key).

Theorem 2 claims that for any $\epsilon>0$ there exists a protocol that is successful with probability $(1-\epsilon)$, and the size of the key is equal to (6), which gives for the profile (11) the value $1.5n$. Moreover, this value of the key is optimal (up to an additive term $O(\log n)$).

It was shown in [14] that a secret key of this size can be obtained in an *omniscience* protocol. In this protocol, the parties broadcast messages so that each of them learns completely the entire triple of inputs (x,y,z). The total length of the broadcasted messages bits is less than $\mathrm{C}(x,y,z)$, so an eavesdropper can learn only a partial information on the inputs. More specifically, communication complexity of the omniscience protocol is (8), which is $(3k-4.5)n$ for a triple satisfying (11). The gap between $\mathrm{C}(x,y,z)\stackrel{\lg}{=}(3k-3)n$ and the amount of the divulged information is used to produce the secret key of size $1.5n$.

The omniscience protocol used in [14] provides an *upper bound* on the communication complexity of secret key agreement. In what follows we prove the matching *lower bound* (for protocols with simultaneous messages) and show that $(3k-4.5)n$ is the optimal communication complexity for a protocol of secret key agreement protocols with simultaneous messages for inputs satisfying (11). The proof follows the scheme sketched in Section 3. The first ingredient of this proof is Lemma 3 (see p. 7).

Sketch of proof of Lemma 3. This lemma is a relativized version of [14, Theorem 4.2], where s is used as an oracle. One can follow the argument from [14] step by step, substituting s as a supplementary condition in each term of Kolmogorov complexity appearing in the proof. □

Corollary 3. *Consider a communication protocol with three parties where Alice is given x, Bob is given y, and Charlie is given z. Denote by m_C the concatenation of all messages broadcasted by Charlie during the communication. If the parties agree on a secret key w on which the eavesdropper gets no information (even given access to the messages sent by all parties), then* $\mathrm{C}(w) \leq^{\lg} \mathrm{I}(x : y \mid r, m_C)$.

Proof. We apply Lemma 3 substituting m_C instead of the public information s. □

Now we are ready to prove our main result.

Theorem 4 rephrased. *Let Alice, Bob, and Charlie be given x, y, and z respectively such that (x, y, z) is a hyperedge of the hypergraph $G = (V_1, V_2, V_3, H)$ from Proposition 1 (the pairwise disjoint self-orthogonal directions in a $(k+2)$-dimensional vector space over $\mathbb{F}_{2^n}$). We consider non-interactive communication protocols where Alice, Bob, and Charlie send messages m_A, m_B, and m_C respectively and produce a secret key w with the optimal complexity* $\mathrm{C}(w) \stackrel{\lg}{=} 1.5n$. *Then* $\mathrm{C}(m_A) \geq^{\lg} (k-1.5)n$, $\mathrm{C}(m_B) \geq^{\lg} (k-1.5)n$, $\mathrm{C}(m_C) \geq^{\lg} (k-1.5)n$, *and the communication complexity of the protocol is at least $(3k-4.5)n - O(\log n)$, which matches the communication complexity of the omniscience protocol.*

Proof. To simplify the notation, we ignore the bits r provided by the public source of randomness and explain the proof for deterministic protocols. Our argument trivially relativizes given any instance of random bits r independent of (x, y) (which is true with a probability close to 1), cf. the full proof of Theorem 6.

From Corollary 3 we know that the size of the key (in our case $1.5n$) cannot be greater than $\mathrm{I}(x : y \mid m_C)$. By the construction of the tri-expander, $\mathrm{I}(x : y) \stackrel{\lg}{=} n$. Therefore, the difference between $\mathrm{I}(x : y)$ and $\mathrm{I}(x : y \mid m_C)$ is at least $0.5n$.

Lemma 5. *For all binary strings x, y, z it holds* $\mathrm{I}(x : y \mid s) - \mathrm{I}(x : y) \leq^{\lg} \mathrm{I}(s : xy)$.

(See the proof of the lemma in Appendix C.) We combine Corollary 3 with Lemma 5 and obtain $\mathrm{I}(m_C : xy) \geq^{\lg} 0.5n$.

Now we apply Theorem 5 to the bipartite graph G_3 associated with the tri-expander G (see p. 9); here Charlie plays the role of Speaker, and Alice and Bob together play the role of Listener. Since $\mathrm{I}(m_C : xy) \geq^{\lg} 0.5n$, we obtain $\mathrm{C}(m_C) \geq^{\lg} \mathrm{C}(z \mid x, y) + 0.5n \stackrel{\lg}{=} (k-1.5)kn$. A similar argument applies to $\mathrm{C}(m_A)$ and $\mathrm{C}(m_B)$, and we are done. □

7 Secret key agreement: a lower bound for all symmetric profiles

Proof of Theorem 3. If the complexity profile of (x, y, z) is symmetric then it can be specified by a triple of parameters α, β, γ,

$$\begin{cases} \mathrm{C}(x \mid y, z) \stackrel{\lg}{=} \mathrm{C}(y \mid x, z) \stackrel{\lg}{=} \mathrm{C}(z \mid x, y) \stackrel{\lg}{=} \alpha, \\ \mathrm{I}(x : y \mid z) \stackrel{\lg}{=} \mathrm{I}(x : z \mid y) \stackrel{\lg}{=} \mathrm{I}(y : z \mid x) \stackrel{\lg}{=} \beta,\ \mathrm{I}(x : y : z) \stackrel{\lg}{=} \gamma. \end{cases} \tag{17}$$

In Theorem 4 we proved that communication complexity (8) of the omniscience protocol is optimal in case $\alpha = (k-2)n$, $\beta = n$, and $\gamma = 0$. We reduce the problem with arbitrary α, β, γ to the special case settled in Theorem 4. We split this reduction into three steps, as shown in the following lemma.

Lemma 6. *If communication complexity* (8) *is optimal (in the worst case) for some triples of inputs (x, y, z) with complexity profile* (3) *then*

- *(a) for every positive $\delta \leq n$, communication of the omniscience protocol is also optimal (also in the worst case) for some triples of inputs (x', y', z') with complexity profile*

$$\begin{cases} \mathrm{C}(x' \mid y', z') \stackrel{\lg}{=} \mathrm{C}(y' \mid x', z') \stackrel{\lg}{=} \mathrm{C}(z' \mid x', y') \stackrel{\lg}{=} \alpha - \delta, \\ \mathrm{I}(x' : y' \mid z') \stackrel{\lg}{=} \mathrm{I}(x' : z' \mid y') \stackrel{\lg}{=} \mathrm{I}(y' : z' \mid x') \stackrel{\lg}{=} \beta,\ \mathrm{I}('x : y' : z) \stackrel{\lg}{=} \gamma, \end{cases} \tag{18}$$

- *(b) for every positive δ, communication of the omniscience protocol is also optimal for some triples of inputs (x', y', z') with complexity profile*

$$\begin{cases} \mathrm{C}(x' \mid y', z') \stackrel{\lg}{=} \mathrm{C}(y' \mid x', z') \stackrel{\lg}{=} \mathrm{C}(z' \mid x', y') \stackrel{\lg}{=} \alpha, \\ \mathrm{I}(x' : y' \mid z') \stackrel{\lg}{=} \mathrm{I}(x' : z' \mid y') \stackrel{\lg}{=} \mathrm{I}(y' : z' \mid x') \stackrel{\lg}{=} \beta,\ \mathrm{I}(x' : y' : z) \stackrel{\lg}{=} \gamma + \delta. \end{cases} \tag{19}$$

Let us consider the special case $\alpha \stackrel{\lg}{=} (k-2)n$, $\beta \stackrel{\lg}{=} n$, $\gamma \stackrel{\lg}{=} 0$ (as in Theorem 4). Then

- *(c) for every positive $\delta \le \beta/2$, communication of the omniscience protocol is also optimal for some triples of inputs (x', y', z') with complexity profile*

$$\begin{cases} \mathrm{C}(x' \mid y', z') \stackrel{\lg}{=} \mathrm{C}(y' \mid x', z') \stackrel{\lg}{=} \mathrm{C}(z' \mid x', y') \stackrel{\lg}{=} \alpha, \\ \mathrm{I}(x' : y' \mid z') \stackrel{\lg}{=} \mathrm{I}(x' : z' \mid y') \stackrel{\lg}{=} \mathrm{I}(y' : z' \mid x') \stackrel{\lg}{=} \beta + \delta,\ \mathrm{I}('x : y' : z) \stackrel{\lg}{=} -3\delta; \end{cases} \tag{20}$$

In Lemma 6 we show that the existence of a "too efficient" protocol for (19), (20), (18) would imply a "too efficient protocol" for (11), which is impossible due to Theorem 4. The proof is based on repeated application of Muchnik's theorem on conditional descriptions ([30]), which basically claims that for all strings $a, b_1, \dots, b_\ell$ and for every number $m \le \mathrm{C}(a)$ there exists a "digital fingerprint" of a of length m that looks maximally random conditional on each b_j. Technically, this means that for some a' we have

$$\mathrm{C}(a') \stackrel{\lg}{=} m,\ \mathrm{C}(a' \mid a) \stackrel{\lg}{=} 0, \text{ and } \mathrm{C}(a' \mid b_j) = \min\{\mathrm{C}(a' \mid b_j), m\} \text{ for } j = 1, \dots, \ell.$$

The proof of this lemma uses mostly techniques of Kolmogorov complexity that are not specific for communication problems, see Appendix E.

It is not hard to verify that starting with a triple (x, y, z) from Theorem 4 and then applying the reductions from Lemma 6, we can obtain any realizable profiles (3). Indeed, we begin with a triple of pairwise orthogonal directions (x, y, z) with $\alpha = (k-2)n, \beta = n, \gamma = 0$ for a suitable n and k, then apply Lemma 6 (b) or Lemma 6 (c) to get a triple (x', y', z') with a suitable $I(x' : y' : z')$ (case (b) serves to make the triple mutual information positive, and case (c) is needed if we want to make it negative), and further apply Lemma 6 (a) to trim the value of α.

Thus, Theorem 4 implies optimality of (8) not only for triples with a pretty specific complexity profile but for triples of inputs (x, y, z) with arbitrary symmetric complexity profile (3). □

8 Upper bound for interactive protocols

In this section we show that the communication complexity (8) is not optimal for multi-round protocols where the parties can actually interact with each other.

Proposition 2. *In the setting of Theorem 4 there is a* multi-round *communication protocol* (*not a* simultaneous messages *protocol*) *with communication complexity $(2k - 2.5)n + O(\log n)$, where the parties agree on a secret key of the optimal size $1.5n - O(\log n)$.*

Sketch of proof of Proposition 2. We adapt the *omniscience* protocol from [14]. In what follows we assume that random hash-functions are chosen with the public source of randomness (e.g., one may assume that random hashing is the multiplication by a randomly chosen matrix).

In the first round, Alice and Bob send messages $m_A = m_A(x, r)$ and $m_B = m_B(y, r)$ (random hash-values of x and y), each of length $(k - 1.5)n + O(\log n)$ such that Charlie given (m_A, m_B, z) can reconstruct the pair (x, y). Then Charlie sends a message m_C that is another random hash-value of (x, y) of length $0.5n + O(\log n)$.

With a high probability (for a randomly chosen hash-function), the values m_B and m_C are enough for Alice to reconstruct y, and the values m_A and m_C are enough for Bob to reconstruct x. Thus, at the end of communication, with high probability each party knows (x, y). At the same time, the adversary learns from the communication at most $|m_A| + |m_B| + |m_C| = (2k - 2.5)n + O(\log n)$ bits of information.

Now each party applies to (x, y) another (independently chosen) random hash function and obtains a hash-value $w = \mathrm{hash}(x, y)$ of length

$$\mathrm{C}(x, y) - |m_A| - |m_B| - |m_C| - O(\log n) = 1.5n - O(\log n).$$

With a high probability the obtained w is incompressible conditional on the data accessible to the eavesdropper (the messages of the parties and the public random bits). □

Remark 8. In the omniscience protocol, we may define random hashing as random linear mappings, i.e., each hash-values can be computed as the product over $\mathbb{F}_2$ of a bit vector by a randomly chosen binary matrix of the appropriate dimension. These matrices can be made publicly known: we can obtain these random matrices from the public source of random bits, and this does not reveal any information about the secret key to the adversary. Using more sophisticated constructions of hash-functions, we could reduce the number of used random bits (although this improvement is not necessary to prove Theorem 2 in the model with a public source of randomness).

9 Conclusion and open problems

We proved that the standard omniscience protocol provides the optimal worst-case communication complexity of the problem of secret key agreement (with three parties) in the class of protocols with simultaneous messages. A general open problem is to study the limits of our approach. In particular, for the class of multi-round communication protocols, the value (8) is no longer the optimal communication complexity of secret key agreement (Theorem 2). Our technique implies *some* lower bounds for communication complexity of interactive protocols, but it does not match the known upper bounds. Thus, a natural open problem is to settle the communication complexity of multi-party secret key agreement for multi-round protocols. It would be also interesting to extend our results to the communication model with private sources of randomness. Another open problem is to get rid of Lemma 6 and find a more direct proof of Theorem 3 with a more flexible construction of a tri-expander.

References

[1] Diffie, Whitfield; Hellman, Martin E. New Directions in Cryptography. IEEE Transactions on Information Theory. 22 (6): (1976) 644-654.

[2] Ralph C. Merkle. Secure Communications Over Insecure Channels. Communications of the ACM. 21 (4): (1978) 294-299.

[3] B. Smith. Pre-and post-quantum Diffie–Hellman from groups, actions, and isogenies. In Arithmetic of Finite Fields: 7th International Workshop, WAIFI 2018, Bergen, Norway, June 14-16, 2018, Revised Selected Papers 7 (pp. 3-40). Springer International Publishing.

[4] Bennett, C. H.; Bessette, F.; Brassard, G.; Salvail, L.; Smolin, J. (1992). Experimental Quantum Cryptography. Journal of Cryptology. 5 (1): 3-28.

[5] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. Proceedings of the Royal Society A: Mathematical, Physical and engineering sciences, 461(2053):207–235, 2005.

[6] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. Reviews of modern physics, 81(2):865, 2009.

[7] Tanya Ignatenko and Frans MJ Willems. Biometric security from an information-theoretical perspective. Foundations and Trends® in Communications and Information Theory, 7(2-3):135-316, 2012.

[8] Bloch, M., Günlü, O., Yener, A., Oggier, F., Poor, H.V., Sankar, L. and Schaefer, R.F., 2021. An overview of information-theoretic security and privacy: Metrics, limits and applications. IEEE Journal on Selected Areas in Information Theory, 2(1), pp.5-22.

[9] Dodis, Y., Kanukurthi, B., Katz, J., Reyzin, L. and Smith, A., 2012. Robust fuzzy extractors and authenticated key agreement from close secrets. IEEE Transactions on Information Theory, 58(9), pp.6207-6222.

[10] Y. Z. Ding. Error correction in the bounded storage model. In 2nd Theory of Cryptography Conference - TCC 2005, volume 3378 of LNCS, 578-599. Springer, 2005.

[11] Y. Dodis and A. Smith. Correcting errors without leaking partial information. In 37th Annual ACM Symposium on Theory of Computing (STOC), 654-663. ACM Press, 2005.

[12] Grünwald, Peter, and Paul Vitányi. Shannon information and Kolmogorov complexity. arXiv:cs/0410002 (2004).

[13] Ming Li and Paul Vitányi. An introduction to Kolmogorov complexity and its applications. Springer, 4 edition, 2019.

[14] Andrei Romashchenko, Marius Zimand. An Operational Characterization of Mutual Information in Algorithmic Information Theory. J. ACM 66(5): 38:1-38:42 (2019).

[15] Emirhan Gürpınar and Andrei Romashchenko. Communication Complexity of the Secret Key Agreement in Algorithmic Information Theory. arXiv:2004.13411. A preliminary version presented in MFCS 2020.

[16] Alexander Shen, Vladimir Uspensky, and Nikolay Vereshchagin. Kolmogorov complexity and algorithmic randomness, volume 220. American Mathematical Soc., 2017.

[17] Shai Evra, Konstantin Golubev, and Alexander Lubotzky. Mixing properties and the chromatic number of ramanujan complexes. International Mathematics Research Notices, 22: 11520-11548 (2015).

[18] Joel Friedman and Avi Widgerson. On the second eigenvalue of hypergraphs. Combinatorica. 15(1): 43-65 (1995).

[19] L. Antunes, S. Laplante, A. Pinto, and L. Salvador. Cryptographic security of individual instances. In Information Theoretic Security: Second International Conference, ICITS (2007).

[20] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. Bulletin of the American Mathematical Society 43(4): 439-561 (2006).

[21] Eyal Kushilevitz and Noam Nisan. Communication Complexity. Cambridge University Press, 2006.

[22] Imre Csiszár and Prakash Narayan. Secrecy capacities for multiple terminals. IEEE Trans. Inform. Theor. 50(12): 3047-3061 (2004).

[23] Alexei Chernov, Andrej Muchnik, Andrei Romashchenko, Alexander Shen, and Nikolai Vereshchagin. Upper semi-lattice of binary strings with the relation "x is simple conditional to y". Theoretical Computer Science, 271(1-2): 69-95 (2002).

[24] Rudolf Ahlswede and Imre Csiszár. Common randomness in information theory and cryptography. I. Secret sharing. IEEE Transactions on Information Theory, 39(4): 1121-1132 (1993).

[25] Ueli M. Maurer. Secret key agreement by public discussion from common information. IEEE Transactions on Information Theory, 39(3): 733-742 (1993).

[26] Ilan Newman. Private vs. common random bits in communication complexity. Inf. Process. Lett. 39(2), 67-71 (1991).

[27] László Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In 27th Annual Symposium on Foundations of Computer Science. 337-347 (1986).

[28] Alexander K. Zvonkin and Leonid A. Levin. The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. Russian Mathematical Surveys, 25(6): 83-124 (1970).

[29] Andrei Kolmogorov. Three approaches to the quantitative definition of information. Problems of information transmission, 1(1): 1-7 (1965).

[30] Andrej Muchnik, Conditional complexity and codes, Theoretical Computer Science, 271(1-2): 97-109 (2002)

[31] Bruno Bauwens, Optimal probabilistic polynomial time compression and the Slepian-Wolf theorem: tighter version and simple proofs. arXiv:1802.00750 (2018)

[32] Marius Zimand, Kolmogorov complexity version of Slepian-Wolf coding. In Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing. 22-32 (2017)

[33] Alexander Shen, a personal communication. September 11, 2023.

[34] Mitali Bafna, Badih Ghazi, Noah Golowich, Madhu Sudan. Communication-rounds tradeoffs for common randomness and secret key generation. In Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, 1861-1871 (2019)

[35] Noah Golowich and Madhu Sudan. Round complexity of common randomness generation: The amortized setting. In Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 1076-1095 (2020)

[36] Madhu Sudan, Himanshu Tyagi, and Shun Watanabe. Communication for generating correlation: A unifying survey. IEEE Transactions on Information Theorym 66(1): 5-37 (2019)

[37] A. J. Hoffman, On the line graph of a projective plane. Proceedings of the American Mathematical Society 16(2): 297-302 (1965)

[38] Damiel Hammer, Andrei Romashchenko, Alexander Shen, and Nikolai Vereshchagin. Inequalities for Shannon entropy and Kolmogorov complexity. Journal of Computer and System Sciences 60(2): 442-464 (2000)

[39] Andrei Romashchenko, Pairs of Words with Nonmaterializable Mutual Information. Problems of Information Transmission 36(1): 3-20 (2000)

[40] An. A., Muchnik and A. E. Romashchenko. Stability of properties of Kolmogorov complexity under relativization. Problems of information transmission 46(1): 38-61 (2010)

A Preliminaries: Kolmogorov complexity in some more detail

Let M be a Turing Machine with two input tapes and one output tape. We say that p is a program that prints a string x given y (a description of x conditional on y) if M prints x on the pair of inputs (p, y). *Kolmogorov complexity* of x *conditional on* y relative to M is defined as

$$\mathrm{C}_M(x \mid y) = \min\{|p| : M(p, y) = x\}.$$

The invariance theorem (see [29]) claims that there exists an *optimal* Turing machine U such that for every other Turing machine V there is a number c_V such that for all x and y

$$\mathrm{C}_U(x \mid y) \le \mathrm{C}_V(x \mid y) + c_V.$$

Thus, the algorithmic complexity of x relative to U is minimal up to an additive constant. In the rest of the paper we fix an optimal machine U, omit the subscript U and define *Kolmogorov complexity* of x conditional on y as

$$\mathrm{C}(x \mid y) := \mathrm{C}_U(x \mid y).$$

Kolmogorov complexity $\mathrm{C}(x)$ of a string x (without a condition) is defined as the Kolmogorov complexity of x conditional on the empty string. We fix an arbitrary computable bijection between binary strings and all finite tuples of binary strings and define Kolmogorov complexity of a tuple $\langle x_1, \dots, x_k\rangle$ as Kolmogorov complexity of the code of this tuple. For brevity we denote this complexity by $\mathrm{C}(x_1, \dots, x_k)$. Similarly, we can fix a bijection between binary strings and elements of finite fields, polynomials over finite fields, directions in vector spaces over finite fields, etc., and talk about Kolmogorov complexities of these objects (implying Kolmogorov complexity of their codes). We use the conventional notation

$$\mathrm{I}(x : y) := \mathrm{C}(x) + \mathrm{C}(y) - \mathrm{C}(x, y) \text{ and } \mathrm{I}(x : y \mid z) := \mathrm{C}(x \mid z) + \mathrm{C}(y \mid z) - \mathrm{C}(x, y \mid z)$$

(mutual information and conditional mutual information for a pair) and

$$\mathrm{I}(x : y : z) := \mathrm{I}(x : y) - \mathrm{I}(x : y \mid z)$$

(the triple mutual information). The Kolmogorov–Levin theorem, [28], claims that for all x, y

$$\mathrm{C}(x, y) \stackrel{\text{lg}}{=} \mathrm{C}(x \mid y) + \mathrm{C}(y).$$

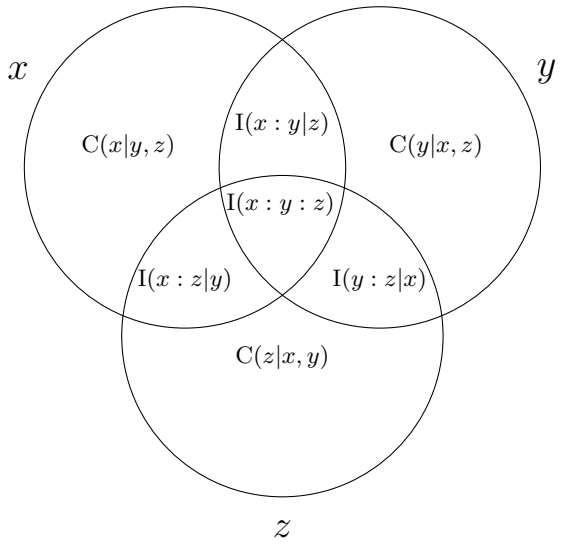


Figure 5: Complexity profile for a triple x, y, z. On this diagram it is easy to observe several standard equations:
• $\mathrm{C}(x) \stackrel{\lg}{=} \mathrm{C}(x \mid y, z) + \mathrm{I}(x : y \mid z) + \mathrm{I}(x : z \mid y) + \mathrm{I}(x : y : z)$
• $\mathrm{C}(x, y) \stackrel{\lg}{=} \mathrm{C}(x \mid y, z) + \mathrm{C}(y \mid x, z) + \mathrm{I}(x : y \mid z) + \mathrm{I}(x : z \mid y) + \mathrm{I}(y : z \mid x) + \mathrm{I}(x : y : z)$
• $\mathrm{C}(x \mid y) \stackrel{\lg}{=} \mathrm{C}(x \mid y, z) + \mathrm{I}(x : z \mid y)$
• $\mathrm{I}(x : y) \stackrel{\lg}{=} \mathrm{I}(x : y \mid z) + \mathrm{I}(x : y : z)$
• $\mathrm{I}(x : yz) \stackrel{\lg}{=} \mathrm{I}(x : y \mid z) + \mathrm{I}(x : z \mid y) + \mathrm{I}(x : y : z)$
and so on; all these equations are valid up to $O(\log(|x| + |y| + |z|))$.

Using the Kolmogorov–Levin theorem it is not hard to show that

$$\mathrm{I}(x : y : z) \stackrel{\lg}{=} \mathrm{C}(x) + \mathrm{C}(y) + \mathrm{C}(z) - \mathrm{C}(x, y) - \mathrm{C}(x, z) - \mathrm{C}(y, z) + \mathrm{C}(x, y, z)$$

and, therefore,

$$\mathrm{I}(x : y : z) \stackrel{\lg}{=} \mathrm{I}(x : z) - \mathrm{I}(x : z \mid y) \stackrel{\lg}{=} \mathrm{I}(y : z) - \mathrm{I}(y : z \mid x).$$

These relations can be observed on a Venn-like diagram, see Fig. 5.

A string x is said to be (almost) *incompressible* given y if $C(x \mid y) \geq^{\lg} |x|$, and x and y are said to be *independent*, if $\mathrm{I}(x : y) \stackrel{\lg}{=} 0$. For every n, the majority of binary strings of length n are almost incompressible; the vast majority of pairs of strings x and y of length n are independent.

For a pair of strings (x, y) we call by its *complexity profile* the triple of numbers $(\mathrm{C}(x), \mathrm{C}(y), \mathrm{C}(x, y))$. Due to the Kolmogorov–Levin theorem, the complexity profile of a pair is determined (up to additive error terms $O(\log(|x| + |y|))$) by the triple of numbers $(\mathrm{C}(x \mid y), \mathrm{C}(y \mid x), \mathrm{I}(x : y))$. Indeed,

$$\mathrm{C}(x) \stackrel{\lg}{=} \mathrm{C}(x \mid y) + \mathrm{I}(x : y),\ C(y) \stackrel{\lg}{=} \mathrm{C}(y \mid x) + \mathrm{I}(x : y),\ C(x, y) \stackrel{\lg}{=} \mathrm{C}(x \mid y) + \mathrm{C}(y \mid x) + \mathrm{I}(x : y).$$

Similarly, for a triple of strings (x, y, z) we define its *complexity profile* as the vector with 7 components

$$(\mathrm{C}(x), \mathrm{C}(y), \mathrm{C}(z), \mathrm{C}(x, y), \mathrm{C}(x, z), \mathrm{C}(y, z), \mathrm{C}(x, y, z)).$$

This profile can be equivalently specified (again, up to additive logarithmic error terms) by the numbers

$$\mathrm{C}(x \mid y, z), \mathrm{C}(y \mid x, z), \mathrm{C}(z \mid x, y), \mathrm{I}(x : y \mid z), \mathrm{I}(x : z \mid y), \mathrm{I}(y : z \mid x), \mathrm{I}(x : y : z),$$

see Fig. 5.

In general, for an n-tuple of string $(x_1, \ldots, x_n)$, its complexity profile is the vector that consists of $2^n - 1$ components $\mathrm{C}(x_{i_1}, \ldots, x_{i_s})$ for all non-empty tuples $1 \leq i_1 < \ldots < i_s \leq n$. In Fig. 1 we show diagrams illustrating the complexity profiles for Examples 2-3 and Proposition 1

For a survey of the basic properties of Kolmogorov complexity we refer the reader to the introductory chapters in [13] and [16].

B Bound of the spectral gap for the tri-expander

In the proof of Lemma 4 we use the fact that the bipartite graphs associated with the hypergraph from Proposition 1 are highly symmetric, and these symmetries simplify the computation of the eigenvalues. To guarantee this property, we have imposed restrictions that may seem artificial: the characteristic of the field is 2, we take into consideration only self-orthogonal vectors, and the direction $(1, \ldots, 1)$ is not included in the set V.

Proof of Lemma 4. By construction, the graphs G_1, G_2, G_3 are isomorphic. So we only need to compute the eigenvalues of G_1. Let us begin with the case when k is an odd number, and the vector $(1, 1, \ldots, 1)$ is not self-orthogonal.

In the bipartite graph $G_1 = (L, R, E)$ the left part of vertices L coincides with the space of all self-orthogonal directions V, the right part R consists of pairs of self-orthogonal directions (y, z) that are mutually orthogonal and $y \neq z$, and the set of hyperedges H consists of the triples (x, y, z) of self-orthogonal directions that are pairwise distinct and mutually orthogonal.

To compute $\#L$, we count the number of directions in $\mathcal{L}_{so}$. To this end, we divide the total number of non-zero vectors in the $(k+1)$-dimensional space $\mathcal{L}_{so}$ by the number of vectors in each equivalence class (the number of non-zero elements in the field), which gives $\#L = \Theta(2^{kn})$.

For each $y \neq (1, \ldots, 1)$, the space of vectors $z \in \mathcal{L}_{so}$ that are orthogonal to y is a subspace of co-dimension 1 in $\mathcal{L}_{so}$. To count the number of directions in this subspace, we need again divide the total number of non-zero vectors by the number of non-zero elements in the field. We obtain that $\#R = \Theta(2^{kn}) \cdot \Theta(2^{(k-1)n})$.

To find D_R, we count the number of directions in $\mathcal{L}_{so}$ that are orthogonal to two directions y, z (that cannot coincide with $(1, \ldots, 1)$), i.e., the directions in a subspace of co-dimension 2, which gives $D_R = \Theta(2^{(k-2)n})$. Similarly, we obtain $D_L = \Theta(2^{(k-1)n}) \cdot \Theta(2^{(k-2)n})$.

Observe that $\#H = \#L \cdot D_L = \#R \cdot D_R = \Theta(2^{kn}) \cdot \Theta(2^{(k-1)n}) \cdot \Theta(2^{(k-2)n})$.

Now we can compute the eigenvalues. We denote $M = \begin{pmatrix} 0 & A \\ A^\top & 0 \end{pmatrix}$ the adjacency matrix of the graph. We estimate the eigenvalues of $A \cdot A^\top$, which is the matrix of paths of length 2 in the graph, starting and finishing in L. Starting at some $x \in L$, we can go to some $(y, z) \in R$ and then either come back to the same x, or to end up in a different $x' \in L$. For a fixed x, the number of paths

$$x \to (y, z) \to x$$

is equal to D_L (any (y, z) matching x serves as the middle point of the path). For $x \neq x'$, the number of paths

$$x \to (y, z) \to x'$$

is equal to the number of mutually orthogonal pairs (y, z) that are both orthogonal to x and x', and all four directions x, x', y, z are distinct. This is similar to the computation of D_L but the co-dimensions are incremented: we have $\Theta(2^{(k-2)n}) \cdot \Theta(2^{(k-3)n})$ such pairs.

We denote by I the identity matrix and by J the matrix of all ones. Those matrices are both symmetric. We get

$$\begin{aligned} A \cdot A^\top &= D_L \cdot I + \Theta(2^{(k-2)n} \cdot 2^{(k-3)n}) \cdot (J - I) \\ &= \Theta(2^{(k-1)n} \cdot 2^{(k-2)n}) \cdot I + \Theta(2^{(k-2)n} \cdot 2^{(k-3)n}) \cdot J. \end{aligned}$$

Observe that I and J have a common basis of eigenvectors. Indeed, for the matrix I all vectors in the space are eigenvectors (and the only eigenvalue is 1 with multiplicity $\#L$). The eigenvalues of J are the number $\#L$ (of multiplicity 1) and 0 (of multiplicity $\#L - 1$). Therefore, the eigenvectors of $A \cdot A^\top$ are

$$\begin{aligned} \lambda_1 &= \Theta(2^{(k-2)n} \cdot 2^{(k-3)n}) \cdot \#L = \Theta(2^{kn} \cdot 2^{(k-2)n} \cdot 2^{(k-3)n}) = \Theta(2^{(3k-5)n}) \\ \lambda_2 &= \ldots = \lambda_{\#L} = \Theta(2^{(k-1)n} \cdot 2^{(k-2)n}) = \Theta(2^{(2k-3)n}) = \Theta(D_L). \end{aligned}$$

(Observe that λ_1 can be found directly as $D_L \cdot D_R$.) The eigenvalues of M are the square roots of those of $A \cdot A^\top$.

If k is even, the computation is similar, but on each step we should subtract from the set of self-orthogonal directions the vectors collinear with $(1, \ldots, 1)$. □

C Useful information inequalities

Proof of Lemma 5. We need to prove that

$$\mathrm{I}(x : y \mid s) \leq^{\lg} \mathrm{I}(x : y) + \mathrm{I}(s : xy). \tag{21}$$

Observe that

$$\mathrm{I}(x : y) \stackrel{\lg}{=} \mathrm{I}(x : y \mid s) + \mathrm{I}(x : y : s)$$

and

$$\mathrm{I}(s : xy) \stackrel{\lg}{=} \mathrm{I}(s : x) + \mathrm{I}(s : y) - \mathrm{I}(x : y : s).$$

Therefore, (21) rewrites to

$$I(x : y \mid s) \leq^{\lg} \mathrm{I}(x : y \mid s) + \mathrm{I}(x : y : s) + \mathrm{I}(s : x) + \mathrm{I}(s : y) - \mathrm{I}(x : y : s).$$

This inequality is always true since the terms $\mathrm{I}(s : x)$ and $\mathrm{I}(s : y)$ are non-negative. □

Lemma 7. *For all* x, y, x', y', r
(i) $\mathrm{I}(x' : y \mid z) \leq^{\lg} \mathrm{I}(x : y \mid z) + \mathrm{C}(x' \mid x)$.
(ii) $\mathrm{I}(x' : y' \mid z) \leq^{\lg} \mathrm{I}(x : y \mid z) + \mathrm{C}(x' \mid x) + \mathrm{C}(y' \mid y)$.
(iii) $\mathrm{I}(x' : y' \mid z, r) \leq^{\lg} \mathrm{I}(x : y \mid z, r) + \mathrm{C}(x' \mid x, r) + \mathrm{C}(y' \mid y, r)$.

Proof. To prove (i), we observe that by the chain rule for the mutual information

$$\mathrm{I}(x, x' : y \mid z) \stackrel{\lg}{=} \mathrm{I}(x : y \mid z) + \mathrm{I}(x' : y \mid x, z) \stackrel{\lg}{=} \mathrm{I}(x' : y \mid z) + \mathrm{I}(x : y \mid x', z).$$

Therefore,

$$\begin{array}{rcl} \mathrm{I}(x' : y \mid z) & \stackrel{\lg}{=} & \mathrm{I}(x : y \mid z) + \mathrm{I}(x' : y \mid x, z) - \mathrm{I}(x : y | x', z) \\ & \leq^{\lg} & \mathrm{I}(x : y \mid z) + \mathrm{I}(x' : y \mid x, z) \\ & \leq^{\lg} & \mathrm{I}(x : y \mid z) + \mathrm{C}(x' \mid x, z) \\ & \leq^{\lg} & \mathrm{I}(x : y \mid z) + \mathrm{C}(x' \mid x). \end{array}$$

To prove (ii) we apply the same argument twice (at first we replace x' by x and then replace y' by y),

$$\mathrm{I}(x' : y' \mid z) \leq^{\lg} \mathrm{I}(x : y' \mid z) + \mathrm{C}(x' \mid x) \leq^{\lg} \mathrm{I}(x : y \mid z) + \mathrm{C}(x' \mid x) + \mathrm{C}(y' \mid y).$$

The proof of (iii) is a "relativized" version of the proof of (ii); we only need to add r to the condition in all term of Kolmogorov complexity in the argument. □

D Lemma on relativization

Kolmogorov complexity can be relativized: for any oracle $\mathcal{O}$, we may define Kolmogorov complexity $\mathrm{C}^{\mathcal{O}}(x)$, $\mathrm{C}^{\mathcal{O}}(x \mid y)$ in terms of a universal decompressor that can access $\mathcal{O}$. In case when $\mathcal{O}$ is represented by a finite string s, the relativization has a simple meaning: $\mathrm{C}^{\mathcal{O}}(x) = \mathrm{C}(x \mid s) + O(1)$ and $\mathrm{C}^{\mathcal{O}}(x \mid y) = \mathrm{C}(x \mid y, s)$. Having fixed an oracle, we can pose the problem of secret key agreement in terms of the relativized Kolmogorov complexity.

The lower bounds on communication complexity of secret key agreement that we have discussed followed a pretty constructive scheme: we defined a set of input data sets H such that

- most triples $(x, y, z) \in H$ have complexity profile close to the required parameters, and
- we show that any secret key agreement that succeeds on most $(x, y, z) \in H$ must have large communication complexity.

This type of argument easily relativizes. Indeed, we can show that an oracle contains negligible information on most $(x, y, z) \in H$; therefore, the relativization does not change significantly the complexity profile for most triples in H. It follows that a successful secret key agreement scheme for triples from H in the sense of the relativized Kolmogorov complexity can be used as a secret key agreement scheme in the sense of the standard (non-relativized) Kolmogorov complexity. Thus, if we had a lower bound on the communication complexity of a successful protocol in the non-relativized setting, substantially the same bound applies to the relativized version

of the problem. In other words, *relativization cannot make the problem of secret key agreement easier* (cannot reduce communication complexity).

However, this argument does not imply that relativization cannot make the communication complexity of the problem *harder*. Indeed, the relativization conditional on an oracle $\mathcal{O}$ can provide us with completely new tuples (x, y, z) with the required complexity profile. These new input data sets may have unusual combinatorial properties, and *a priori* they might require a longer communication to agree on a secret key. In what follows we show that this is not the case: for a fixed complexity profile, relativization cannot increase communication complexity of secret key agreement.

In this section we use the technique of "clones" that was developed in [38, 39, 40]. We need to recall several definitions. For a tuple $(x_1, \ldots, x_n)$, its *complexity profile* is the vector of $2^n - 1$ values of Kolmogorov complexities $\mathrm{C}(x_{i_1}, \ldots, x_{i_k})$ for all $1 \le i_1 < \ldots i_k \le n$; the *extended complexity profile* includes (besides the same $2^n - 1$ values of unconditional complexities) the vector of conditional complexities $\mathrm{C}(x_{i_1}, \ldots, x_{i_k} \mid x_{j_1}, \ldots, x_{j_s})$ for all disjoint sets of indices $1 \le i_1 < \ldots i_k \le n$ and $1 \le j_1 < \ldots j_s \le n$.

For a tuple $(x_1, \ldots, x_n)$, the set of its *clones* denoted $\mathrm{Clone}(x_1, \ldots, x_n)$ is defined as the set of all $(x'_1, \ldots, x'_n)$ such that the extended complexity profile of $(x'_1, \ldots, x'_n)$ is component-wise not greater than the extended complexity profile of $(x_1, \ldots, x_n)$. It is known (see [39, 40]) that

- **substantiality:** $\log\left(\#\mathrm{Clone}(x_1, \ldots, x_n)\right) \stackrel{\lg}{=} \mathrm{C}(x_1, \ldots, x_n)$, and
- **uniformity:** if the set of all indices is split into two parts,

$$\{1, \ldots, n\} = \{i_1, \ldots, i_k\} \sqcup \{j_1, \ldots, j_s\},$$

then for every set of strings $x'_{j_1}, \ldots, x'_{j_s}$, there are at most $2^{\mathrm{C}(x_{i_1}, \ldots, x_{i_k} \mid x_{j_1}, \ldots, x_{j_s})+1}$ tuples $x'_{i_1}, \ldots, x'_{i_k}$ such that the n-tuple combined of $x'_{j_1}, \ldots, x'_{j_s}$ and $x'_{i_1}, \ldots, x'_{i_k}$ belongs to $\mathrm{Clone}(x_1, \ldots, x_n)$.

Similarly, we can define the extended complexity profile and the set of clones using Kolmogorov complexity relativized conditional on an oracle. For the relativized clones we have the same properties of substantiality and uniformity.

Lemma 8. *Let π be a communication protocol of secret key agreement for three participants (with inputs of length n), with public randomness ($m = O(n)$ public random bits). Assume that there exists an oracle $\mathcal{O}$ and a complexity profile $\bar{p} \in \mathbb{N}^7$ such that for some triple of inputs (x, y, z) with*

$$(\mathrm{C}^{\mathcal{O}}(x), \mathrm{C}^{\mathcal{O}}(y), \mathrm{C}^{\mathcal{O}}(z), \mathrm{C}^{\mathcal{O}}(x, y), \mathrm{C}^{\mathcal{O}}(x, z), \mathrm{C}^{\mathcal{O}}(y, z), \mathrm{C}^{\mathcal{O}}(x, y, z)) = \bar{p}$$

the protocol π fails (with a probability at least $1/2$) to obtain a common secret key. The secrecy failure means that given oracle $\mathcal{O}$, Kolmogorov complexity of the key conditional on the transcript $\mathrm{transcript}_\pi(x, y, z, r)$ and the string r sampled by the public source of random bits is below some threshold ℓ,

$$\mathrm{C}(\text{produced key} \mid r, \mathrm{transcript}_\pi(x, y, z, r)) < \ell.$$

Then there exists a triple of inputs (x', y', z') whose non-relativized complexity profile

$$\mathrm{C}(x'), \mathrm{C}(y'), \mathrm{C}(z'), \mathrm{C}(x', y'), \mathrm{C}(x', z'), \mathrm{C}(y', z'), \mathrm{C}(x', y', z')$$

is component-wise $O(\log n)$-close to $\bar{p}$, and the protocol π fails (also with a probability at least $1/2$) to obtain a common secret key (here the secrecy is understood without relativization, with a threshold $\ell' = \ell + O(\log n)$).

Proof. We begin the proof with some notation. Let (x, y, z) be a triple of inputs from the statement of the theorem. We denote by $\bar{q} = \bar{q}(x, y, z)$ the extended complexity profile of this triple. For each triple of inputs (x, y, z) and for each m-bit string r we denote by

$$\mathrm{transcript}_\pi(x, y, z, r)$$

the transcript of the protocol applied to these inputs with public random bits r and by

$$\mathrm{result}_\pi(x, y, z, r)$$

the key produced by the parties (if the parties fail to agree on a common key, we let w be the empty word).

If π fails (in the sense of the relativized Kolmogorov complexity) on some triple of inputs $(x, y, z) \in \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n$, it means that for a half of all bit strings $r \in \{0,1\}^m$, when the protocol is applied to these inputs,

(i) either Alice, Bob, and Charlie fail to agree on a common key,

(ii) or they do agree on one and the same key but this key is not secret, i.e., its relativized complexity is below the threshold, $\mathrm{C}^{\mathcal{O}}(\mathrm{result}_\pi(x,y,z,r) \mid \mathrm{transcript}_\pi(x,y,z,r),r) < \ell$.

We denote by $F \subset \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n$ the set of all triples of inputs (x,y,z) on which π fails, and $F' := F \cap \mathrm{Clone}(\bar{q})$, i.e., the set of all failure inputs that have extended complexity profile component-wise below $\bar{q}$.

We say that a tuple of strings

$$(x,y,z,r,\mathrm{transcript}_\pi(x,y,z,r),\mathrm{result}_\pi(x,y,z,r)) \tag{22}$$

is a *positive certificate* for (x,y,z) if all three parties agree on the same key and

$$\mathrm{C}^{\mathcal{O}}(\mathrm{result}_\pi(x,y,z,r) \mid r,\mathrm{transcript}_\pi(x,y,z,r)) \geq \ell.$$

Otherwise, this tuple is called a *negative certificate* for (x,y,z). By definition, π fails on (x,y,z) if for a half of all bit strings $r \in \{0,1\}^m$, the tuple (22) is a negative certificate.

We say that w is *compatible* with (r,t), if there exists at least one negative certificate (x,y,z,r,t,w). Observe that for each r and t there are less than 2^ℓ strings w such that

$$\mathrm{C}^{\mathcal{O}}(w \mid r,t) < \ell.$$

Therefore, for each (r,t) there are less than 2^ℓ strings w compatible with (r,t).

We know that F' is not empty (by the condition of the lemma, at least one triple (x,y,z) causes a failure of π). Observe that

- F' inherits the property of *uniformity* from $\mathrm{Clone}(x,y,z)$,
- we can enumerate the set F' given access to $\mathcal{O}$,
- since (x,y,z) can be specified by its ordinal number in this enumeration,

$$\mathrm{C}^{\mathcal{O}}(x,y,z) \leq^{\lg} \log \#F'$$

or, equivalently, $\#F' \geq 2^{\mathrm{C}^{\mathcal{O}}(x,y,z)-O(\log n)}$.

Now define a combinatorial structure that looks "similar" to the triple of sets

$$(F', \text{positive certificates for } F', \text{negative certificates for } F'), \tag{23}$$

but the oracle $\mathcal{O}$ is not involved in the definition. Having fixed an arbitrary set of tuples $G \subset \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n$ and the protocol π, we can compute the set $S(G)$ of all tuples (22) for every $(x,y,z) \in G$. We consider possible splits of S into two classes, $S = S_+ \sqcup S_-$ and say that (r,t) is G-compatible with w, if there is at least one tuple $(x,y,z,r,t,w) \in S_-$ (for some $(x,y,z) \in G$). Such a split is *valid* if

(a) for each $(x,y,z) \in G$ and for a half of $r \in \{0,1\}^m$, the tuple

$$(x,y,z,r,\mathrm{transcript}_\pi(x,y,z,r),\mathrm{result}_\pi(x,y,z,r))$$

belongs to S_-, and

(b) for each (r,t) there are less than 2^ℓ strings w that is G-compatible with (r,t).

We say that a valid split (G,S_+,S_-) is *similar* to the original triple (23) if

- **a variant of substantiality:** $\log \#G \geq^{\lg} \mathrm{C}^{\mathcal{O}}(x,y,z)$,
- **a variant of uniformity:** the cardinalities of sections and projections of G are not greater than the exponent of the corresponding complexity term in $\bar{q}$ (similarly to the property of uniformity of $F' \subset \mathrm{Clone}^{\mathcal{O}}(x,y,z)$ as formulated above).

Obviously, triples of sets satisfying the definition of similarity exist, e.g., the original set F' with its positive and negative certificates is similar to itself. Let (G^0, S^0_+, S^0_-) be the very first (e.g., in the lexicographical order) triple of sets respecting the presented requirements.

Although we cannot enumerate the elements of F' without access to the oracle $\mathcal{O}$, the sets (G^0, S^0_+, S^0_-) can be found algorithmicaly given only the protocol π and the components of the extended vector $\bar{q}$. Observe that for most triples $(x', y', z') \in G_0$, the extended complexity profile (non-relativized one) is $O(\log n)$-close to $\bar{q}$. Condition (b) above implies that for each w that is G^0-compatible with (r, t) we have $\mathrm{C}(w \mid r, t) \leq^{\lg} \ell$ (once again, this Kolmogorov complexity term is not relativized).

Thus, most $(x', y', z') \in G_0$ have the required (non-relativized) complexity profile, and by the construction of triples similar to (23), our protocol π fails on these triples with a probability $> 1/2$. This observation concludes the proof. □

Remark 9. If a communication protocol leaks minor information on the key to the adversary,

$$\mathrm{C}(\text{key} \mid \text{transcript}, \text{public randomness}) = |\text{key}| - \delta,$$

we can improve the secrecy by taking a random hash $\text{key}' = \text{hash}(\text{key}, \text{public random bits})$ such that the new key' is $\delta + O(1)$ shorter than original key but

$$\mathrm{C}(\text{key}' \mid \text{transcript}, \text{public randomness}) = |\text{key}'| - O(1),$$

see [14]. Thus, if we can agree on a mildly secure key of an asymptotically optimal size, we can also agree on a strongly secure key of approximately the same size.

Remark 10. Lemma 8 can be understood in the counter-positive way: if Alice, Bob, and Charlie can efficiently agree on a secret key for triples of inputs (x, y, z) with some specific complexity profile (in the sense of the standard non-relativized Kolmogorov complexity), they can do the same in the sense of Kolmogorov complexity relativized conditional on oracle $\mathcal{O}$. In other words, relativization changes the set of inputs with a specified complexity profile but it does not make the problem of secret key agreement more difficult.

E Proof of Lemma 6

Let us recall Muchnik's theorem on conditional descriptions and its version proven by Bauwens and Zimand.

Theorem 7. *(a) [30] For every string a and for all strings $b_1, \ldots, b_\ell$ and for every number $m \leq \mathrm{C}(a)$ there exists a "digital fingerprint" of a of length m that looks maximally random conditional on each b_j. Technically, this means that for some $\tilde{a}$ we have*

$$\mathrm{C}(\tilde{a}) \stackrel{\lg}{=} m,\ \mathrm{C}(\tilde{a} \mid a) \stackrel{\lg}{=} 0, \text{ and } \mathrm{C}(\tilde{a} \mid b_j) = \min\{\mathrm{C}(a \mid b_j), m\} \text{ for } j = 1, \ldots, \ell.$$

(b) [31, 32] (see also the Single Source Compression Theorem *in [14]) Moreover, such a "fingerprint" can be constructed pretty explicitly: given the length of string a, the numbers $\mathrm{C}(a \mid b_j)$, and m, one can construct an algorithm* Code *such that the conditions from (a) are valid for the vast majority of strings $\tilde{a} = \mathrm{Code}(a, r)$, where the probability is taken over the choice of a string r of length $O(\log(|a| + |b_1| + \ldots + |b_\ell|))$.*

Proof of Lemma 6. **Proof of (a).** We apply Theorem 7 for $\ell = 1$, with $a = x$ and $b_1 = \langle y, z \rangle$, and $m = \delta$ and obtain a string $\tilde{x}$ such that

$$\mathrm{C}(\tilde{x}) \stackrel{\lg}{=} \delta,\ \mathrm{C}(\tilde{x} \mid x) \stackrel{\lg}{=} 0,\ \mathrm{C}(\tilde{x} \mid y, z) \stackrel{\lg}{=} \delta.$$

In a similar way (applying again Theorem 7) we obtain $\tilde{y}$ and $\tilde{z}$ such that

$$\mathrm{C}(\tilde{y}) \stackrel{\lg}{=} \delta,\ \mathrm{C}(\tilde{y} \mid y) \stackrel{\lg}{=} 0,\ \mathrm{C}(\tilde{y} \mid x, z) \stackrel{\lg}{=} \delta \text{ and } \mathrm{C}(\tilde{z}) \stackrel{\lg}{=} \delta,\ \mathrm{C}(\tilde{z} \mid z) \stackrel{\lg}{=} 0,\ \mathrm{C}(\tilde{z} \mid x, y) \stackrel{\lg}{=} \delta.$$

A routine check shows that the triple (x, y, z) *conditional on* $\langle \tilde{x}, \tilde{y}, \tilde{z} \rangle$ has complexity profile (18).

Let π' be a communication protocol for the profile (18). Given (x, y, z), Alice, Bob, and Charlie can do as follows: each of them computes a fingerprint $\tilde{x}, \tilde{y}, \tilde{z}$ for x, y, and z respectively, and broadcast them. Then they proceed with a protocol π' applied to (x, y, z) and the complexity profile (18) (for Kolmogorov complexity relativized conditional on $\tilde{x}, \tilde{y}, \tilde{z}$). If the protocol succeeds, they obtain a secret key w that is incompressible given the public random bits and the full transcript of the combined protocol, which consists of the transcript of π', the

public random bits, *and the broadcasted strings* $\tilde{x}, \tilde{y}, \tilde{z}$. Thus, we obtain a communication protocol for the original (x, y, z), whose communication complexity is equal to the communication complexity of π' increased by 3δ bits (the total length of $\tilde{x}, \tilde{y}, \tilde{z}$ broadcasted at the first stage).

We know the optimal communication complexity of secret key agreement for the original (x, y, z) due to Theorem 4. Therefore, communication complexity of π' (for the relativized complexity profile) cannot be better than (8), which is in our case $3\alpha + \frac{3}{2}\beta - 3\delta$ (the optimal communication complexity for (x, y, z) decreased by 3δ). It remains to use Lemma 8 and conclude that the communication complexity of secret key agreement for (x', y', z') having the non-relativized complexity profile (18) cannot be better than $3\alpha + \frac{3}{2}\beta - 3\delta$.

Proof of (b). Let v be a string of δ bits such that $I(x, y, z : v) \stackrel{\lg}{=} 0$, and

$$x' = \langle x, v\rangle,\ y' = \langle y, v\rangle,\ z' = \langle z, v\rangle. \tag{24}$$

It is easy to verify that complexity profile of (x', y', z') matches (19). From Theorem 2 it follows that the optimal size of a secret key that three parties can agree on when given x', y', z' as inputs is

$$\begin{array}{l} \frac{1}{2}\left(I(x' : y' \mid z') + I(x' : z' \mid y') + I(y' : z' \mid z')\right) + I(x' : y' : z') \\ \stackrel{\lg}{=} \frac{3}{2}\beta + \gamma + \delta \stackrel{\lg}{=} \frac{1}{2}\left(I(x : y \mid z) + I(x : z \mid y) + I(y : z \mid z)\right) + I(x : y : z) + \delta. \end{array} \tag{25}$$

(i.e., the size of the key for the triple of inputs (x, y, z) plus δ). Our aim is to show that such a protocol requires communication complexity at least

$$\mathrm{C}(x', y', z') - \frac{1}{2}\big(I(x' : y' \mid z') + I(x' : z' \mid y') + I(y' : z' \mid z')\big) - I(x' : y' : z') = 3\alpha + \frac{3}{2}\beta.$$

Assume for the sake of contradiction that there is protocol π that achieves the goal with communication complexity $3\alpha + \frac{3}{2}\beta - \epsilon$. In what follows we construct a protocol π' that allows to construct an optimal size secret key with the same communication complexity for the original inputs (x, y, z).

In the new protocol Alice, Bob, and Charlie take from the (common) public source of random bits a string of δ bits v and define x', y', z' as in (24). With an overwhelming probability we have $I(x, y, z : v) \stackrel{\lg}{=} 0$, so we have (19). Then Alice, Bob, and Charlie proceed as in protocol π and find a common key w of size (25). As π is a valid protocol of secret key agreement, the key w has zero mutual information with the transcript t of the protocol. However, we loose the conditional of secrecy when v is public (which is in our case a part of the public source of random bits accessible to the attacker). However, we may restore the secrecy by reducing the size of the key. We apply Theorem 7 and construct $w' = \mathrm{Code}(w, r')$ (where r' is a string of public random bits of logarithmic size) such that

$$\mathrm{C}(w') = \frac{3}{2}\beta + \gamma,\ \mathrm{I}(w' : v, t) \stackrel{\lg}{=} 0$$

with a high probability (over the choice of r'). The produced w' can be taken as a secret key. The new protocol has communication complexity $3\alpha + \frac{3}{2}\beta - \epsilon$ (the same as π), and we get a contradiction with Theorem 4 unless $\epsilon \stackrel{\lg}{=} 0$.

Proof of (c). Let (x, y, z) be a hyperedge of tri-expander, as in the proof of Theorem 4. We will transform this triple in a different triple of inputs (x', y', z') using the following trick suggested by Alexander Shen, [33]. We apply Theorem 7 with $\ell = 3$ for $a = x$, $b_1 = y$, $b_2 = z$, $b_3 = \langle y, z\rangle$ and $m = \mathrm{C}(x) - \delta$, and obtain an x' such that

$$\mathrm{C}(x') = \mathrm{C}(x) - \delta,\ \mathrm{C}(x' \mid x) \stackrel{\lg}{=} 0,\ \mathrm{C}(x' \mid y) \stackrel{\lg}{=} \mathrm{C}(x \mid y),\ \mathrm{C}(x' \mid z) \stackrel{\lg}{=} \mathrm{C}(x \mid z),\ C(x' \mid y, z) \stackrel{\lg}{=} \mathrm{C}(x \mid y, z).$$

In a similar way, we obtain y' and z' such that

$$\mathrm{C}(y') = \mathrm{C}(y) - \delta,\ \mathrm{C}(y' \mid y) \stackrel{\lg}{=} 0,\ \mathrm{C}(y' \mid x) \stackrel{\lg}{=} \mathrm{C}(y \mid x),\ \mathrm{C}(y' \mid z) \stackrel{\lg}{=} \mathrm{C}(y \mid z),\ C(y' \mid x, z) \stackrel{\lg}{=} \mathrm{C}(y \mid x, z)$$

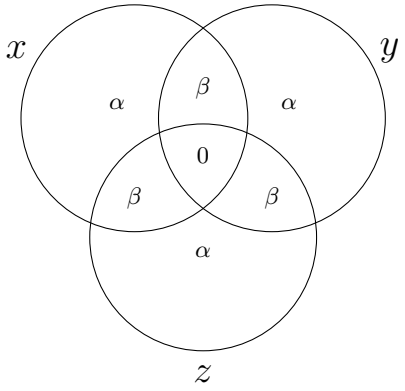
and

$$\mathrm{C}(z') = \mathrm{C}(z) - \delta,\ \mathrm{C}(z' \mid z) \stackrel{\lg}{=} 0,\ \mathrm{C}(z' \mid x) \stackrel{\lg}{=} \mathrm{C}(z \mid x),\ \mathrm{C}(z' \mid y) \stackrel{\lg}{=} \mathrm{C}(z \mid y),\ C(z' \mid x, y) \stackrel{\lg}{=} \mathrm{C}(z \mid x, y).$$

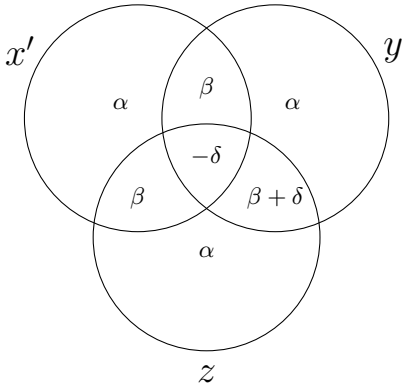
It is not hard to verify that the triple (x', y', z') has complexity profile (20), see Fig. 6.

From Theorem 2 it follows that the optimal size of a secret key that three parties can agree on when given x', y', z' as inputs is

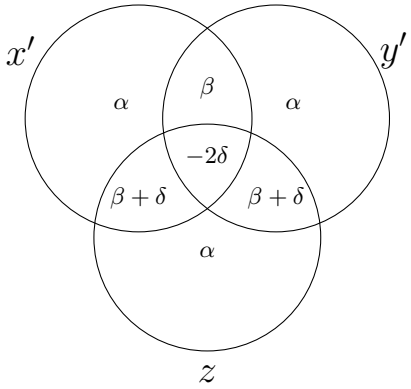
$$\frac{1}{2}\left(I(x' : y' \mid z') + I(x' : z' \mid y') + I(y' : z' \mid z')\right) + I(x' : y' : z') \stackrel{\lg}{=} \frac{3\beta}{2} - \frac{3\delta}{2}.$$



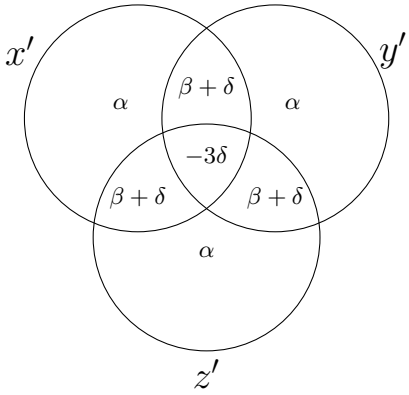
(a) $\mathrm{C}(x \mid y,z) \stackrel{\lg}{=} \mathrm{C}(y \mid x,z) \stackrel{\lg}{=} \mathrm{C}(z \mid x,y) \stackrel{\lg}{=} \alpha$,
$\mathrm{I}(x : y \mid z) \stackrel{\lg}{=} \mathrm{I}(x : z \mid y) \stackrel{\lg}{=} \mathrm{I}(y : z \mid x) \stackrel{\lg}{=} \beta$,
$\mathrm{I}(x : y : z) \stackrel{\lg}{=} 0$.



(b) $\mathrm{C}(x'|y,z) \stackrel{\lg}{=} \mathrm{C}(y|x',z) \stackrel{\lg}{=} \mathrm{C}(z|x',y) \stackrel{\lg}{=} \alpha$,
$\mathrm{I}(x' : y|z) \stackrel{\lg}{=} \mathrm{I}(x' : z|y) \stackrel{\lg}{=} \beta$, $\mathrm{I}(y : z|x') \stackrel{\lg}{=} \beta + \delta$,
$\mathrm{I}(x' : y : z) \stackrel{\lg}{=} -\delta$.



(c) $\mathrm{C}(x'|y',z) \stackrel{\lg}{=} \mathrm{C}(y'|x',z) \stackrel{\lg}{=} \mathrm{C}(z|x',y') \stackrel{\lg}{=} \alpha$,
$\mathrm{I}(x' : y'|z) \stackrel{\lg}{=} \beta$, $\mathrm{I}(x' : z|y') \stackrel{\lg}{=} \mathrm{I}(y' : z|x') \stackrel{\lg}{=} \beta + \delta$,
$\mathrm{I}(x' : y' : z) \stackrel{\lg}{=} -2\delta$.



(d) $\mathrm{C}(x'|y',z') \stackrel{\lg}{=} \mathrm{C}(y'|x',z') \stackrel{\lg}{=} \mathrm{C}(z'|x',y') \stackrel{\lg}{=} \alpha$,
$\mathrm{I}(x' : y'|z') \stackrel{\lg}{=} \mathrm{I}(x' : z'|y') \stackrel{\lg}{=} \mathrm{I}(y' : z'|x') \stackrel{\lg}{=} \beta + \delta$,
$\mathrm{I}(x' : y' : z') \stackrel{\lg}{=} -3\delta$.

Figure 6: Complexity profile for Muchnik's fingerprints of x, y, z.

This means (see Corollary 3) that Charlie must send a message m_C such that

$$\mathrm{I}(x' : y' \mid m_C) \geq^{\lg} \frac{3\beta}{2} - \frac{3\delta}{2}.$$

Observe that $\mathrm{I}(x' : y') \stackrel{\lg}{=} \beta - 2\delta$. Therefore, the mutual information between m_C and (x', y') must be greater than the difference between $\mathrm{I}(x' : y' \mid m_C)$ and $\mathrm{I}(x' : y')$,

$$\mathrm{I}(m_C : x', y') \geq^{\lg} \left(\frac{3\beta}{2} - \frac{3\delta}{2}\right) - (\beta - 2\delta) \stackrel{\lg}{=} \frac{\beta}{2} + \frac{\delta}{2}$$

(as in the proof of Theorem 4). If (x', y', z') are obtained from a hyperedge of a tri-expander, then we can apply Corollary 2 and conclude that

$$\mathrm{C}(m_C) \geq^{\lg} \mathrm{C}(z' \mid x'y') + \frac{\beta}{2} + \frac{\delta}{2} = \alpha + \frac{\beta}{2} + \frac{\delta}{2}.$$

A similar argument gives

$$\mathrm{C}(m_A) \geq \alpha + \frac{\beta}{2} + \frac{\delta}{2} \text{ and } \mathrm{C}(m_B) \geq \alpha + \frac{\beta}{2} + \frac{\delta}{2}$$

for the messages sent by Alice and Bob respectively. By summing up these bounds we conclude that the total length of all messages must be at least

$$3\alpha + \frac{3\beta}{2} + \frac{3\delta}{2},$$

which is exactly the communication complexity of the omniscience protocol. □