Secret Keys from Channel Noise

Hadi Ahmadi, Reihaneh Safavi-Naini

Department of Computer Science, University of Calgary, Canada
{hahmadi, rei}@ucalgary.ca

Abstract. We study the problem of unconditionally secure Secret Key Establishment (SKE) when Alice and Bob are connected by two noisy channels in opposite directions, and the channels are eavesdropped by Eve. We consider the case that Alice and Bob do not have any sources of initial randomness at their disposal. We start by discussing special cases of interest where SKE is impossible, and then provide a simple SKE construction over a binary symmetric channel that achieves some rates of secret key. We next focus on the Secret Key (SK) capacity, i.e., the highest rate of secure and reliable key establishment (in bits per channel use) that the parties can achieve. Relying on the existence of capacity-achieving coding schemes, we propose a multi-round SKE protocol, called the *main protocol*, that proves a lower bound on the SK capacity. The main protocol consists of an initialization round, followed by repeated use of a two-round SKE protocol, called the *basic protocol*. We also provide an upper bound on the SK capacity and show that the two bounds coincide when channels do not leak information to the adversary. We apply the results to the case that communicants are connected by binary symmetric channels.

1 Introduction

In cryptography, it is commonly assumed that parties have access to sources of randomness that serve their randomized algorithms and protocols. It is also common to assume that this randomness is *perfect*, i.e., randomness is represented as a sequence of independently and uniformly random bits. For example, in the Diffie-Hellman (DH) key agreement protocol the two parties require uniformly random strings to generate the exponents. Noting that perfect randomness is hard to obtain and, in many scenarios, the distribution of the random source is either biased or unknown, Dodis and Spencer [17] initiated the study of building cryptographic primitives using *imperfect* random sources. They focussed on symmetric-key encryption and message authentication, and showed that in both cases the corresponding sources do not require perfect randomness.

In practice, generating randomness with high entropy needs specialized hardware and/or software as well as access to complex processes that could be hard to obtain in many cases, including when devices with low computational resources are considered. A natural question is then, whether the need for a separate random source can be eliminated from a particular cryptographic task. Obviously, cryptography is not possible without randomness (uncertainty). For devices with communication capability however, channel noise is an attractive *resource* for providing randomness.

In a traditional communication system, information (randomness) sources and communication channels are considered as two different types of resources. Physical communication channels are noisy and can be viewed as a potential resource to provide randomness in cryptographic systems. Wyner's pioneering work [34] showed that channel noise can be used to provide perfect security in message transmission, and in fact replace the role of the shared secret key in Shannon's model [30] of perfect security. This work started a long line of research that relies on channel noise for constructing cryptographic primitives, and it shares the vision of Crépeau and Kilian [12] that, *"Noise, on the other hand, breeds disorder, uncertainty, and confusion. Thus, it is the cryptographer's natural ally."*

Wyner's work and, to our knowledge, all cryptographic systems that use noisy channels as a resource, however, also assume that the parties in the system have access to independent sources of initial randomness. In this paper, we initiate the study of cryptographic systems without making this assumption. We consider the case that the algorithms have fixed hardwired constant strings, such as identification strings that are publicly known, and there is no other resource for randomness except channel noise. One can ask whether, in such a setting, a particular cryptographic primitive exists and, if it does, whether it is sufficiently efficient to be of practical interest. The answer to this question would depend on the required functionality of the primitive, and the system description (including the communication environment and the adversary framework). In this paper, we focus on the basic task of Secret Key Establishment (SKE) in the presence of a passive adversary and pose the following question:

Question 1. *Can Alice and Bob establish a shared secret key, without having access to initial randomness, by communicating over noisy channels that leak information to an eavesdropping adversary, Eve? In the case of a positive answer, are there efficient constructions to generate secret keys in practice?*

Here, we mean that Alice and Bob have neither independent nor correlated randomness initially. To the best of our knowledge, this paper is the first work to consider SKE with no initial randomness.

1.1 Our work

We focus on Question 1 and study SKE over a pair of independent Discrete Memoryless Broadcast Channels (DMBCs). A DMBC is a channel that provides noisy outputs to multiple receivers. We assume there is one DMBC from Alice to Bob and Eve, and one from Bob to Alice and Eve. We refer to this setup as 2DMBC and assume that this is the only method of communication in the system. SKE in this setup has been studied in [3]; however, again, it was assumed that Alice and Bob have access to sources of randomness.

We assume Alice and Bob each have a fixed string, $\mathbf{a}$ and $\mathbf{b}$, respectively. We also assume a full-duplex model of communication where, in *each channel use*, Aice and Bob each sends one symbol over her/his DMBC and, in *each communication round*, each sends a message of the same length over their respective DMBC. This communication model is used to simplify the presentation of our results; the results can be adapted to half-duplex channels where, in each communication round, either Alice or Bob sends a message.

Impossibility results: Beyond doubt, SKE without initial randomness is impossible if the channels between the parties are noise free. This observation also holds in the computational setting, e.g., for the DH protocol. This is because all parameters in the system are deterministic and, assuming Eve has at least the computational capability of Alice and Bob, she can execute the same algorithms as theirs to derive the key. This implies that using error correcting codes to construct reliable communication channels removes the possibility of SKE without initial randomness.

In Section 3, we discuss special cases of 2DMBC where SKE is impossible despite the existence of noise in the system. These special cases include (1) only one-way communication is possible, (2) one DMBC is completely noise free, and (3) one DMBC is noisy but returns two identical outputs. We note that SKE in the above cases have been already studied [14, 15, 23] under the assumption that initial randomness is available to the parties. The goals of these studies, however, was to show the possibility of SKE in the corresponding settings.

SKE Construction: We give a positive answer to Question 1 by considering an example scenario where each DMBC consists of two independent (errors over the two channels are independent) Binary Symmetric Channels (BSCs), one from the sender to the receiver and one from the sender to Eve, with bit error probabilities p_1 and p_2, respectively. We propose a two-round SKE protocol that uses three simple primitives, a von Neumann randomness extractor, a binary error-correcting code, and a universal hash function. The protocol works as follows. In round 1, Alice sends a constant sequence, e.g. an all-zero sequence, to Bob; Bob receives a noisy string and uses the von Neumann extractor to derive a uniformly random binary sequence from it. In round 2, Bob splits the uniform sequence into two sub-sequences, encodes them separately, and sends the codewords to Alice. Alice decodes her received sequence to find the two sub-sequences. Finally, Alice and Bob apply universal hashing to the sub-sequences to derive a secret key that is secure against Eve. Overall, the protocol computation time includes one run of the extractor, two runs of encoding and decoding functions, and one run of the universal hashing function. Using full-duplex communication channel allows Alice and Bob to independently initiate one instance of the protocol, and so effectively double the secret key rate per (duplex) channel use; this will of course double the computation cost.

Bounds on the SK capacity: We formalize the 2DMBC model and focus on the general description of a SKE protocol over a 2DMBC. We define the *Secret Key (SK) rate* of a protocol Π as the average number of shared random bits per channel use that Alice and Bob can securely and reliably generate by using Π. The *Secret Key (SK) capacity* of a 2DMBC is the highest SK rate that all possible SKE protocols can achieve. This leads to the following question:
Question 2. *What is the SK capacity of a 2DMBC?*

Towards answering Question 2, we provide lower and upper bounds on the SK capacity of a 2DMBC. We prove the lower bound by showing that there exists a SKE construction that achieves the bound. We describe a multi-round

SKE protocol, referred to as the *main protocol*, that consists of an *initialization round*, followed by repeated use of a two-round protocol that we call the *basic protocol*.

The initialization round bootstraps the main protocol by providing Alice and Bob with some pieces of "independent randomness" that is obtained from channel noise. By independent randomness, we mean a random variable that is independent of all random variables accessible to other parties. The randomness is derived from channel noise after one round of communication and is required for executing one iteration of the basic protocol. Each iteration of the basic protocol only uses the fresh randomness derived in the previous iteration. An execution of the basic protocol simultaneously serves two purposes: it (1) generates new pieces of independent randomness for Alice and Bob to be used in the next iteration, and (2) establishes one part of the shared secret key. The basic protocol uses two new primitives that we refer to as *secure block code* and *secure equipartition.* A secure block code is a deterministic primitive, consisting of a block code and a key derivation function that provides Alice and Bob with a part of the secret key. A secure equipartition is a tool to derive new independent randomness from channel noise which is hidden in the noisy received sequence. This randomness is independent of the channel input and Eve's view. The lower bound proof relies on the existence of these two primitives.

In each iteration of the basic protocol, the number of derived key bits and the number of channel uses are fixed; therefore, one can associate a fixed key rate for each iteration of the protocol. During the initialization round however, no secret key bit is derived. Since the SK rate of the main protocol is the average number of the final secret key bits per channel use, the channel uses in the initialization round can be amortized over the number of the consecutive invocations of the basic protocol and hence the SK rate tends towards that of a single basic protocol execution. One may propose other protocols for key establishment in the setting considered in this paper; an example of such a protocol is given in Section 1.2. Nonetheless, the main protocol described in this paper achieves the highest rate among the known constructions, hence resulting in a tighter lower bound on the SK capacity.

The lower bound shows that positive SK rates are achievable when both DMBCs are in favor of the legitimate parties, i.e., compared to Eve, the legitimate parties receive a less noisy version of the transmitted messages. More interestingly, it shows that this condition, although sufficient, is not necessary and *there are cases where both DMBCs are in favor of Eve, yet it is possible to establish secure shared key.*

We also provide an upper bound on the SK capacity by bounding the highest SK rate of a general multi-round SKE protocol. We show that the lower and the upper bounds coincide in the case that the channels do not leak any information to the adversary. This corresponds to the problem of common randomness generation over independent noisy channels, studied in [31], where the common randomness capacity was derived. In other words, the results in this paper match those in [31] under this special condition.

Discussion: The communication scenario considered in this paper naturally occurs in real life. All physical channels are noisy and in most cases, in particular in wireless communication, they are easy to eavesdrop. Assuming no initial perfect randomness for Alice and Bob is also natural when communicating nodes do not have additional hardware or access to complex random processes (e.g. processing time in a large computer system). In particular, mobile devices and their communication capabilities, match the setting considered in this paper. Our results show that, in the absence of initial randomness, nodes can start with constant strings such as their pre-stored IDs and "distill" randomness from channel noise.

Our work initiates a new direction for research: possibility and construction of cryptographic primitives when the only resource for randomness is channel noise. We note that converting a cryptographic primitive that uses noisy channel as a resource and allows Alice and Bob to have initial randomness, to the case that they do not have such randomness is not straightforward. As mentioned above, in some cases, the construction in the latter setting becomes impossible and, in cases such as this work, although SKE is possible, efficient constructions that achieve the lower bound or sufficiently high secret key rate, can become challenging.

The lower bound proof given in this paper, uses an existential argument: we do not give a construction that achieves the bound and can be used in practice. However, attempts to design efficient while optimal primitives for secure equipartition and secure block code can be directly applied to the main SKE protocol design to achieve SK rates close to the lower bound. This is an interesting direction for future research similar to the work in [8] that applies theoretical SKE results in [23, 34] to practice.

The SKE construction given for binary symmetric channels can be viewed as a relaxed version of the main protocol where a simplified one-round basic protocol is used only once. The von Neumann extractor plays the role of (secure) equipartition in deriving independent randomness while the combination of coding and universal hashing is to replace the secure block code. Using these computationally efficient yet non-optimal primitives results in SK rates that are well below the lower bound. We further discuss this in Section 6.

1.2 Related work

The problem considered in this paper has relations to a number of previous studied areas, in particular, secure message transmission and key agreement over noisy channels, key agreement over public discussion channels using correlated randomness, key extraction from weak keys, and common randomness generation over noisy channels. In the following, we briefly clarify these relations.

Exploiting channel noise to provide security functionalities is pioneered by Wyner [34] who proposed an alternative to Shannon's model of secure communication [30]. In Wyner's model, Alice and Bob do not have any initial shared key; they are however, connected by a noisy channel that is wiretapped and allows Eve to only receive a degraded version of what Bob receives. Wyner showed that it is possible to exploit channel noise to transmit messages with *perfect secrecy*. Wyner's definition of perfect secrecy is in line with Shannon's definition in the information-theoretic setting, i.e., requiring Eve's complete uncertainly about the transmitted message, given what she receives through her wiretap channel. Wyner's work initiated a long line of research on utilizing channel noise to construct information theoretically secure cryptographic primitives including SKE [1, 14, 22, 23, 29], Oblivious Transfer (OT) [12, 13, 27], and Bit Commitment (BC) schemes [5, 7]. In all these works however, access to initial randomness is assumed and removing this assumption will require revisiting the results and examining the existence of the primitives. For instance, secure message transmission in the original Wyner's model will not be possible without Alice having access to a random source.

Maurer [23], concurrently with Ahlswede and Csiszár [1], studied the problem of key agreement over a public discussion channel when Alice and Bob have initial correlated randomness. The correlated randomness may be obtained from correlated sources or communication over noisy channels. They determined lower and upper bounds on the SK capacity in this setting and showed conditions under which key agreement may or may not be possible. Key agreement using correlated randomness and a one-way noisy channel has been discussed in [22, 29] and it is shown that Alice and Bob can benefit from both resources (correlated sources and a noisy channel) to establish shared secret keys.

A related line of research considered stronger adversaries, i.e., active adversaries who can tamper with communication over public channels. Maurer and Wolf [25] revisited the results in [1, 23] in the active adversary setting and proved a number of possibility and impossibility results. Followup work considered key agreement (also referred to as key extraction) over public channels when Alice and Bob initially share a weak key [26, 28] or close randomness [18, 21].

The following two works are closely related to the setting in this paper, whereas neither provides results that are applicable to this setting. Venkatesan and Anantharam [31] considered shared randomness generation over a pair of independent DMCs and acquired the common randomness capacity of the channels. This is the first attempt to design communication primitives with no initial randomness. Authors noted that their results could not be applied to the case that the DMCs are eavesdropped by Eve – the setting that is considered in this paper.

SKE over a pair of independent DMBCs was considered in [3], where bounds on the SK capacity were provided. The constructions, however, assumed availability of free independent randomness to the parties, without which the corresponding proofs will not be valid. Assuming no initial randomness, one may use the results in [3] to design an SKE protocol as follows. Alice and Bob first execute an initialization round (e.g., using secure equipartition proposed in this paper) to derive the required amount of independent randomness, and then use the protocol given in [3] to establish a secret key. Compared to this protocol nevertheless, our main protocol potentially increases the SK rate up to two times, through iteration. The particular novelty of the basic protocol compared to the protocol in [3] is that, it combines the dual tasks of secure key derivation and fresh randomness generation (using secure equipartition).

1.3 Notation

We use calligraphic letters ($\mathcal{X}$), uppercase letters (X), and lowercase letters (x) to denote finite alphabets, Random variables (RVs), and their realizations over sets, respectively. $\mathcal{X}^n$ is the set of all sequences of length n (so called

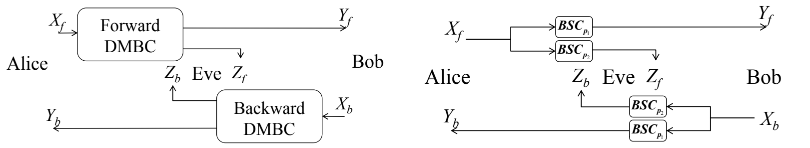
n-sequences) with elements from $\mathcal{X}$. $X^n = (X_1, X_2, \ldots, X_n) \in \mathcal{X}^n$ denotes a random n-sequence in $\mathcal{X}^n$. In case there is no confusion about the length, we use $\mathbf{X}$ to denote a random sequence and $\mathbf{x}$ to denote a realization in $\mathcal{X}^n$. While describing a multiple round protocol, we use $X^{n:r}$ (or $\mathbf{X}^{:r}$) to indicate a random n-sequence that is sent, received, or obtained in round r. For the RVs X, Y, and Z, we use $X \leftrightarrow Y \leftrightarrow Z$ to denote a Markov chain between them in the given order. '$||$' denotes the concatenation of two sequences. For a value x, we use $(x)_+$ to show $\max\{0, x\}$ and, for an integer N, we use $[N]$ to show the set of integers $\{1, 2, \ldots, N\}$. For two integers N and M, $N.M$ denotes their integer multiplication. All logarithms are in base 2 and, for $0 \le p \le 1$, $h(p) = -p \log p - (1-p)\log(1-p)$ denotes the binary entropy function.

1.4 Paper organization

Section 2 describes SKE over 2DMBCs and delivers the security definitions. In Section 3, we provide some impossibility results for special cases of 2DMBC setting and an example of a simple SKE construction for BSCs together with an estimation of the secret key rate. Section 4 summarizes our main results on the SK capacity. In Section 5, we describe the main protocol that achieves the lower bound. Section 6 studies the SKE results for the case of BSCs, and Section 7 concludes the paper.

2 The Secret Key Establishment Problem

A Discrete Memoryless Channel (DMC) $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ is a communication channel that, for any input symbol $X \in \mathcal{X}$, returns an output $Y \in \mathcal{Y}$ according to the distribution $P_{Y|X}$ and independently of other symbols. A Discrete Memoryless Broadcast Channel (DMBC) $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{YZ|X})$ is a channel that, for an input symbol $X \in \mathcal{X}$, returns two output symbols $Y \in \mathcal{Y}$ and $Z \in \mathcal{Z}$ according to the distribution $P_{YZ|X}$ and independently of other symbols. In the 2DMBC setup, shown in Fig. 1(a), there is a forward DMBC from Alice to Bob and Eve, denoted by $(\mathcal{X}_f, \mathcal{Y}_f, \mathcal{Z}_f, P_{Y_f,Z_f|X_f})$, and a backward DMBC from Bob to Alice and Eve, denoted by $(\mathcal{X}_b, \mathcal{Y}_b, \mathcal{Z}_b, P_{Y_bZ_b|X_b})$. The parties have deterministic computation systems. We describe SKE in the full-duplex model of communication where in each round Alice and Bob both can send messages.



(a) General 2DMBC (b) 2DMBC with independent BSCs

Fig. 1. The 2DMBC setup (a) in general and (b) in the case of independent BSCs

To establish a secret key, Alice and Bob follow a SKE protocol with t communication rounds where, in round r, each channel is used n_r times. The protocol is specified by a sequence of deterministic round function pairs, $(f_r, g_r)_{r=1}^{t-1}$, and a pair of deterministic key derivation functions (ϕ_A, ϕ_B) such that

$$f_r : \mathcal{Y}_f^{\sigma_{r-1}} \to \mathcal{X}_f^{n_r}, \qquad \phi_A : \mathcal{Y}_f^n \to \mathcal{S} \cup \{\perp\}, \tag{1}$$

$$g_r : \mathcal{Y}_b^{\sigma_{r-1}} \to \mathcal{X}_b^{n_r}, \qquad \phi_B : \mathcal{Y}_b^n \to \mathcal{S} \cup \{\perp\}, \tag{2}$$

where $\sigma_j = \sum_{i=0}^{j} n_i$, $\perp$ denotes the error symbol, and $n = \sigma_{t-1}$ is the total number of channel uses at the end of the protocol. The protocol takes as input a pair, $(\mathbf{a}, \mathbf{b}) \in \mathcal{X}_f^{n_0} \times \mathcal{X}_b^{n_0}$, of constant (publicly known) sequences. In a communication round r, Alice and Bob send the n_r-sequences $\mathbf{X}_f^{:r}$ and $\mathbf{X}_b^{:r}$ and receive $\mathbf{Y}_b^{:r}$ and $\mathbf{Y}_f^{:r}$, respectively. Eve receives $(\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})$. The input sequences are calculated as

$$\mathbf{X}_f^{:r} = \begin{cases} \mathbf{a}, & r = 0 \\ f_r(V_A^{:r-1}) & 1 \le r \le t-1 \end{cases}, \qquad \mathbf{X}_b^{:r} = \begin{cases} \mathbf{b}, & r = 0 \\ g_r(V_B^{:r-1}) & 1 \le r \le t-1 \end{cases}. \tag{3}$$

$V_A^{:r-1}$, $V_B^{:r-1}$, and $V_E^{:r-1}$ are, respectively, the views of Alice, Bob and Eve, at the end of round $r-1$, i.e.,

$$V_A^{:r-1} = (\mathbf{Y}_b^{:i})_{i=1}^{r-1}, \quad V_B^{:r-1} = (\mathbf{Y}_f^{:i})_{i=1}^{r-1}, \quad \text{and} \quad V_E^{:r-1} = (\mathbf{Z}_f^{:i}, \mathbf{Z}_b^{:i})_{i=1}^{r-1}. \tag{4}$$

By view of a party, we mean the randomness that they collect through the protocol execution. We do not include constants and deterministic functions that are applied to the variables in the views, since they do not result in new information (randomness). When the t rounds of communication are completed, Alice and Bob calculate their secret keys respectively as

$$S_A = \phi_A(V_A^{:t-1}), \quad \text{and} \quad S_B = \phi_B(V_B^{:t-1}). \tag{5}$$

Let $View_E = V_E^{:t-1}$ be Eve's view at the end of the protocol.

Definition 1. *For $R_{sk} \geq 0$ and $0 \leq \delta \leq 1$, the SKE protocol Π is* (R_{sk}, δ)-secure *if there exists a random variable $S \in \mathcal{S}$ such that the following requirements are satisfied:*

$$\text{Randomness:} \qquad \frac{H(S)}{n} \geq R_{sk} - \delta, \tag{6a}$$

$$\text{Reliability:} \qquad \Pr(S_A = S_B = S) \geq 1 - \delta, \tag{6b}$$

$$\text{Secrecy:} \qquad \frac{H(S|View_E)}{H(S)} \geq 1 - \delta. \tag{6c}$$

Definition 2. *The* Secret-Key (SK) capacity *C_{sk} is defined as the largest $R_{sk} \geq 0$ such that, for any arbitrarily small $\delta > 0$, there exists an (R_{sk}, δ)-secure SKE protocol.*

Remark 1. The above definition of SK capacity follows [34] and later [1, 14, 22, 23, 29]. It is referred to as *the weak SK capacity* since it only requires Eve's uncertainty about the secret key to be negligible in "rate". In contrast, in the "strong" SK capacity [24], Eve's total uncertainty must be negligible. Maurer and Wolf [24] showed that for the settings in [14, 23, 34], the weak definition can be replaced by the strong without sacrificing the SK capacity. We believe that a similar result can be proved for the setting in this paper, using an argument similar to [24]. We will show this in our future work.

3 SKE in special cases of 2DMBC

3.1 Impossibility results for special cases

We revisit a number of well-studied SKE scenarios that can be viewed as special cases of 2DMBC. We argue that, without initial randomness available to parties, SKE is impossible in these cases irrespective of the channel specification.

One-way communication: Consider a case that one of the DMBCs, say the backward DMBC, always returns constant values at its outputs. This is the same as assuming a one-way communication over the forward channel. Irrespective of the protocol, Alice will never have a single bit of randomness in her view and, without randomness, she cannot have a secret key. Note that this special case is essentially the one-way DMBC setting of Csiszár and Körner [14], with the difference that no initial randomness is provided to the parties.

One channel is noiseless and public: Without loss of generality, assume that the backward DMBC is noiseless and public. For any SKE protocol as described in Section 2, we have $\mathbf{X}_b^{:r} = \mathbf{Y}_b^{:r} = \mathbf{Z}_b^{:r}$ for each round r. This suggests that, at the end of the protocol, Eve's view includes Alice's view (see (4)). Eve can simply use Alice's key derivation function ϕ_A on her view to calculate S_A and so there will not exist any variable $S \in \mathcal{S}$ as the secret key that satisfies the requirements in (6). One can find a more precise argument by studying the upper bound, provided in Section 4, for this special case. It is interesting to note that this argument is also valid when in addition to the one-way DMBC a free "two-way" public discussion channel exists. This is the setting that was studied by Maurer in [23] and was proved to allow positive SK rates when parties have access to initial randomness.

One channel is noisy but returns two identical outputs: Assume that this property holds for the backward DMBC. In this case, $\mathbf{X}_b^{:r}$ may be different from the outputs and we only have $\mathbf{Y}_b^{:r} = \mathbf{Z}_b^{:r}$. Nevertheless, this is sufficient to argue that Eve's view includes Alice's view; hence, the impossibility of SKE.

3.2 An SKE protocol for binary symmetric channels

Assume that the 2DMBC consists of four independent binary symmetric channels as illustrated in Fig. 1(b). The main channels from Alice to Bob and vice versa have bit error probability p_1, while both Eve's channels have bit error probability p_2. Furthermore, Alice has an all-zero sequence of length m, $\mathbf{a} = \underline{0}^m$. We describe a two-round SKE construction that uses the primitives described below.

The von Neumann randomness extractor [32]: This extractor takes a binary sequence of even length and outputs a variable-length sequence that has uniform distribution. For an input Bernoulli sequence $\mathbf{Y} = (Y_1Y_2, Y_3Y_4, \ldots, Y_{m-1}Y_m)$ of even length m, where $P(Y_i = 1) = p$, the von Neumann extractor divides the sequence into $m/2$ pairs of bits and uses the following mapping on each pair

$$00 \rightarrow \Lambda, \qquad 01 \rightarrow 0, \qquad 10 \rightarrow 1, \qquad 11 \rightarrow \Lambda,$$

where Λ represents no output. The output sequence is the concatenation of the mapped bits. It is easy to observe that the extractor is computationally efficient and the output bits are independently and uniformly distributed.

While the von Neumann extractor does not return a fixed-length output, it can be used to design a primitive $Ext : \{0,1\}^m \rightarrow \{0,1\}^l \cup \{\perp\}$ that derives an l-bit uniform string from an m-bit Bernoulli sequence. The Ext function runs the von Neumann extractor on the m-bit sequence $\mathbf{Y}$. If the output length is less l, it returns $\perp$; otherwise, it returns the first l bits of the output. The probability that, for an m-bit Bernoulli sequence (with $P(Y_i) = p$), Ext returns $\perp$ equals

$$\Pr(\mathcal{E}rr_{ext}) = \sum_{i=0}^{l-1} \binom{\frac{m}{2}}{i} (2p(1-p))^i \,(1-2p(1-p))^{\frac{m}{2}-i}. \tag{7}$$

An (n, k) binary error correcting channel code: We denote the encoding and the decoding functions by $Enc : \{0,1\}^k \rightarrow \{0,1\}^n$ and $Dec : \{0,1\}^n \rightarrow \{0,1\}^k$, respectively. There are efficient (n, k) error correcting codes that can correct nearly up to $t = (n-k)/2$ bits of error. When used over a binary symmetric channel with error probability p, the decoding error probability of such codes equals the probability that the number of errors is greater than t, i.e.,

$$\Pr(\mathcal{E}rr_{enc}) \geq \Pr(n_{err} > t) = \sum_{i=t+1}^{n} \binom{n}{i} p^i \,(1-p)^{n-i}. \tag{8}$$

Universal class of hash functions: A class $\mathcal{H}$ of (hash) functions $h : \mathcal{A} \rightarrow \mathcal{B}$ is universal [9], if for any distinct pair of inputs $x_1, x_2 \in \mathcal{A}$, the equality $h(x_1) = h(x_2)$ happens with probability at most $1/|\mathcal{B}|$, provided that h is uniformly at random selected from $\mathcal{H}$. For the purpose of our SKE construction design, we use the following universal class of computationally efficient hash functions, proposed in [33],

$$\mathcal{H} = \{h_c : GF(2^k) \rightarrow \{0,1\}^s, \quad c \in GF(2^k)\},$$

where $h_c(x)$ returns the first s bits of $c.x$, and the multiplication is over the polynomial representation of $GF(2^k)$.

Protocol description: Using the above primitives, the SKE protocol proceeds as follows. Alice sends her constant sequence $\mathbf{X}_f = \mathbf{a} = (\underline{0})^m$ over the forward DMBC. Bob and Eve receive the m-sequences $\mathbf{Y}_f$ and $\mathbf{Z}_f$ (m is even). Bob views this as an m-bit Bernoulli sequence, $\mathbf{Y}_f = (Y_{f,1}, \ldots, Y_{f,m})$, with $P(Y_{f,i} = 1) = p_1$ and finds $\mathbf{U} = Ext(\mathbf{Y}_f)$. If $\mathbf{U} = \perp$, the error $\mathcal{E}rr_{ext}$ occurs; otherwise, Bob splits the l-bit $\mathbf{U}$ into two independent and uniform k-bit sequences $\mathbf{U}_1$ and $\mathbf{U}_2$, where $k = l/2$. He calculates the n-bit codewords $\mathbf{X}_{1b} = Enc(\mathbf{U}_1)$ and $\mathbf{X}_{2b} = Enc(\mathbf{U}_2)$ and sends them over the backward DMBC; Alice and Eve receive $(\mathbf{Y}_{1b}, \mathbf{Y}_{2b})$ and $(\mathbf{Z}_{1b}, \mathbf{Z}_{2b})$, respectively. Alice calculates the k-sequences $\hat{\mathbf{U}}_1 = Dec(\mathbf{Y}_{1b})$ and $\hat{\mathbf{U}}_2 = Dec(\mathbf{Y}_{2b})$. The error event $\mathcal{E}rr_{enc1}$ (resp. $\mathcal{E}rr_{enc2}$) occurs when $\hat{\mathbf{U}}_1 \neq \mathbf{U}_1$ (resp. $\hat{\mathbf{U}}_2 \neq \mathbf{U}_2$). Next, Alice and Bob use universal hashing for privacy amplification, i.e., to derive keys that are secure against Eve. The secret key is $S = h_C(\mathbf{U}_1)$ where $C = \mathbf{U}_2$. Bob calculates $S_B = S$ and Alice calculates $S_A = h_{\hat{C}}(\hat{\mathbf{U}}_1)$ where $\hat{C} = \hat{\mathbf{U}}_2$.

Analysis of randomness, reliability, and secrecy: The above protocol provides Alice and Bob with s uniformly random bits of key. The rate of key establishment is calculated as the number of the key bits divided by the number of channel uses, i.e., $R_{sk} = \frac{s}{m+2n}$.

Regarding the reliability requirement (6b), we observe that $S_A = S_B = S$ holds if none of the errors $\mathcal{E}rr_{ext}$, $\mathcal{E}rr_{enc1}$, and $\mathcal{E}rr_{enc2}$ occurs. This gives

$$\Pr(S_A = S_B = S) \geq 1 - \Pr(\mathcal{E}rr_{ext}) - \Pr(\mathcal{E}rr_{enc1}) - \Pr(\mathcal{E}rr_{enc2}), \tag{9}$$

where $\Pr(\mathcal{E}rr_{ext})$, $\Pr(\mathcal{E}rr_{enc1}) = \Pr(\mathcal{E}rr_{enc2})$ are obtained from (7) and (8) for $p = p_1$, respectively. For an arbitrarily small $\delta > 0$, we can, for instance, choose the parameters m, l, n, and $k = l/2$ such that each of the above error probabilities is at most $\delta/3$ and so (6b) is satisfied.

To argue the secrecy of the key, we use the following lemma.

Lemma 1. *[6, Corollary 4] For a random k-sequence $\mathbf{U}_1$, if the conditional Rényi entropy $R(\mathbf{U}_1|\mathbf{Z}_{1b} = \mathbf{z})$ is lower bounded by s_0 and $S = h_C(\mathbf{U}_1)$ for a uniformly random C, then*

$$H(S|\mathbf{Z}_{1b} = \mathbf{z}, C) \geq s - \frac{2^{s-s_0}}{\ln 2}. \tag{10}$$

Since the channels are memoryless, for large enough n, from asymptotic equipartition property (AEP) for the sequences $\mathbf{U}_1$ and $\mathbf{Z}_{1b}$ (see, e.g., [10, Chapter 3]), we can replace the Rényi entropy $R(\mathbf{U}_1|\mathbf{Z}_{1b} = \mathbf{z})$ in the above by the Shannon entropy as $H(\mathbf{U}_1|\mathbf{Z}_{1b})$, which we calculate below.

$$H(\mathbf{U}_1|\mathbf{Z}_{1b}) = H(\mathbf{U}_1) - H(\mathbf{Z}_{1b}) + H(\mathbf{Z}_{1b}|\mathbf{U}_1) = k - H(\mathbf{Z}_{1b}) + nh(p_2) \geq k - n(1 - h(p_2)). \tag{11}$$

Using Lemma 1 and letting $s_0 = k - n(1 - h(p_2))$, we calculate Eve's uncertainty about the secret key as

$$\begin{aligned} &H(S|\mathbf{Z}_{1b}, \mathbf{Z}_{2b}, \mathbf{Z}_f) \overset{(a)}{=} H(S|\mathbf{Z}_{1b}, \mathbf{Z}_{2b}) \geq H(S|\mathbf{Z}_{1b}, \mathbf{U}_2) = H(S|\mathbf{Z}_{1b}, C) \geq s - \frac{2^{s-k+n(1-h(p_2))}}{\ln 2} \\ &\Rightarrow \frac{H(S|\mathbf{Z}_{1b}, \mathbf{Z}_{2b}, \mathbf{Z}_f)}{H(S)} \geq 1 - \frac{2^{s-k+n(1-h(p_2))}}{s \ln 2} \end{aligned} \tag{12}$$

Equality (a) holds since the randomness in $\mathbf{Z}_f$ comes only from Eve's BSC noise that is independent of all the variables including $(S, \mathbf{Z}_{1b}, \mathbf{Z}_{2b})$. For an arbitrarily small $\delta > 0$, we can choose the parameters k, n and s for (12) such that the secrecy requirement in (6c) holds.

Table 1 shows the construction parameters for SKE over binary symmetric channels with $p_1 = 0.1$ and $p_2 = 0.2$ when the secret key length is $s = 100$ and the security parameter δ has different values. According to this table, the achievable SK rate by this construction is about $R_{sk} = 0.015$ bits per channel use.

δ	n	k	l	m	R_{sk}
10^{-1}	404	300	600	5230	0.0166
10^{-2}	458	330	660	5430	0.0158
10^{-3}	508	358	716	5590	0.0151
10^{-4}	560	388	776	5730	0.0146

Table 1. The SKE construction parameters with respect to different values of δ for $s = 100$.

Remark 2. In each round of the above construction,z either Alice or Bob sends a sequence over the channel. Assuming the full-duplex communication model, Alice and Bob can follow another run of the protocol in parallel, this time with Bob as the initiator. This will double the secret key rate and so, for the values of $p_1 = 0.1$ and $p_2 = 0.2$, the SK rate achievable by this construction is around 0.03 bit per channel use.

Remark 3. The aim of the above construction is to show the feasibility of efficient SKE with no initial randomness. We have chosen simple primitives for the ease of explanation. Using more complex primitives in the above construction, one may achieve higher secret key rates.

4 Bounds on the SK capacity

We provide lower and upper bounds on the SK capacity as defined in Section 2. Let the RVs X_f, Y_f, Z_f and X_b, Y_b, Z_b correspond to the channel probability distributions $P_{Y_f,Z_f|X_f}$ and $P_{Y_b,Z_b|X_b}$, respectively.

Theorem 1. *The SK capacity is lower bounded as*

$$C_{sk}^{2DMBC} \geq \max_{\mu \geq 0, P_{X_f}, P_{X_b}} \{Lbound_A + Lbound_B\}, \tag{13}$$

where

$$Lbound_A = \frac{1}{1+\mu}\left(\mu(I(Y_b;X_b) - I(Y_b;Z_b)) + \gamma_1(I(X_f;Y_f) - I(X_f;Z_f))_+\right), \tag{14}$$

$$Lbound_B = \frac{1}{1+\mu}\left(\mu(I(Y_f;X_f) - I(Y_f;Z_f)) + \gamma_2(I(X_b;Y_b) - I(X_b;Z_b))_+\right), \tag{15}$$

for

$$\gamma_1 = \min\{1, \frac{H(Y_b|X_b,Z_b) + \mu(H(Y_b|X_b) - H(X_f))}{I(X_f;Y_f)}\}, \tag{16}$$

$$\gamma_2 = \min\{1, \frac{H(Y_f|X_f,Z_f) + \mu(H(Y_f|X_f) - H(X_b))}{I(X_b;Y_b)}\}, \tag{17}$$

such that

$$H(Y_b|X_b,Z_b) > \mu H(X_f), \qquad I(X_f;Y_f) > \mu H(Y_b|X_b), \tag{18}$$

$$H(Y_f|X_f,Z_f) > \mu H(X_b), \qquad I(X_b;Y_b) > \mu H(Y_f|X_f). \tag{19}$$

Proof. See Section 5 and Appendix A.

The lower bound (13) is achieved by the so-called main protocol. The main protocol consists of an initialization round followed by iteration of a two round protocol, called the basic protocol. Each iteration of the basic protocol uses some randomness and generates new randomness for the next iteration, together with a new part of secret key. The initialization round provides the initial randomness for the first iteration of the basic protocol. As the number of iterations increases, the SK rate of the main protocol approaches the lower bound, which is, in fact, the SK rate of the basic protocol. In the full-duplex channel model, the basic protocol proceeds as two parallel instances of a two-round sub-protocol: The first (resp. second) instance is initiated by Alice (resp. Bob) and achieves the key rate $Lbound_A$ (resp. $Lbound_B$), for fixed values μ, P_{X_f}, and P_{X_b} that are chosen to maximize (13). Each of the key rates, $Lbound_A$ and $Lbound_B$, is the sum of two terms, each corresponding to the key rate achievable in one round of the basic protocol (see (14)-(15)). The real value μ is the ratio between the number of channel uses in the first and the second rounds, e.g., $\mu = 0$ implies no channel use in the first round, implying a one-round basic protocol. The real values γ_1 and γ_2 are to relate the amount of achievable key rate as a function of the randomness obtained from channel noise.

As mentioned above, each round of the basic protocol generates some key rates. The keys rate achieved by the second round depends on the DMBC parameters (i.e., $I(X_f;Y_f) - I(X_f;Z_f)$ and $I(X_b;Y_b) - I(X_b;Z_b)$), and the key rate achieved in the first round depends on the "inverse" DMBC (see Definition 8) parameters (i.e., $I(Y_f;X_f) - I(Y_f;Z_f)$ and $I(Y_b;X_b) - I(Y_b;Z_b)$). We refer to Section 5 for more details. When the DMBCs are in favor of Alice and Bob, i.e., $I(X_f;Y_f) - I(X_f;Z_f)$ and $I(X_b;Y_b) - I(X_b;Z_b)$ are positive, $Lbound_A$ and $Lbound_B$ will be positive by simply choosing $\mu = 0$. This implies a positive SK capacity. When the channels are in favor of Eve, the lower bound may remain positive (for some values of $\mu > 0$) if one of the inverse DMBCs is in favor of Alice and Bob. The study of the lower bound for BSCs in Section 6 shows clearly the existence positive SK rates in the latter case (see Fig. 3).

Theorem 2. *The SK capacity is upper bounded as*

$$C_{sk}^{2DMBC} \leq \max_{P_{X_f}, P_{X_b}} \{Ubound_A + Ubound_B\}, \tag{20}$$

where

$$Ubound_A = \min\{H(Y_b|X_b,Z_b), I(X_f;Y_f|Z_f)\}, \quad \textit{and} \quad Ubound_B = \min\{H(Y_f|X_f,Z_f), I(X_b;Y_b|Z_b)\}. \tag{21}$$

Proof. See Appendix B.

The above upper bound also proves the SKE impossibility results for the special cases discussed in Section 3.1. In the case of one way communication, e.g., when the backward channel returns constant values at its ouputs, both terms $I(X_b;Y_b|Z_b)$ and $H(Y_b|X_b,Z_b)$ equal zero, implying a zero upper bound on SK rates. The same argument can be used to prove impossibility when the backward channel is noiseless and public or it is noisy but returns identical outputs to Alice and Eve.

Theorem 3 shows that the two bounds coincide when the two DMBCs do not leak information. The resulting value matches the common randomness capacity of a pair of independent DMCs, given in [31].

Theorem 3. *When the DMBCs do not leak information to Eve, the bounds coincide and the SK capacity equals*

$$C_{sk}^{2DMBC} = \max_{P_{X_f},P_{X_b}} \{\min\{H(Y_b|X_b), I(X_f;Y_f)\} + \min\{H(Y_f|X_f), I(X_b;Y_b)\}\}. \tag{22}$$

Proof. See Appendix C.

5 The main SKE Protocol: Achieving the Lower Bound

We noted that the bound in Theorem 1 is achieved by the *main protocol.* The main protocol has $2t+1$ rounds and does not need any initial randomness. The protocol starts with an initialization round (round 0) that provides Alice and Bob with some amount of independent randomness. The initialization round is followed by t iterations of a two-round protocol, called the *basic protocol.* Each iteration of the basic protocol takes some independent randomness from Alice and Bob and returns to them a part of the secret key as well as new pieces of independent randomness. The independent randomness that is produced in iteration $1 \le r \le t-1$ (resp. round 0) will be used in iteration $r+1$ (resp. iteration 1). The secret key parts are finally concatenated to give the final secret key. In a deeper look, the basic protocol proceeds as two parallel instances of a key agreement sub-protocol, one initiated by Alice and one initiated by Bob. Each instance of the sub-protocol uses a part of the randomness provided by Alice and Bob, and partially contributes to the secret key. More details are provided in Section 5.2.

The main protocol relies on the existence of two primitives, referred to as *secure equipartition* and *secure block code.* In the following, we define these primitives, show their existence, and then describe the main protocol.

5.1 Preliminaries

Definition 3. *For a probability distribution P_X over the set $\mathcal{X}$, a sequence $x^n \in \mathcal{X}^n$ is called ϵ-typical if*

$$|-\frac{1}{n}\log P(x^n) - H(X)| < \epsilon,$$

where $P(x^n)$ is calculated as

$$P(x^n) = \prod_{i=1}^{n} P(x_i).$$

Definition 4. *An (n,M,ϵ)*-block code *for the DMC $(\mathcal{X},\mathcal{Y},P_{Y|X})$ is a set $\{(c_i,\mathcal{C}_i)_{i=1}^M\}$ such that $c_i \in \mathcal{X}^n$, $(\mathcal{C}_i)_{i=1}^M$ partitions $\mathcal{Y}^n$, and $P_{Y|X}^n(Y^n = \mathcal{C}_i|X^n = c_i) \ge 1-\epsilon$.*

Block codes are used to promise reliable communication over noisy channels (DMCs). The following lemma shows the existence of block codes, for a DMC $(\mathcal{X},\mathcal{Y},P_{Y|X})$, that achieve reliable communication rates up to $I(X;Y)$.

Lemma 2. *For any P_X, $R_c < I(X;Y)$, and large enough n, there exists an (n,M,ϵ)-block code for the DMC $(\mathcal{X},\mathcal{Y},P_{Y|X})$ with ϵ-typical codewords $c_i \in \mathcal{X}^n$ such that $M = \lfloor 2^{nR_c} \rfloor$ and $\epsilon = 2^{n(R_c - I(X;Y))} \to 0$.*

Proof. See e.g. [10, 20].

We define a *secure block code* for a DMBC as the composition of a block code and a function that we refer to as a *key derivation function.* A secure block code can be used by two parties, connected through a DMBC, to generate a secret key securely.

Definition 5. *An* (n, M, K, ϵ)-secure block code*, with* $K \leq M$*, for the DMBC* $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{YZ|X})$ *consists of an* (n, M, ϵ)*-block code for the DMC* $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ *as above, a partition of* $(c_i)_{i=1}^M$ *into* $(\mathcal{K}_j)_{j=1}^K$*, and a key derivation function* $\phi_s : (c_i)_{i=1}^M \to [K]$ *defined as* $\phi_s(c_i) = j$ *iff* $c_i \in \mathcal{K}_j$*, such that if* X^n *is uniformly selected from* $(c_i)_{i=1}^M$ *and* $S = \phi_s(X^n)$ *then* $H(S|Z^n)/\log K \geq 1 - \epsilon$*.*

Although the above definition of a secure block code as a primitive is new to the literature, the work on secure message transmission or key agreement over one-way DMBCs [14,34] implicitly studies the existence of such a primitive. For instance, one can send a message $S \in [K]$ using a secure block code (defined as above), by randomly choosing a codeword in $\phi_s^{-1}(S)$ and sending it over the channel. The receiver decodes the codeword and applies ϕ_s to obtain the secure message. The results in [14, 34] let us conclude the following.

Lemma 3. *For any* P_X*,* $R_c < I(X;Y)$*,* $R_{sc} < R_c - I(X;Z)$*, and large enough* n*, there exists an* (n, M, K, ϵ)*-secure block code for a DMBC* $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{YZ|X})$ *with* ϵ*-typical codewords* c_i *such that* $M = \lfloor 2^{nR_c} \rfloor$*,* $K = \lfloor 2^{nR_{sc}} \rfloor$*, and* $\epsilon = \max\{2^{n(R_c - I(X;Y))}, 2^{n(R_{sc} - (R_c - I(X;Z)))}\} \to 0$*.*

Proof. See [34, Theorem 2] and [14, Corrollary 1].

Lemma 3 indicates that, for the above DMBC, there exists a secure block code that achieves key rates up to $I(X;Y) - I(X;Z)$. In the following, we extend this result by showing that there are sufficiently many secure block codes such that any $X^n \in \mathcal{X}^n$ as input to the channel belongs to at least one of them, with high probability.

Lemma 4. *For any* P_X*,* $R_c < I(X;Y)$*,* $R_{sc} < R_c - I(X;Z)$*, large enough* $R' > H(X) - R_c$*, and large enough* n*, there exist* N *(not necessarily disjoint)* (n, M, K, ϵ)*-secure block codes for the DMBC* $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{YZ|X})$ *with* ϵ*-typical codewords, such that* $M = \lfloor 2^{nR_c} \rfloor$*,* $K = \lfloor 2^{nR_{sc}} \rfloor$*,* $N = \lfloor 2^{nR'} \rfloor$*, and* $\epsilon = \max\{2^{n(R_c - I(X;Y))}, 2^{n(R_{sc} - (R_c - I(X;Z)))}\} \to 0$*; furthermore, the probability that a randomly selected* ϵ*-typical sequence* $X^n \in \mathcal{X}^n$ *belongs to at least one of the codes is at least* $1 - e^{-\gamma}$*, where* $\gamma = 2^{n(R' + R_c - H(X) - \epsilon)} \to \infty$*.*

Proof. See Appendix D.

An equipartition is used to derive uniform randomness from the output of a DMC, such that the randomness that is independent of the input. We remind that, in the SKE construction over BSCs, we used the von Neumann extractor for this purpose.

Definition 6. *An* (M, ϵ)-equipartition *of* $\mathcal{C} \subseteq \mathcal{Y}^n$ *w.r.t.* $c \in \mathcal{X}^n$ *for the DMC* $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ *is a partition* $\{\mathcal{C}(e), \mathcal{C}(1), \ldots, \mathcal{C}(M)\}$ *such that* $P_{Y|X}^n(Y^n = \mathcal{C}(j)|X^n = c)$ *is the same for all* $1 \leq j \leq M$ *and* $P_{Y|X}^n(Y^n = \mathcal{C}(e)|X^n = c) \leq \epsilon$*.*

The following lemma shows that there exists an equipartition for a DMC $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ that can derive randomness rates up to $H(Y|X)$ bits per channel use. This implies the noisier the channel, the higher the achievable rate of randomness that is independent from the channel input.

Lemma 5. *For any* P_X*, typical* $c \in \mathcal{X}^n$*,* $\mathcal{C} \subseteq \mathcal{Y}^n$*, large enough* n*, and* $R_e < H(Y|X)$*, there exists an* (M, ϵ)*-equipartition over the DMC* $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ *such that* $M = \lfloor 2^{nR_e} \rfloor$*,* $\epsilon = 2^{n(R_e - H(Y|X))} \to 0$*. Furthermore, each part has size at most* $2^{n\epsilon}|\mathcal{C}|/M$*.*

Proof. See [31, Lemma 3.2].

For a DMBC, a secure equipartition is used to ensure that the derived randomness is not only independent of the input but also independent of Eve's received sequence. In other words, Eve is uncertain about this random value.

Definition 7. *An* (M, ϵ)-secure equipartition *of* $\mathcal{C} \subseteq \mathcal{Y}^n$ *w.r.t.* $c \in \mathcal{X}^n$ *over the DMBC* $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{YZ|X})$ *is an* (M, ϵ)*-equipartition of* $\mathcal{C}$ *for the DMC* $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ *and a randomness derivation function* $\psi_t : \mathcal{C} \to [M] \cup \perp$ *defined as*

$$\psi_t(y^n) = \begin{cases} j, & y^n \in \mathcal{C}(j) \\ \perp & y^n \in \mathcal{C}(e) \end{cases},$$

such that if $X^n = c$ *and* $T = \psi_t(Y^n)$*, then*

$$H(T|X^n = c, Z^n)/\log M \geq 1 - \epsilon. \tag{23}$$

The following lemma shows the existence of a secure equipartition over the DMBC that achieves randomness rates up to $H(Y|XZ)$ bits per channel use.

Lemma 6. *For any P_X, typical $c \in \mathcal{X}^n$, $\mathcal{C} \subseteq \mathcal{Y}^n$ of size less than $2^{nH(Y)}$, $R_{se} < H(Y|XZ)$, and large enough n, there exists an (M, ϵ)-secure equipartition for the DMBC $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{YZ|X})$ such that $M = \lfloor 2^{nR_{se}} \rfloor$ and*

$$\epsilon = \frac{3I(Y;X,Z)h(\epsilon')}{H(Y|XZ) - \epsilon'} \to 0, \quad \textit{where} \quad \epsilon' = 2^{n(R_{se} - H(Y|XZ))}.$$

Proof. See Appendix E.

To describe of the main protocol, we shall use the notion of an inverse DMBC that implies a virtual channel defined as follows.

Definition 8. *Given a distribution P_X, for a DMBC $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{YZ|X})$, we define its corresponding inverse DMBC as $(\mathcal{Y}, \mathcal{X}, \mathcal{Z}, P_{XZ|Y})$ such that $P_{XZ|Y}$ is calculated as*

$$P_{XZ|Y} = \frac{P_X . P_{YZ|X}}{P_Y}, \qquad \textit{where} \;\; P_Y = \sum_{x,z} P_X . P_{YZ|X}.$$

5.2 Description of the main protocol

Let P_{X_f}, P_{X_b}, and μ be chosen such that the conditions (18) and (19) are satisfied. The conditions can be rephrased as

$$n_2 H(Y_b|X_b, Z_b) \geq n_1(H(X_f) + \alpha), \qquad n_2 I(X_f; Y_f) \geq n_1(H(Y_b|X_b) + \alpha), \tag{24}$$

$$n_2 H(Y_f|X_f, Z_f) \geq n_1(H(X_b) + \alpha), \qquad n_2 I(X_b; Y_b) \geq n_1(H(Y_f|X_f) + \alpha), \tag{25}$$

where $\alpha > 0$ is a sufficiently small real constant, to be determined from δ in the sequel, and n_1 and n_2 are sufficiently large positive integers such that $n_1 = \mu n_2$, and $1/\alpha = o(\min\{n_1, n_2\})$; in other words, $2^{-\alpha \min\{n_1, n_2\}}$ approaches zero.

In the following, we define a number of integer and set parameters and claim the existence of secure block codes and secure equipartitions using these parameters. Next, we describe the construction of the main protocol based on the given primitives. Define

$$\begin{aligned} &R_{1f} = H(X_f) - \alpha, && R_{cf} = I(X_f; Y_f) - \alpha, && R_{scf} = I(X_f; Y_f) - I(X_f; Z_f) - 2\alpha, \\ &R_{ef} = H(Y_f|X_f), && R^+_{ef} = H(Y_f|X_f) + 2\alpha, && R_{sef} = H(Y_f|X_f, Z_f) - \alpha, \\ &R_{scf^{-1}} = I(Y_f; X_f) - I(Y_f; Z_f) - 2\alpha. \end{aligned} \tag{26}$$

We informally describe each of the above quantities as follows. For the forward DMBC, R_{1f} is the (highest) channel input rate, R_{cf} is the rate of reliable transmission, R_{scf} is the rate of secure transmission, R_{ef} is the equipartition rate (or the uncertainty rate of the channel), and R_{sef} is the secure equipartition rate. Note that R_{cf} can also be viewed as the rate of reliable transmission for the inverse forward DMBC (see Definition 8). Finally, $R_{scf^{-1}}$ is the secure transmission rate of the inverse forward DMBC. One can define similar quantities for the backward DMBC.

$$\begin{aligned} &R_{1b} = H(X_b) - \alpha, && R_{cb} = I(X_b; Y_b) - \alpha, && R_{scb} = I(X_b; Y_b) - I(X_b; Z_b) - 2\alpha, \\ &R_{eb} = H(Y_b|X_b), && R^+_{eb} = H(Y_b|X_b) + 2\alpha, && R_{seb} = H(Y_b|X_b, Z_b) - \alpha, \\ &R_{scb^{-1}} = I(Y_b; X_b) - I(Y_b; Z_b) - 2\alpha. \end{aligned} \tag{27}$$

Each iteration of the two-round basic protocol uses the 2DMBC channel n_1 times in the first round and n_2 times in the second round; i.e. in total $n_1 + n_2$. In the second round, Alice (resp. Bob) sends two sequences of lengths n_{21A} and n_{22A} (resp. n_{21B} and n_{22B}), where $n_{21A} + n_{22A} (= n_{21B} + n_{22B}) = n_2$ and,

$$n_{21A} = \frac{1}{R_{cf}} \min\{n_2 R_{cf}, n_2 R_{seb} + n_1 R_{eb} - n_1 R_{1f}\}, \tag{28}$$

$$n_{21B} = \frac{1}{R_{cb}} \min\{n_2 R_{cb}, n_2 R_{sef} + n_1 R_{ef} - n_1 R_{1b}\}. \tag{29}$$

Using the above quantities, we define,

$$\begin{array}{lll}
M_{1A}=\lfloor 2^{n_1 R_{cb}}\rfloor, & M_{21A}=\lfloor 2^{n_{21A}R_{cf}}\rfloor, & \\
K_{1A}=\lfloor 2^{n_1 R_{scb-1}}\rfloor, & K_{21A}=\lfloor 2^{n_{21A}R_{scf}}\rfloor, & \\
N_A=\lfloor 2^{n_1 R_{eb}^{+}}\rfloor, & & \\
L_{1A}=\lfloor 2^{n_1 R_{1f}}\rfloor, & L_{2A}=\lfloor 2^{n_{21A}R_{cf}-n_1 R_{eb}}\rfloor, & L_A=L_{1A}.L_{2A}, \\
\Gamma_{21A}=\min\{L_A,\lfloor 2^{n_{21B}R_{seb}}\rfloor\}, & \Gamma_{22A}=\lfloor 2^{n_{22B}R_{seb}}\rfloor, & \Gamma_A=\Gamma_{21A}.\Gamma_{22A}.
\end{array} \tag{30}$$

$$\begin{array}{lll}
M_{1B}=\lfloor 2^{n_1 R_{cf}}\rfloor, & M_{21B}=\lfloor 2^{n_{21B}R_{cb}}\rfloor, & \\
K_{1B}=\lfloor 2^{n_1 R_{scf-1}}\rfloor, & K_{21B}=\lfloor 2^{n_{21B}R_{scb}}\rfloor, & \\
N_B=\lfloor 2^{n_1 R_{ef}^{+}}\rfloor, & & \\
L_{1B}=\lfloor 2^{n_1 R_{1b}}\rfloor, & L_{2B}=\lfloor 2^{n_{21B}R_{cb}-n_1 R_{ef}}\rfloor, & L_B=L_{1B}.L_{2B}, \\
\Gamma_{21B}=\min\{L_B,\lfloor 2^{n_{21A}R_{sef}}\rfloor\}, & \Gamma_{22B}=\lfloor 2^{n_{22A}R_{sef}}\rfloor, & \Gamma_B=\Gamma_{21B}.\Gamma_{22B}.
\end{array} \tag{31}$$

Using (26)-(30), one can observe that $L_A=\Gamma_A$ and $L_B=\Gamma_B$ in the above. Let the set $\mathcal{X}_{f,\epsilon}^{n_1}=\{\mathbf{x}_{f,1},\ldots,\mathbf{x}_{f,L_{1A}}\}$ be obtained by independently selecting L_{1A} sequences in $\mathcal{X}_f^{n_1}$. Similarly define $\mathcal{X}_{b,\epsilon}^{n_1}=\{\mathbf{x}_{b,1},\ldots,\mathbf{x}_{b,L_{1B}}\}\subseteq\mathcal{X}_b^{n_1}$. Let Alice and Bob have two fixed public integers $u_a\in[\Gamma_{21A}]$ and $u_b\in[\Gamma_{21B}]$ as well as two fixed public sequences $\mathbf{a}\in\mathcal{X}_f^{n_{22A}}$ and $\mathbf{b}\in\mathcal{X}_b^{n_{22B}}$, respectively. Let $u_{A,split}:[\Gamma_{21A}]\times[\Gamma_{22A}]\to[L_{1A}]\times[L_{2A}]$ and $u_{B,split}:[\Gamma_{21B}]\times[\Gamma_{22B}]\to[L_{1B}]\times[L_{2B}]$ be arbitrary bijective mappings.

For given P_{X_f} and P_{X_b}, define the inverse DMBCs $(\mathcal{Y}_f,\mathcal{X}_f,\mathcal{Z}_f,P_{X_f,Z_f|Y_f})$ and $(\mathcal{Y}_b,\mathcal{X}_b,\mathcal{Z}_b,P_{X_b,Z_b|Y_b})$ according to Definition 8. Letting

$$\epsilon=2^{-\min(n_1,n_{21A},n_{21B})\alpha}\to 0 \qquad \text{and} \qquad \gamma=2^{n_1(\alpha-\epsilon)}\to\infty,$$

and using Lemmas 3, 4, and 6 we arrive at the existence of the following primitives to be used in the main protocol.
Secure block codes over the inverse channels (see Lemma 4):

- For the inverse forward DMBC $(\mathcal{Y}_f,\mathcal{X}_f,\mathcal{Z}_f,P_{X_f,Z_f|Y_f})$, there exist N_B $(n_1,M_{1B},K_{1B},\epsilon)$-secure block codes $\{(d_{f,i}^j,\mathcal{D}_{f,i}^j)_{i=1}^{M_{1B}}\ :\ 1\le j\le N_B\}$ with the key derivation functions $\phi_{s,B}^j$, such that a randomly selected ϵ-typical sequence in $\mathcal{Y}_f^n$ is in at least one of the codes with probability at least $1-e^{-\gamma}$.
- For the inverse backward DMBC $(\mathcal{Y}_b,\mathcal{X}_b,\mathcal{Z}_b,P_{X_b,Z_b|Y_b})$, there exist N_A $(n_1,M_{1A},K_{1A},\epsilon)$-secure block codes $\{(d_{b,i}^j,\mathcal{D}_{b,i}^j)_{i=1}^{M_{1A}}\ :\ 1\le j\le N_A\}$ with the key derivation functions $\phi_{s,A}^j$, such that a randomly selected ϵ-typical sequence in $\mathcal{Y}_b^n$ is in at least one of the codes with probability at least $1-e^{-\gamma}$.

Secure block codes and corresponding secure equipartitions over the channels (see Lemmas 3 and 6):

- For the forward DMBC $(\mathcal{X}_f,\mathcal{Y}_f,\mathcal{Z}_f,P_{Y_f,Z_f|X_f})$, there exists an $(n_{21A},M_{21A},K_{21A},\epsilon)$-secure block code $\{(c_{f,i},\mathcal{C}_{f,i})_{i=1}^{M_{21A}}\}$ with the key derivation function $\phi_{s,A}$; furthermore, for each $(c_{f,i},\mathcal{C}_{f,i})$ there exists a (Γ_{21B},ϵ)-secure equipartition $\{\mathcal{C}_{f,i}(e),\mathcal{C}_{f,i}(1),\ldots,\mathcal{C}_{f,i}(\Gamma_{21B})\}$ with the randomness derivation function ψ_B^i.
- For the backward DMBC $(\mathcal{X}_b,\mathcal{Y}_b,\mathcal{Z}_b,P_{Y_b,Z_b|X_b})$, there exists an $(n_{21B},M_{21B},K_{21B},\epsilon)$-secure block code $\{(c_{b,i},\mathcal{C}_{b,i})_{i=1}^{M_{21B}}\}$ with the key derivation function $\phi_{s,B}$; furthermore, for each $(c_{b,i},\mathcal{C}_{b,i})$ there exists a (Γ_{21A},ϵ)-secure equipartition $\{\mathcal{C}_{b,i}(e),\mathcal{C}_{b,i}(1),\ldots,\mathcal{C}_{b,i}(\Gamma_{21A})\}$ with the randomness derivation function ψ_A^i.

Secure equipartitions (for transmission of the constant values) over the channels (see Lemma 6):

- For the forward DMBC $(\mathcal{X}_f,\mathcal{Y}_f,\mathcal{Z}_f,P_{Y_f,Z_f|X_f})$, for $(\mathbf{a},\mathcal{Y}_f)$, there exists a (Γ_{22B},ϵ)-secure equipartition $\{\mathcal{Y}_f(e),\mathcal{Y}_f(1),\ldots,\mathcal{Y}_f(\Gamma_{22B})\}$ with the randomness derivation function ψ_B.
- For the backward DMBC $(\mathcal{X}_b,\mathcal{Y}_b,\mathcal{Z}_b,P_{Y_b,Z_b|X_b})$, for $(\mathbf{b},\mathcal{Y}_b)$, there exists a (Γ_{22A},ϵ)-secure equipartition $\{\mathcal{Y}_b(e),\mathcal{Y}_b(1),\ldots,\mathcal{Y}_b(\Gamma_{22A})\}$ with the randomness derivation function ψ_A.

Using the above primitives, we describe the main protocol below.
The initialization round (round 0): The initialization round proceeds as two parallel instances. The first and the second instances are to derive independent randomness for Bob and Alice, respectively; neither of them, however, produces a secret key. The first instance runs as follows. Alice sends the constant n_2-sequence $\mathbf{X}_f^{:0}=(c_{f,u_a}||a)$ over

the forward DMBC; Bob and Eve receive the noisy versions $\mathbf{Y}_f^{:0} = (\mathbf{Y}_{1f}||\mathbf{Y}_{2f})$ and $\mathbf{Z}_f^{:0}$, respectively. Bob calculates $U_B^{:0} = (\psi_B^{u_a}(\mathbf{Y}_{1f})||\psi_B(\mathbf{Y}_{2f}))$ as independent randomness to be used in the first iteration of the basic protocol. He then splits this into two parts as $(U_{1B}^{:0}, U_{2B}^{:0}) = u_{B,split}(U_B^{:0})$. The first and the second parts are respectively used in the first and the second rounds of iteration 1.

In parallel to the above, the second instance runs as follows. Bob sends the constant n_2-sequence $\mathbf{X}_b^{:0} = (c_{b,u_b}||b)$ over the backward DMBC; Alice and Eve receive $\mathbf{Y}_b^{:0} = (\mathbf{Y}_{1b}||\mathbf{Y}_{2b})$ and $\mathbf{Z}_b^{:0}$, respectively. Alice calculates $U_A^{:0} = (\psi_A^{u_b}(\mathbf{Y}_{1b})||\psi_A(\mathbf{Y}_{2b}))$, as independent randomness, and splits it into $(U_{1A}^{:0}, U_{2A}^{:0}) = u_{A,split}(U_A^{:0})$, where the first and the second parts are respectively used in the first and the second rounds of iteration 1.

The basic protocol (iteration $1 \leq r \leq t$): Each iteration r of the basic protocol proceeds as two parallel instances of a two-round key agreement sub-protocol over the full-duplex communication channel. Each instance runs in two rounds, $2r-1$ and $2r$, where the 2DMBC is used n_1 and n_2 times, respectively. Each instance receives pieces of randomness from Alice and Bob and returns to them a piece of secret key. Furthermore, the first and the second instances are initiated by Alice and Bob and return new pieces of independent randomness to Alice and Bob, respectively. The new randomness is used in the next iteration of the basic protocol. Fig. 2 summarizes the relationship between the random variables that are used in the first instance in iteration r of the basic protocol.

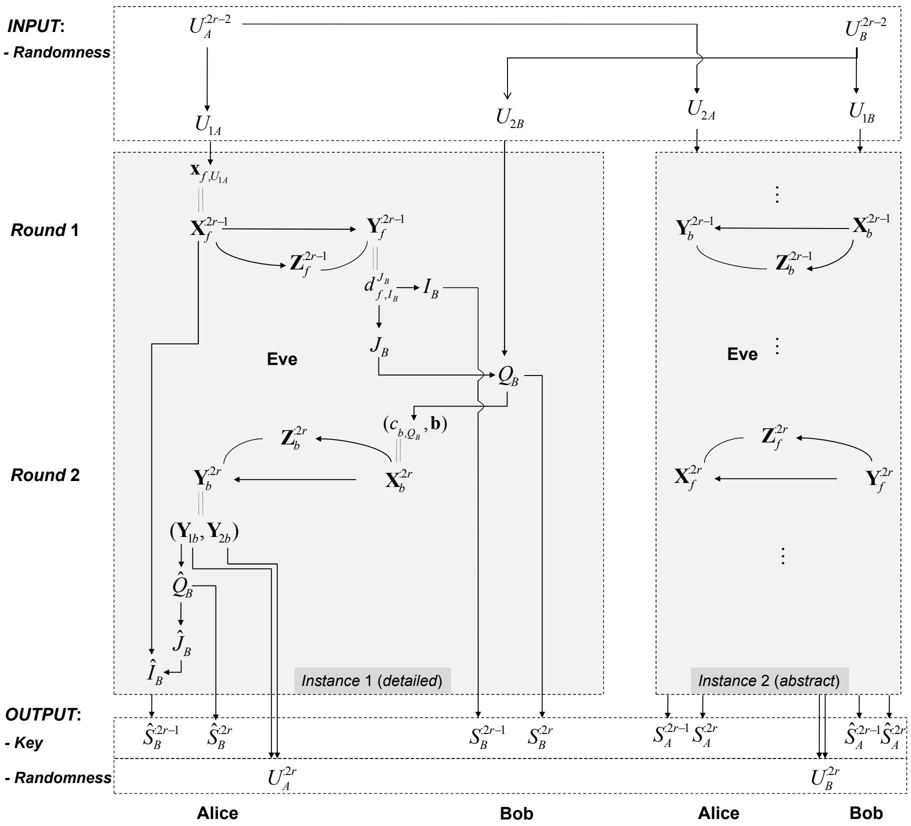


Fig. 2. The relationship between the variables in iteration r of the basic protocol.

We describe the two instances of the key agreement sub-protocol together as follows. Alice and Bob send $\mathbf{X}_f^{:2r-1} = \mathbf{x}_{f,U_{1A}^{:2r-2}}$ and $\mathbf{X}_b^{:2r-1} = \mathbf{x}_{b,U_{1B}^{:2r-2}}$, and receive $\mathbf{Y}_b^{:2r-1}$ and $\mathbf{Y}_f^{:2r-1}$, respectively. Eve also receives $\mathbf{Z}_f^{:2r-1}$ and $\mathbf{Z}_b^{:2r-1}$.

Alice finds (I_A, J_A) such that $\mathbf{Y}_b^{:2r-1} = d_{b,I_A}^{J_A}$, i.e., the I_A-th codeword in the J_A-th secure block code over the inverse backward DMBC; similarly, Bob obtains (I_B, J_B) such that $\mathbf{Y}_f^{:2r-1} = d_{f,I_B}^{J_B}$. Round $2r-1$ may also be interpreted as follows. Alice and Bob have encoded $I_A \in [M_{1A}]$ and $I_B \in [M_{1B}]$ to the codewords $d_{b,I_A}^{J_A}$ and $d_{b,I_B}^{J_B}$; they have sent them over the inverse DMBCs but have not included the information about which block code they belong to. Thus, round

$2r$ is primarily used for sending the block code labels, i.e., $J_A \in [N_A]$ and $J_B \in [N_B]$. The round is also used to send the pieces of randomness, $U_{2A}^{:2r-2} \in [L_{2A}]$ and $U_{2B}^{:2r-2} \in [L_{2B}]$, as well as the deterministic sequences, $\mathbf{a}$ and $\mathbf{b}$.

In the beginning of round $2r$, Alice and Bob respectively calculate $Q_A \in [M_{21A}]$ and $Q_B \in [M_{21B}]$ as (note that $M_{21A} = N_A.L_{2A}$ and $M_{21B} = N_B.L_{2B}$)

$$Q_A = L_{2A} J_A + U_{2A}^{:2r-2}, \quad \text{and} \quad Q_B = L_{2B} J_B + U_{2B}^{:2r-2}. \tag{32}$$

They next use the key derivation functions (in the secure block code) to calculate key parts $S_A^{:2r} = \phi_{s,A}(Q_A)$ and $S_B^{:2r} = \phi_{s,B}(Q_B)$. In this round, Alice and Bob send the n_2-sequences $\mathbf{X}_f^{:2r} = (c_{f,Q_A}||a)$ and $\mathbf{X}_b^{:2r} = (c_{b,Q_B}||b)$ and receive $\mathbf{Y}_b^{:2r} = (\mathbf{Y}_{1b}||\mathbf{Y}_{2b})$ and $\mathbf{Y}_f^{:2r} = (\mathbf{Y}_{1f}||\mathbf{Y}_{2f})$, respectively. Eve also receives $\mathbf{Z}_f^{:2r}$ and $\mathbf{Z}_b^{:2r}$. Using the secure block code for the forward DMBC, Bob obtains $\hat{Q}_A$ such that $\mathbf{Y}_{1f} \in \mathcal{C}_{f,\hat{Q}_A}$ and calculates $\hat{S}_A^{:2r} = \phi_{s,A}(\hat{Q}_A)$; similarly, Alice obtains $\hat{Q}_B$ such that $\mathbf{Y}_{1b} \in \mathcal{C}_{b,\hat{Q}_B}$ and calculates $\hat{S}_B^{:2r} = \phi_{s,B}(\hat{Q}_B)$. To produce randomness for the next iteration, Alice and Bob use their secure equipartitions to calculate $U_A^{:2r} = (\psi_A^{\hat{Q}_B}(\mathbf{Y}_{1b})||\psi_A(\mathbf{Y}_{2b}))$ and $U_B^{:2r} = (\psi_B^{\hat{Q}_A}(\mathbf{Y}_{1f})||\psi_B(\mathbf{Y}_{2f}))$, respectively. The randomness pieces are then split into $(U_{1A}^{:2r}, U_{2B}^{:2r}) = u_{A,split}(U_A^{:2r})$ and $(U_{1B}^{:2r}, U_{2B}^{:2r}) = u_{B,split}(U_B^{:2r})$.

The above calculations are to derive independent randomness and secret key parts from round $2r$. The following is for deriving a key part out of round $2r-1$. Firstly, the parties calculate

$$\hat{U}_{2A}^{:2r-2} = \hat{Q}_A \mod (L_{2A}), \; \hat{J}_A = (\hat{Q}_A - \hat{U}_{2A}^{:2r-2})/L_{2A}, \tag{33}$$

$$\hat{U}_{2B}^{:2r-2} = \hat{Q}_B \mod (L_{2B}), \; \hat{J}_B = (\hat{Q}_B - \hat{U}_{2B}^{:2r-2})/L_{2B}. \tag{34}$$

The quantities $\hat{J}_A \in [N_A]$ and $\hat{J}_B \in [N_B]$ are used to find which secure block codes need to be considered over the inverse DMBCs in round $2r-1$. More precisely, Alice finds $\hat{I}_B$ such that $\mathbf{X}_f^{:2r-1} \in \mathcal{D}_{f,\hat{I}_B}^{\hat{J}_B}$ and Bob finds $\hat{I}_A$ such that $\mathbf{X}_b^{:2r-1} \in \mathcal{D}_{b,\hat{I}_A}^{\hat{J}_A}$. As for the establishment of the secret key part, Alice calculates $S_A^{:2r-1} = \phi_{s,A}^{J_A}(d_{b,I_A}^{J_A})$ and $\hat{S}_B^{:2r-1} = \phi_{s,B}^{\hat{J}_B}(d_{f,\hat{I}_B}^{\hat{J}_B})$, and Bob calculates $\hat{S}_A^{:2r-1} = \phi_{s,A}^{\hat{J}_A}(d_{b,\hat{I}_A}^{\hat{J}_A})$ and $S_B^{:2r-1} = \phi_{s,B}^{J_B}(d_{f,I_B}^{J_B})$.

The total secret key part in iteration r is $(S_A^{:2r-1}, S_A^{:2r}, S_B^{:2r-1}, S_B^{:2r})$. Overall, the main protocol uses the 2DMBC $n = (2t+1)(n_1+n_2)$ times to establish $S = (S_A^{:r}, S_B^{:r})_{r=1}^{2t}$. By following this protocol, Alice calculates $S_A = (S_A^{:r}, \hat{S}_B^{:r})_{r=1}^{2t}$ and Bob calculates $S_B = (\hat{S}_A^{:r}, S_B^{:r})_{r=1}^{2t}$. In Appendix A, we show that the main algorithm satisfies the three requirements given in Definition 1 and achieves the lower bound in Theorem 1. There, we globally refer to the quantities I_A, J_A, I_B, J_B, Q_A, and Q_B for iteration r by using $I_A^{:2r-1}$, $J_A^{:2r-1}$, $I_B^{:2r-1}$, $J_B^{:2r-1}$, $Q_A^{:2r-1}$, and $Q_B^{:2r-1}$, respectively.

6 The SK Capacity in the Case of Binary Symmetric Channels

Consider the case that each DMBC consists of independent BSCs with error probabilities p_1 and p_2, i.e., the special case discussed in Section 3.2 (see Fig. 1(b)). Following the lower bound expression (13) in Theorem 1, and letting X_f and X_b to be uniform binary RVs, we conclude the following lower bound on the SK capacity in the case of BSCs, C_{sk}^{BSC}.

$$C_{sk}^{BSC} \geq 2\max_{\mu\geq 0}\{Lbound\}, \text{ such that} \tag{35}$$

$$Lbound = \tfrac{1}{1+\mu}\left(\mu(h(p_1+p_2-2p_1p_2) - h(p_1)) + \gamma(h(p_2) - h(p_1))_+\right), \tag{36}$$

$$\gamma = \min\{1, \tfrac{h(p_1)}{1-h(p_1)} - \mu\}, \tag{37}$$

$$\mu \leq \min\{h(p_1), \tfrac{1-h(p_1)}{h(p_1)}\}. \tag{38}$$

In general, $\mu \geq 0$ is a non-negative real number. In the following, we see that only three selections of μ that is $\mu \in \{0, M_1, M_2\}$ (with M_1 and M_2 defined in (39)) can lead to the lower bound (35). This makes it easy to calculate the lower bound. By letting

$$M_1 = \frac{h(p_1)}{1-h(p_1)} - 1 \qquad \text{and} \qquad M_2 = \min\{h(p_1), \frac{1-h(p_1)}{h(p_1)}\}, \tag{39}$$

we have $\mu \leq M_2$ as a condition and that (i) if $\mu \leq M_1$, then $\gamma = 1$; (ii) otherwise, $\gamma = \frac{h(p_1)}{1-h(p_1)} - \mu < 1$. Accordingly, we consider the following cases.

Case 1: $h(p_2) \leq h(p_1)$**.** In this case, $(h(p_2) - h(p_1))_+ = 0$ and so *Lbound* is written as

$$\frac{\mu}{\mu+1}\{h(p_1 + p_2 - 2p_1p_2) - h(p_1)\}. \tag{40}$$

This gives that, to maximize *Lbound*, the largest possible μ should be selected, i.e., $\mu = M_2$.

Case 2: $h(p_2) > h(p_1)$**.** We divide this into the following three subcases.
2.1) If $M_2 \leq M_1$, for any $\mu \leq M_2$, the inequality $\mu \leq M_1$ also holds. From (i) above, $\gamma = 1$ and *Lbound* can be expressed as the following weighted average

$$\frac{\mu}{1+\mu}\{(h(p_1 + p_2 - 2p_1p_2) - h(p_1))\} + \frac{1}{1+\mu}\{h(p_2) - h(p_1)\}. \tag{41}$$

Since the first term in the above average is greater or equal to the second term, the average is maximized by selecting the largest possible value for μ that is $\mu = M_2$.
2.2) If $M_2 > M_1 \geq 0$, then we may choose $\mu \leq M_1$ or $M_1 < \mu \leq M_2$.

- For $\mu \leq M_1$, from (i), $\gamma = 1$ and so *Lbound* is expressed the same way as (41). This implies that selecting $\mu = M_1$ (the largest possible value) leads to the maximization of the average.

- For $M_1 < \mu \leq M_2$, from (ii), *Lbound* can be written as the following weighted average

$$\frac{\mu}{\mu+1}\{h(p_1 + p_2 - 2p_1p_2) - h(p_2)\} + \frac{1}{\mu+1}\{\frac{h(p_1)}{1-h(p_1)}(h(p_2) - h(p_1))\}. \tag{42}$$

Depending on the relationship between the first and the second terms of the above average, the maximum is achieved by selecting either the smallest or the largest possible μ in the range $M_1 \leq \mu \leq M_2$, that is either M_1 or M_2, respectively.
2.3) If $M_1 < 0$, then for any $0 \leq \mu \leq M_2$, we have $\mu > M_1$. From (ii), *Lbound* is written the same as (42). However, the smallest and the largest values of μ are 0 and M_2, respectively.

In all cases above, either selection of $\mu \in \{0, M_1, M_2\}$ leads to the maximum achievable rate, i.e. the lower bound in (35) is simplified to

$$C_{sk}^{BSC} \geq 2 \max_{\mu \in \{0, M_1, M_2\}} \{Lbound\}. \tag{43}$$

Following the upper bound (22) in Theorem 2 for the above setting, we arrive at

$$C_{sk}^{BSC} \leq 2 \max\nolimits_{P_{X_f}, P_{X_b}} \{Ubound_A, Ubound_B\}, \text{ where} \tag{44}$$

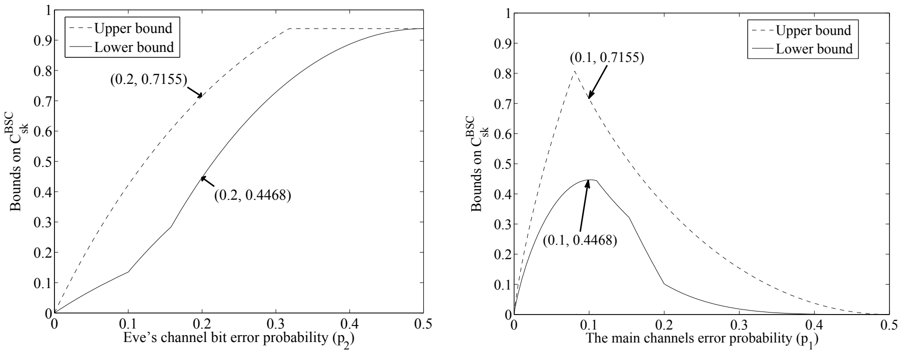
$$Ubound_A = \min\{h(p_1), H(Y_f|Z_f) - h(p_1)\}, \quad \text{and} \quad Ubound_B = \min\{h(p_1), H(Y_b|Z_b) - h(p_1)\}. \tag{45}$$

It is easy to show that, by selecting X_f and X_b to be uniformly random, $Ubound_A$ and $Ubound_B$ reach their highest values, respectively. So, the upper bound can be simplified as

$$C_{sk}^{BSC} \leq 2 \min\{h(p_1), h(p_1 + p_2 - 2p_1p_2) - h(p_1)\}. \tag{46}$$

In Fig. 3, we graph the lower and the upper bounds, in (43) and (46), for different values of the main channels error probability p_1 and Eve's channels error probability p_2. Fig. 3(a) illustrates the changes in the two bounds with respect to $0 \leq p_2 \leq 0.5$ when $p_1 = 0.1$ is fixed. According to this graph, the two bounds coincide when $p_2 = 0$ or when $p_2 = 0.5$. When $p_2 = 0$ all information sent over the 2DMBC is seen by Eve and SKE is impossible; so, both bounds equal zero. When $p_2 = .5$, the setup does not leak any information to Eve and using Theorem 3, the two bounds are expected to coincide. Fig. 3(b) graphs the changes of the two bounds when $0 \leq p_1 \leq 0.5$ and $p_2 = 0.2$ is fixed. This graph shows that when the main channels are noiseless ($p_1 = 0$) or completely noisy ($p_1 = 0.5$), the two bounds coincide at zero and so SKE is impossible. This is expected because in the former case, no randomness exists in the system for Alice and Bob and in the latter, there is no chance of reliable communication. The graphs also show the possibility of SKE even when both DMBCs are in favor of Eve. This can be observed in Fig. 3(a) for values of $0 < p_2 < (p_1 = 0.1)$ and in Fig. 3(b) for values of $(p_2 = 0.2) < p_1 < 0.5$.

In Section 3.2, we have provided an example of a simple and efficient SKE construction. For the values $p_1 = 0.1$ and $p_2 = 0.2$, the construction achieves the SK rate 3%. As depicted in Fig. 3, the lower and the upper bounds on the SK capacity for these values of p_1 and p_2 are about 45% and 72%, respectively. This reveals how the example construction of Section 3.2 works far from optimal achievable rates. As noted earlier, one can improve the performance of the protocol by using more suitable primitives.



(a) The bounds w.r.t p_2 for $p_1 = 0.1$ (b) The bounds w.r.t. p_1 for $p_2 = 0.2$

Fig. 3. The relationship between the two bounds on the SK capacity with respect to p_1 and p_2

7 Conclusion

This paper has raised the question of building cryptographic functionalities over noisy channels when there is no initial randomness available to the parties of a system. We focused on two-party secret key establishment (SKE) where Alice and Bob are connected by two independent noisy broadcast channels and the channels leak some information to an adversary, Eve. We formalized the problem and defined a secure SKE protocol as well as the secret key capacity in this setting, and showed some special cases of this setting where SKE is impossible. We then provided a concrete construction of SKE when channels are binary symmetric and proved the reliability and the security of our construction. We obtained lower and upper bounds on the secret key capacity and showed that they coincide when the channels leak zero information to Eve; this matches the known results in the previous work. For the case that the channels are binary symmetric, we derived the bounds and argued that there is a large gap between the rate achieved by the concrete SKE construction and the rate proved to be achievable by optimal primitives. It would be interesting to design better SKE constructions with higher SK rates. Our work also suggests the question of the possibility of other cryptographic primitives when channel noise is the only source of randomness.

References

1. Ahlswede, R., Csiszár, I.: Common randomness in information theory and cryptography. Part I: secret sharing. IEEE Transaction Information Theory, vol. 39, pp. 1121-1132 (1993)
2. Ahlswede, R., Cai, N.: Transmission, identification, and common randomness capacities for wire-tape channels with secure feedback from the decoder. Book Chapter General Theory of Information Transfer and Combinatorics, LNCS 4123, pp. 258-275 (2006)
3. Ahmadi, H., Safavi-Naini, R.: Secret key establishment over a pair of independent broadcast channels. In: International Symposium Information Theory and its Application, pp. 185-190 (2010) Full version on the arXiv preprint server, arXiv:1001.3908
4. Ahmadi, H., Safavi-Naini, R.: New results on key establishment over a pair of independent broadcast channels. In: International Symposium Information Theory and its Application, 2010. Full version on the arXiv preprint server, arXiv:1004.4334v1
5. Barros, J., Imai, H., Nascimento, A.C.A., Skludarek, S.: Bit commitment over Gaussian channels. In: IEEE International Symposium Information Theory, pp. 14371441 (2006)
6. Bennett, C.H., Brassard, G., Crepeau, C., Maurer, U.M.: Generalized privacy amplification. IEEE Transaction Information Theory, vol. 41, pp. 1915-1923 (1995)
7. Bloch, M., Barros, J., McLaughlin, S.: Practical information-theoretic commitment. In: Allerton Conference Communication, Control, and Computing, pp. 1035-1039 (2007)
8. Bloch, M., Barros, J., Rodrigues, M.R.D., McLaughlin, S.W.: Wireless information theoretic security. IEEE Transaction Information Theory, vol. 54, pp. 2515-2534 (2008)
9. Carter, L. Wegman, M.N.: Universal Classes of Hash Functions. Journal of Computer and System Sciences 18, pp. 143-154 (1979)
10. Cover, T.M., Thomas, J.A.: *Elements of information theory.* Wiley-IEEE, Edition 2 (2006)
11. Crandall, R., Pomerance, C.: *Prime numbers: a computational perspective*, Springer, Edition 2 (2005)

12. Crépeau, C., Kilian, J.: Weakening security assumptions and oblivious transfer. In: Crypto, LNCS 403, pp. 27 (1990)
13. Crépeau, C., Morozov, K., Wolf, S.: Efficient unconditional oblivious transfer from almost any noisy channel. In: Security in Communication Networks, LNCS 3352, pp. 4759 (2004)
14. Csiszár, I., Körner, J.: Broadcast channels with confidential messages. IEEE Transaction Information Theory, vol. 24, pp. 339-348 (1978)
15. Csiszár, I., Narayan P.: Common randomness and secret key generation with a helper. IEEE Transaction Information Theory, vol. 46, pp. 344-366 (2000)
16. Diffie, W., Hellman, M.E.: New directions in cryptography. IEEE Transaction Information Theory, vol. 22, pp. 64454 (1976)
17. Dodis, Y., Spencer, J.: On the (non)universality of the one-time pad. In: IEEE Annual Symposium FOCS, pp. 376388 (2002)
18. Dodis, Y., Wichs, D.: Non-malleable extractors and symmetric key cryptography from weak secrets. In: ACM Symposium Theory of Computing, pp. 601-610 (2009)
19. Elias, P.: The efficient construction of an unbiased random sequence, Annals of Mathematical Statistics, vol. 43, pp. 864-870 (1972)
20. Gallager, R.G.: *Information theory and teliable communication*, New York: Wiley (1968)
21. Kanukurthi, B., Reyzin, L.: Key agreement from close secrets over unsecured channels. In: Eurocrypt, LNCS 5479, pp. 206-223 (2009)
22. Khisti, A., Diggavi, S., Wornell, G.: Secret key generation with correlated sources and noisy channels. IEEE International Symposium Information Theory, pp. 1005-1009 (2008)
23. Maurer, U.: Secret key agreement by public discussion from common information. IEEE Transaction Information Theory, vol. 39, pp. 733-742 (1993)
24. Maurer, U., Wolf, S.: Information-theoretic key agreement: from weak to strong secrecy for free. In: Eurocrypt, LNCS 1807, pp. 351-351 (2000)
25. Maurer, U., Wolf, S.: Secret-key agreement over unauthenticated public channels - part I: definitions and a completeness. IEEE Transaction Information Theory, vol. 49, pp. 822-831 (2003)
26. Maurer, U., Wolf, S.: Secret-key agreement over unauthenticated public channels - part III: privacy amplification. IEEE Transaction Information Theory, vol. 49, pp. 839-851 (2003)
27. Nascimento A.C.A., Winter, A.: On the oblivious transfer capacity of noisy correlations. In: IEEE International Symposium Information Theory, pp. 1871-1875 (2006)
28. Renner, R., Wolf, S.: Unconditional authenticity and privacy from an arbitrarily weak secret. In: Crypto, LNCS 2729, pp. 78-95 (2003)
29. Prabhakaran, V., Eswaran, K., Ramchandran, K.: Secrecy via sources and channels - a secret key - secret message rate trade-off region. In: IEEE International Symposium Information Theory, pp. 1010-1014 (2008)
30. Shannon, C.E.: Communication theory of secrecy systems, Bell System Technical Journal, vol. 28, pp. 656-715 (1948)
31. Venkatesan, S., Anantharam, V.: The common randomness capacity of a pair of independent discrete memoryless channels. IEEE Transaction Information Theory, vol. 44, pp. 215224 (1998)
32. von Neumannm, J.: Various techniques used in connection with random digits. National Bureau of Standards Applied Math Series, vol. 12, pp. 36-38 (1951)
33. Wegman, M.N., Carter, J.L.: New hash functions and their use in authentication and set equality. Journal of Computer and System Sciences, vol. 22, pp. 265-279 (1981)
34. Wyner, A.D.: The wire-tap channel. Bell System Technical Journal, vol. 54, pp. 1355-1367 (1975)

A Analysis of the main protocol (Section 5)

In this section, we prove that the main protocol can achieve rates up to the lower bound, while satisfying the three conditions, required in Definition 1.

A.1 Reliability analysis: proving (6b)

We define the error event $\mathcal{E}rr$, which is true if at least one of the following happens.

– For an $1 \leq r \leq t$, at the end of round $2r-1$, Alice fails to find (I_A, J_A) such that $Y_b^{n_1:2r-1} = d_{b,I_A}^{J_A}$ or Bob fails to find (I_B, J_B) such that $Y_f^{n_1:2r-1} = d_{f,I_B}^{J_B}$. We refer to this event as $\mathcal{E}_r^1$, which indicates the failure in finding appropriate secure block codes over the inverse channels.

– For an $1 \leq r \leq t$, in round $2r$, Alice calculates $\hat{Q}_B^{:2r-1} \neq Q_B^{:2r-1}$ or $\hat{I}_B \neq I_B$ or Bob calculates $\hat{Q}_A^{:2r-1} \neq Q_A^{:2r-1}$ or $\hat{I}_A \neq I_A$. We refer to this event as $\mathcal{E}_r^2$, which indicates the decoding error in using the secure block codes.
– For an $0 \leq r \leq t-1$, at the end of round $2r$, Alice calculates $U_A^{:2r} = \perp$ or Bob calculates $U_B^{:2r} = \perp$. We refer to this event as $\mathcal{E}_r^3$, which shows the error in using the secure equipartitions.

The probability of each of the above events can be made arbitrarily small, thanks to the properties of the secure block codes and the secure equipartitions used in the protocol. In precise, we have the following upper bounds on the error event probabilities for each iteration $1 \leq r \leq t$ of the basic protocol, assuming that no error occurs in round 0 and all iterations up to $r-1$. Regarding the two sets of secure block codes for the inverse forward and backward channels, we have $\Pr(\mathcal{E}_r^1) \leq 2e^{-\gamma}$. The error event $\mathcal{E}_r^2$ corresponds to four decoding functions of the secure block codes of Alice and Bob over the channels, which implies $\Pr(\mathcal{E}_r^1) \leq 4\epsilon$. Finally, the secure partitions used by Alice and Bob give $\Pr(\mathcal{E}_r^3) \leq 2\epsilon$. The total error probability is calculated as follows. Let $\mathcal{E}_0^1$, $\mathcal{E}_0^2$, and $\mathcal{E}_t^3$ be always false.

$$\begin{aligned}
\Pr(\mathcal{E}rr) &= \Pr\left(\bigcup_{r=0}^{t}(\mathcal{E}_r^1 \cup \mathcal{E}_r^2 \cup \mathcal{E}_r^3)\right) \\
&= \Pr\left(\mathcal{E}_0^1 \cup \mathcal{E}_0^2 \cup \mathcal{E}_0^3 \cup \bigcup_{r=1}^{t}\left[(\mathcal{E}_r^1 \cup \mathcal{E}_r^2 \cup \mathcal{E}_r^3) \cap \bigcap_{i=0}^{r-1}\overline{(\mathcal{E}_r^1 \cup \mathcal{E}_r^2 \cup \mathcal{E}_r^3)}\right]\right) \\
&= \Pr\left(\mathcal{E}_0^3\right) + \sum_{r=1}^{t}\Pr\left((\mathcal{E}_r^1 \cup \mathcal{E}_r^2 \cup \mathcal{E}_r^3) \cap \bigcap_{i=0}^{r-1}(\overline{\mathcal{E}_r^1} \cap \overline{\mathcal{E}_r^2} \cap \overline{\mathcal{E}_r^3})\right) \\
&\leq \Pr\left(\mathcal{E}_0^3\right) + \sum_{r=1}^{t}\Pr\left((\mathcal{E}_r^1 \cup \mathcal{E}_r^2 \cup \mathcal{E}_r^3) | \bigcap_{i=0}^{r-1}(\overline{\mathcal{E}_r^1} \cap \overline{\mathcal{E}_r^2} \cap \overline{\mathcal{E}_r^3})\right) \\
&\leq \Pr\left(\mathcal{E}_0^3\right) + \sum_{r=1}^{t}\Pr\left(\mathcal{E}_r^1 | \bigcap_{i=0}^{r-1}(\overline{\mathcal{E}_r^1} \cap \overline{\mathcal{E}_r^2} \cap \overline{\mathcal{E}_r^3})\right) + \Pr\left(\mathcal{E}_r^2 | \bigcap_{i=0}^{r-1}(\overline{\mathcal{E}_r^1} \cap \overline{\mathcal{E}_r^2} \cap \overline{\mathcal{E}_r^3})\right) + \sum_{r=1}^{t-1}\Pr\left(\mathcal{E}_r^3 | \bigcap_{i=0}^{r-1}(\overline{\mathcal{E}_r^1} \cap \overline{\mathcal{E}_r^2} \cap \overline{\mathcal{E}_r^3})\right) \\
&\leq 2\epsilon + 2te^{-\gamma} + 4t\epsilon + 2(t-1)\epsilon \leq 6t\epsilon + 2te^{-2^{n_1\alpha/2}} \leq 7t\epsilon.
\end{aligned} \tag{47}$$

By selecting t to be polynomially increasing with $\min\{n_1, n_{21A}, n_{21B}\}$, $t\epsilon$ approaches zero for large enough n_1, n_{21A}, n_{2A}. This proves that for any arbitrarily $\delta > 0$ and sufficiently small α, we can find n_1 and n_2 such that $7t\epsilon < \delta$ and so

$$\Pr(S_A = S_B = S) \geq 1 - \Pr(\mathcal{E}rr) \geq 1 - 7t\epsilon > 1 - \delta. \tag{48}$$

A.2 Randomness analysis: proving (6a)

The entropy of the secret key S can be bounded from below as

$$H(S) \geq \Pr(\overline{\mathcal{E}rr})H(S|\overline{\mathcal{E}rr}) \geq (1 - 7t\epsilon)H(S|\overline{\mathcal{E}rr}). \tag{49}$$

We hereafter assume that no error has occurred and calculate the entropy of S based on this assumption. Assuming no error implies that each secure equipartition function gives an independent and uniform RV in the domain. In other words, ψ_A^j, ψ_A, ψ_B^j, and ψ_B return independent and uniform RVs in $[\mathit{\Gamma}_{21A}]$, $[\mathit{\Gamma}_{22A}]$, $[\mathit{\Gamma}_{21B}]$, and $[\mathit{\Gamma}_{22B}]$, respectively. So, for each iteration $1 \leq r \leq t$, the variables $U_A^{:2r-2}$ and $U_B^{:2r-2}$ are uniformly distributed and independent of the variables in any round less or equal to round $2r-2$. Since each execution of the basic protocol runs two key agreement procedures independently in parallel, the variables of these procedure are also independent. This implies that, for all $(i_A^{:2r-1}, q_A^{:2r-1}, i_B^{:2r-1}, q_B^{:2r-1})_{r=1}^t$ in $([M_{1A}] \times [M_{21A}] \times [M_{1B}] \times [M_{21B}])^t$,

$$\begin{aligned}
&\Pr\left(\bigcap_{r=1}^{t}(I_A^{:2r-1}, Q_A^{:2r-1}, I_B^{:2r-1}, Q_B^{:2r-1}) = (i_A^{:2r-1}, q_A^{:2r-1}, i_B^{:2r-1}, q_B^{:2r-1})\right) \\
&\qquad = \prod_{r=1}^{t}\Pr\left((I_A^{:2r-1}, Q_A^{:2r-1}, I_B^{:2r-1}, Q_B^{:2r-1}) = (i_A^{:2r-1}, q_A^{:2r-1}, i_B^{:2r-1}, q_B^{:2r-1})\right) \\
&\qquad = \prod_{r=1}^{t}\Pr\left((I_A^{:2r-1}, Q_A^{:2r-1}) = (i_A^{:2r-1}, q_A^{:2r-1})\right).\Pr\left((I_B^{:2r-1}, Q_B^{:2r-1}) = (i_B^{:2r-1}, q_B^{:2r-1})\right).
\end{aligned} \tag{50}$$

and hence, for all $(s_A^{:2r-1}, s_A^{:2r}, s_B^{:2r-1}, s_B^{:2r})_{r=1}^t$ in $([K_{1A}] \times [K_{21A}] \times [K_{1B}] \times [K_{21B}])^t$,

$$\Pr\left(\bigcap_{r=1}^{t}(S_A^{:2r-1}, S_A^{:2r}, S_B^{:2r-1}, S_B^{:2r}) = (s_A^{:2r-1}, s_A^{:2r}, s_B^{:2r-1}, s_B^{:2r})\right)$$
$$= \prod_{r=1}^{t} \Pr\left((S_A^{:2r-1}, S_A^{:2r}) = (s_A^{:2r-1}, s_A^{:2r})\right) . \Pr\left((S_B^{:2r-1}, S_B^{:2r}) = (s_B^{:2r-1}, s_B^{:2r})\right). \quad (51)$$

This leads to

$$H(S) = H\left((S_A^{:2r-1}, S_A^{:2r}, S_B^{:2r-1}, S_B^{:2r})_{r=1}^t\right) = \sum_{r=1}^{t} H(S_A^{:2r-1}, S_A^{:2r}) + H(S_B^{:2r-1}, S_B^{:2r}). \quad (52)$$

To continue the calculation above, we first discuss the RVs $I_A^{:2r-1}$, $J_A^{:2r-1}$, and $U_{2A}^{:2r-2}$. For all $i \in [M_{1A}]$ and all $j \in [N_A]$, we have

$$\Pr\left((I_A^{:2r-1}, J_A^{:2r-1}) = (i,j)\right) \leq \Pr(Y_b^{n_1:2r-1} = d_{b,i}^j) \leq 2^{-n_1(H(Y_b)-\epsilon)}, \quad (53)$$

where the last inequality follows from AEP and that $d_{b,i}^j$ is ϵ-typical w.r.t. Y_b. Since $U_A^{:2r-2} \in [\Gamma_{21A}] \times [\Gamma_{21B}]$ has a uniform distribution in, the two parts of it $U_{1A}^{:2r-2} \in [L_{1A}]$ and $U_{2A}^{:2r-2} \in [L_{2A}]$ are also uniformly distributed, i.e., specifically for $U_{2A}^{:2r-2}$,

$$\forall u \in [L_{2A}]: \quad \Pr\left(U_{2A}^{:2r-2} = u\right) = \frac{1}{L_{2A}}. \quad (54)$$

We conclude that, for all $i \in [M_{1A}]$ and all $q \in [M_{21A}]$, letting $u = q \mod (L_{2A})$ and $j = (q-u)/L_{2A}$, we have (see (26), (31), and (32))

$$\begin{aligned}
\Pr\left((I_A^{:2r-1}, Q_A^{:2r-1}) = (i,q)\right) &= \Pr\left((I_A^{:2r-1}, J_A^{:2r-1}, U_{2A}^{:2r-2}) = (i,j,u)\right) \\
&\overset{(a)}{=} \Pr\left((I_A^{:2r-1}, J_A^{:2r-1}) = (i,j)\right) . \Pr\left(U_{2A}^{:2r-2} = u\right) \\
&= \frac{1}{L_{2A}} \Pr\left((I_A^{:2r-1}, J_A^{:2r-1}) = (i,j)\right) \\
&\leq \frac{1}{L_{2A}} 2^{-n_1(H(Y_b)-\epsilon)} = 2^{-n_1 I(X_b;Y_b) - n_{21A} I(X_f;Y_f) + n_1\epsilon} \\
&= \frac{2^{n_1\epsilon + n_{21A}\alpha}}{M_{1A} M_{21A}}.
\end{aligned} \quad (55)$$

Equality (a) holds since $(I_A^{:2r-1}, J_A^{:2r-1})$ and $U_{2A}^{:2r-2}$ are independent. The continuity of the entropy function gives

$$H(I_A^{:2r-1}, Q_A^{:2r-1}) \geq \log(M_{1A} M_{21A}) - n_{21A}\alpha - n_1\epsilon. \quad (56)$$

From the property of functions ϕ and ϕ^j (see Definition 5) and the description of the protocol, we can write

$$H(S_A^{:2r-1}, S_A^{:2r}) \geq \log(K_{1A} K_{21A}) - n_{21A}\alpha - n_1\epsilon = n_1 R_{scb^{-1}} + n_{21A} R_{scf} - n_{21A}\alpha - n_1\epsilon. \quad (57)$$

One can follow a similar approach to above to show

$$H(S_B^{:2r-1}, S_B^{:2r}) \geq \log(K_{1B} + K_{21B}) - n_1\epsilon = n_1 R_{scf^{-1}} + n_{21B} R_{scb} - n_{21B}\alpha - n_1\epsilon. \quad (58)$$

Using (52) and (57) in (58), we can write

$$\begin{aligned}
\frac{H(S)}{n} &= \frac{t(n_1 R_{scb^{-1}} + n_{21A} R_{scf} - n_{21A}\alpha - n_1\epsilon) + t(n_1 R_{scf^{-1}} + n_{21B} R_{scb} - n_{21B}\alpha - n_1\epsilon)}{(t+1)(n_1+n_2)} \\
&\geq \frac{t}{(t+1)(\mu+1)} \left(\mu R_{scb^{-1}} + \frac{n_{21A}}{n_2} R_{scf} + \mu R_{scf^{-1}} + \frac{n_{21B}}{n_2} R_{scb} - 2\alpha - 2\mu\epsilon\right) \\
&\geq \frac{t}{(t+1)(1+\mu)} \left((\mu R_{scb^{-1}} + \gamma_1 R_{scf}) + (\mu R_{scf^{-1}} + \gamma_2 R_{scb}) - 2(1+\mu)\alpha\right)
\end{aligned} \quad (59)$$

where $\mu = \frac{n_1}{n_2}$ and γ_1 and γ_2 are as defined in the theorem. Thus, for an arbitrarily given $\delta > 0$, we can choose α, t, n_1, and n_2 such that

$$\frac{H(S)}{n} > Lbound_A + Lbound_B - \delta, \quad (60)$$

with $Lbound_A$ and $Lbound_B$ as defined in the theorem.

A.3 Secrecy analysis: proving (6c)

Denote by $V_E^{:r}$ and $SK^{:r}$ Eve's view and the total secret key established at the end of round r, respectively.

$$\begin{aligned}
H(S|V_E^{:2t}) &= H(S) - I\left(SK^{:2t}; V_E^{:2t}\right) \\
&= H(S) - I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; V_E^{:2t}\right) - I\left(SK^{:2t-2}; V_E^{:2t}|(S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}\right) \\
&= H(S) - I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; V_E^{:2t}\right) - I\left(SK^{:2t-2}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}|(S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}\right) \\
&\quad - I\left(SK^{:2t-2}; V_E^{:2t-2}|(S_A^{:r}, S_B^{:r}, \mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) \\
&\geq H(S) - I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; V_E^{:2t}\right) - I\left(SK^{:2t-2}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}|(S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}\right) \\
&\quad - I\left(U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2}|SK^{:2t-2}\right) - I\left(SK^{:2t-2}; V_E^{:2t-2}\right), \qquad (61)
\end{aligned}$$

where the last inequality holds since

$$\begin{aligned}
I\left(SK^{:2t-2}; V_E^{:2t-2}|(S_A^{:r}, S_B^{:r}, \mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) &\leq I\left(SK^{:2t-2}, U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2}|(S_A^{:r}, S_B^{:r}, \mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) \\
&\overset{(a)}{\leq} I\left(SK^{:2t-2}, U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2}\right) \\
&= I\left(U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2}|SK^{:2t-2}\right) + I\left(SK^{:2t-2}; V_E^{:2t-2}\right). \qquad (62)
\end{aligned}$$

Inequality (a) is due to the Markov chain $(SK^{:2t-2}, V_E^{:2t-2}) \leftrightarrow (U_A^{:2t-2}, U_B^{:2t-2}) \leftrightarrow (S_A^{:r}, S_B^{:r}, \mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}$. There are 5 terms on the right hand of (61). In the sequel, we calculate the second, the third, and the fourth terms separately and show that they are all arbitrarily small.

The second term in (61)

$$\begin{aligned}
&I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; V_E^{:2t}\right) \\
&= I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) + I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; V_E^{:2t-2}|(\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) \\
&\leq I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) + I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}, U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2}, \mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}|(\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) \\
&\overset{(a)}{\leq} I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) + I\left(U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2}, \mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}\right) \\
&\overset{(b)}{=} I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) + I\left(U_A^{:2t-2}, U_B^{:2t-2}; \mathbf{Z}_f^{:2t-2}, \mathbf{Z}_b^{:2t-2}, \mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}\right) \\
&\overset{(c)}{\leq} I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}\right) + (\log \Gamma_A + \log \Gamma_B)\epsilon \\
&\overset{(d)}{=} I\left(S_A^{:2t-1}, S_A^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}\right) + I\left(S_B^{:2t-1}, S_B^{:2t}; \mathbf{Z}_f^{:2t-1}, \mathbf{Z}_b^{:2t}\right) + \log(\Gamma_A \Gamma_B)\epsilon. \qquad (63)
\end{aligned}$$

Inequality (a) is due to the Markov chain

$$(\mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}, V_E^{:2t-2}) \leftrightarrow (U_A^{:2t-2}, U_B^{:2t-2}) \leftrightarrow (S_A^{:r}, S_B^{:r}, \mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t},$$

equality (b) is due to

$$V_E^{:2t-2} \leftrightarrow (\mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}, \mathbf{Z}_f^{:2t-2}, \mathbf{Z}_b^{:2t-2}) \leftrightarrow (U_A^{:2t-2}, U_B^{:2t-2}),$$

inequality (c) follows from the property of secure equipartitions (see (23)), and equality (d) holds due to the independency of the variables. We shall show that the first two terms of (63) are small. Using (26) and (30)),

$$\begin{aligned}
&I\left(S_A^{:2t-1}, S_A^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}\right) \\
&= I\left(S_A^{:2t-1}, S_A^{:2t}, \mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}\right) - I\left(\mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}|S_A^{:2t-1}, S_A^{:2t}\right) \\
&= I\left(\mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}\right) - I\left(\mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}|S_A^{:2t-1}, S_A^{:2t}\right) \\
&\leq I\left(\mathbf{Y}_b^{:2t-1}; \mathbf{Z}_b^{:2t-1}\right) + I\left(\mathbf{X}_f^{:2t}; \mathbf{Z}_f^{:2t}\right) - I\left(\mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}|S_A^{:2t-1}, S_A^{:2t}\right) \\
&\leq n_1(I(Y_b; Z_b) + \epsilon) + n_{21A}(I(X_f; Z_f) + \epsilon) - H\left(\mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t}|S_A^{:2t-1}, S_A^{:2t}\right) + H\left(\mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t}|\mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}, S_A^{:2t-1}, S_A^{:2t}\right). \qquad (64)
\end{aligned}$$

The last inequality follows from AEP. Using the proof for the existence of capacity achieving codes along with Fano's inequalities gives us that the last term in the above is at most $(n_1 + n_{21A})\delta_1$ for some arbitrarily small δ_1 (see e.g.,

Appendix E or [3, Proof of Lemma 2]). For the the rest we write

$$
\begin{aligned}
& I\left(S_A^{:2t-1}, S_A^{:2t}; \mathbf{Z}_b^{:2t-1}, \mathbf{Z}_f^{:2t}\right) \\
& \leq n_1 I(Y_b; Z_b) + n_{21A} I(X_f; Z_f) - H\left(\mathbf{Y}_b^{:2t-1}, \mathbf{X}_f^{:2t} | S_A^{:2t-1}, S_A^{:2t}\right) + (n_1 + n_{21A})(\epsilon + \delta_1) \\
& = n_1 I(Y_b; Z_b) + n_{21A} I(X_f; Z_f) - H\left(I_A^{:2t-1}, Q_A^{:2t-1} | S_A^{:2t-1}, S_A^{:2t}\right) + (n_1 + n_{21A})(\epsilon + \delta_1) \\
& = n_1 I(Y_b; Z_b) + n_{21A} I(X_f; Z_f) - H(I_A^{:2t-1}, Q_A^{:2t-1}) + H(S_A^{:2t-1}) + H(S_A^{:2t}) + (n_1 + n_{21A})(\epsilon + \delta_1) \\
& \leq n_1 I(Y_b; Z_b) + n_{21A} I(X_f; Z_f) - \log(M_1 M_{21A}) + \log(K_1 K_{21A}) + (n_1 + n_{21A})(\epsilon + \delta_1) \\
& = n_1 I(Y_b; Z_b) + n_{21A} I(X_f; Z_f) + n_1(R_{scb^{-1}} - R_{cb}) + n_{21A}(R_{scf} - R_{cf}) + (n_1 + n_{21A})(\epsilon + \delta_1) \\
& = n_1 I(Y_b; Z_b) + n_{21A} I(X_f; Z_f) - n_1(I(Y_b; Z_b) + \alpha) - n_{21A}(I(X_f; Z_f) + \alpha) + (n_1 + n_{21A})(\epsilon + \delta_1) \\
& = (n_1 + n_{21A})(\epsilon + \delta_1 - \alpha) \leq (n_1 + n_{21A})\delta_2,
\end{aligned}
\tag{65}
$$

for an arbitrarily small δ_2. Similarly, one can show

$$
I\left(S_B^{:2t-1}, S_B^{:2t}; \mathbf{Z}_f^{:2t-1}, \mathbf{Z}_b^{:2t}\right) \leq (n_1 + n_{21B})\delta_3,
\tag{66}
$$

for an arbitrarily small δ_3. This gives that (63) is bounded as

$$
I\left((S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}; V_E^{:2t}\right) \leq (n_1 + n_2)\delta_4,
\tag{67}
$$

for some arbitrarily small δ_4.

The third term in (61)

$$
\begin{aligned}
& I\left(SK^{:2t-2}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t} | (S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}\right) \\
& \leq I\left(SK^{:2t-2}, \mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}; (\mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t}, U_A^{:2t-2}, U_B^{:2t-2} | (S_A^{:r}, S_B^{:r})_{r=2t-1}^{2t}\right) \\
& \overset{(a)}{\leq} I\left(\mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}; U_A^{:2t-2}, U_B^{:2t-2}\right) \\
& \overset{(b)}{=} I\left(\mathbf{X}_f^{:2t-2}; U_B^{:2t-2}\right) + I\left(\mathbf{X}_b^{:2t-2}; U_A^{:2t-2}\right) \\
& \leq \log(\Gamma_A . \Gamma_B)\epsilon.
\end{aligned}
\tag{68}
$$

Inequality (a) is due to the Markov chain

$$
SK^{:2t-2} \leftrightarrow (\mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}) \leftrightarrow (U_A^{:2t-2}, U_B^{:2t-2}) \leftrightarrow (S_A^{:r}, S_B^{:r}, \mathbf{Z}_f^{:r}, \mathbf{Z}_b^{:r})_{r=2t-1}^{2t},
$$

and equality (b) follows from the independence of the variables.

The fourth term in (61)

$$
\begin{aligned}
I\left(U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2} | SK^{:2t-2}\right) & \leq I\left(U_A^{:2t-2}, U_B^{:2t-2}; V_E^{:2t-2}, \mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2} | SK^{:2t-2}\right) \\
& \overset{(a)}{\leq} I\left(U_A^{:2t-2}, U_B^{:2t-2}; \mathbf{Z}_f^{:2t-2}, \mathbf{Z}_b^{:2t-2}, \mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}\right) \\
& \overset{(b)}{=} I\left(U_A^{:2t-2}; \mathbf{Z}_b^{:2t-2}, \mathbf{X}_b^{:2t-2}\right) + I\left(U_B^{:2t-2}; \mathbf{Z}_f^{:2t-2}, \mathbf{X}_f^{:2t-2}\right) \\
& \leq \log(\Gamma_A . \Gamma_B)\epsilon.
\end{aligned}
\tag{69}
$$

Inequality (a) is due to the Markov chain

$$
(SK^{:2t-2}, V_E^{:2t-2}) \leftrightarrow (\mathbf{X}_f^{:2t-2}, \mathbf{X}_b^{:2t-2}, \mathbf{Z}_f^{:2t-2}, \mathbf{Z}_b^{:2t-2}) \leftrightarrow (U_A^{:2t-2}, U_B^{:2t-2}),
$$

and equality (b) follows from the independence of the variables.
Using (67)-(69) in (61), we arrive at

$$
H(S|V_E^{:2t}) \geq H(S) - I\left(SK^{:t-2}; V_E^{:t-2}\right) - (n_1 + n_2)\delta_5,
\tag{70}
$$

for some arbitrarily small δ_5. Repeating the above steps t times, lets us conclude

$$H(S|V_E^{:2t}) \geq H(S) - t(n_1+n_2)\delta_5, \tag{71}$$

which proves, for appropriate selection of parameters,

$$\frac{H(S|V_E^{:2t})}{H(S)} \geq 1 - \frac{t(n_1+n_2)\delta_5}{H(S)} > 1-\delta. \tag{72}$$

B Proof of Theorem 2: upper bound

Let Π be an (R_{sk},δ)-secure t-round protocol that achieves the SK rate R_{sk} as described in Section 2 for an arbitrarily small $\delta > 0$. Using (6c) and Fano's inequality for (6b), we have

$$I(S;V_E^{:t-1}) = H(S) - H(S|View_E) \leq \delta H(S), \quad H(S|S_A) \leq h(\delta) + \delta H(S), \quad H(S|S_B) \leq h(\delta)+\delta H(S) \tag{73}$$

Considering (73), we write the entropy of S as

$$\begin{aligned} H(S) &= I(S;S_A) + H(S|S_A) + I(S;V_E^{:t-1}) - I(S;V_E^{:t-1}) \\ &\leq I(S;S_A|V_E^{:t}) + H(S|S_A) + I(S;V_E^{:t-1}) \qquad (74) \\ &\leq H(S_A|V_E^{:t-1}) + h(\delta) + 2\delta H(S) \\ &\leq H(V_A^{:t-1}|V_E^{:t-1}) + h(\delta) + 2\delta H(S). \qquad (75) \end{aligned}$$

Similarly

$$H(S) \leq H(V_B^{:t-1}|V_E^{:t-1}) + h(\delta) + 2\delta H(S), \tag{76}$$

$$H(S) \leq H(V_A^{:t-1}, V_B^{:t-1}|V_E^{:t-1}) + h(\delta) + 2\delta H(S), \tag{77}$$

and, from (74),

$$\begin{aligned} H(S) &\leq I(S,S_B;S_A|V_E^{:t-1}) + H(S|S_A) + I(S;V_E^{:t-1}) \\ &= I(S_B;S_A|V_E^{:t-1}) + I(S;S_A|S_B,V_E^{:t-1}) + H(S|S_A) + I(S;V_E^{:t-1}) \\ &\leq I(S_A;S_B|V_E^{:t-1}) + H(S|S_B) + H(S|S_A) + I(S;V_E^{:t-1}) \\ &\leq I(V_A^{:t-1};V_B^{:t-1}|V_E^{:t-1}) + 2h(\delta) + 3\delta H(S). \qquad (78) \end{aligned}$$

Choose the RVs (X_f,Y_f,Z_f) and (X_b,Y_b,Z_b) such that they correspond to the 2DMBC probability distributions and

$$P_{X_f} = \frac{1}{n}\sum_{r=0}^{t-1}\sum_{i=1}^{n_r} P_{X_{f,i}^{:r}}, \qquad P_{X_b} = \frac{1}{n}\sum_{r=0}^{t-1}\sum_{i=1}^{n_r} P_{X_{b,i}^{:r}},$$

Below, we study each of the inequalities (75)-(78), respectively, to obtain four upper bounds on the entropy of the key S produced by the SKE protocol Π.

$$\begin{aligned} H(V_A^{:t-1}|V_E^{:t-1}) &\leq H(V_A^{:t-1}|V_E^{:t-1}) \\ &= \sum_{r=0}^{t-1} H(Y_b^{n_r:r}|V_A^{:r-1}, V_E^{:t-1}) \\ &\leq \sum_{r=0}^{t-1} H(Y_b^{n_r:r}|Z_b^{n_r:r}) \\ &\leq nH(Y_b|Z_b). \qquad (79) \end{aligned}$$

Similarly

$$H(V_A^{:t-1}|V_E^{:t-1}) \leq nH(Y_f|Z_f). \tag{80}$$

$$
\begin{aligned}
H(V_A^{:t-1}, V_B^{:t-1}|V_E^{:t-1}) &= \sum_{r=0}^{t-1} H(Y_f^{n_r:r}, Y_b^{n_r:r}|V_A^{:r-1}, V_B^{:r-1}, V_E^{:t-1}) \\
&= \sum_{r=0}^{t-1} H(Y_f^{n_r:r}, Y_b^{n_r:r}|V_A^{:r-1}, V_B^{:r-1}, X_f^{n_r:r}, X_b^{n_r:r}, V_E^{:t-1}) \\
&= \sum_{r=0}^{t-1} (H(Y_f^{n_r:r}|X_f^{n_r:r}, Z_f^{n_r:r}, Z_b^{n_r:r}) + H(Y_b^{n_r:r}|X_b^{n_r:r}, Z_f^{n_r:r}, Z_b^{n_r:r})) \\
&\leq n(H(Y_f|X_f, Z_f) + H(Y_b|X_b, Z_b)). \qquad (81)
\end{aligned}
$$

Finally,

$$
\begin{aligned}
I(V_A^{:t-1}; V_B^{:t-1}|V_E^{:t-1}) &= H(V_A^{:t-1}|V_E^{:t-1}) + H(V_B^{:t-1}|V_E^{:t-1}) - H(V_A^{:t-1}, V_B^{:t-1}|V_E^{:t-1}) \\
&\leq n(I(X_f; Y_f|Z_f) + H(X_b; Y_b|Z_b)). \qquad (82)
\end{aligned}
$$

Combining the results gives

$$
\begin{aligned}
H(S) &\leq n \min\{H(Y_f|Z_f), H(Y_b|Z_b), (H(Y_f|X_f, Z_f) + H(Y_b|X_b, Z_b)), (I(X_f; Y_f|Z_f) + H(X_b; Y_b|Z_b))\} + 2h(\delta) + 3\delta H(S) \\
&= n(\min\{H(Y_f|X_f, Z_f), I(X_b; Y_b|Z_b)\} + \min\{H(Y_b|X_b, Z_b), I(X_f; Y_f|Z_f)\}) + 2h(\delta) + 3\delta H(S). \qquad (83)
\end{aligned}
$$

From (6a) and (83), we conclude the following upper bound on R_{sk}

$$
\begin{aligned}
R_{sk} &< \frac{1}{n} H(S) + \delta \\
&< \min\{H(Y_f|X_f, Z_f), I(X_b; Y_b|Z_b)\}] + \min\{H(Y_b|X_b, Z_b), I(X_f; Y_f|Z_f)\} + \delta + 2h(\delta) + 3\delta H(S) \\
&\leq \min\{H(Y_f|X_f, Z_f), I(X_b; Y_b|Z_b)\}] + \min\{H(Y_b|X_b, Z_b), I(X_f; Y_f|Z_f)\}.
\end{aligned}
$$

The last inequality holds since δ is arbitrarily small. □

C Proof of Theorem 3

When the channel leaks zero information to Eve, we have $I(X_f, Y_f; Z_f) = I(X_b, Y_f; Z_b) = 0$. Following the lower bound (13), we choose $\mu \approx 0$ and lower bound the SK capacity as follows. Note that for this selection of μ the conditions (18) and (19) always hold.

$$
C_{sk}^{2DMBC} \geq \max_{P_{X_f}, P_{X_b}} \{Lbound_A + Lbound_B\}, \qquad (84)
$$

where

$$
Lbound_A = (\gamma_1 I(X_f; Y_f)), \qquad \gamma_1 = \min\{1, \frac{H(Y_b|X_b)}{I(X_f; Y_f)}\}, \qquad (85)
$$

$$
Lbound_B = (\gamma_2 I(X_b; Y_b)), \qquad \gamma_2 = \min\{1, \frac{H(Y_f|X_f)}{I(X_b; Y_b)}\}. \qquad (86)
$$

The above can be written as

$$
Lbound_A = \min\{H(Y_b|X_b), I(X_f; Y_f)\}, \qquad Lbound_B = \min\{H(Y_f|X_f), I(X_b; Y_b)\}. \qquad (87)
$$

The first and the second term above equal $Ubound_A$ and $Ubound_B$ in the upper bound (22). □

D Proof of Lemma 4

The proof would be an extension of that of Lemma 3; thus, a complete proof is omitted. In brief, the existence of N secure block codes is proved similarly to the existence of one secure block code, following Shannon's random coding argument. Let $\{(d_i^j, \mathcal{D}_i^j)_{i=1}^M,\ 1 \le j \le N\}$ with the key derivation functions $(\phi_{s,B}^j)_{j=1}^N$ represent N codebooks whose codewords of length n are identically and independently distributed (i.i.d.) according to the probability P_X, i.e.,

$$\forall i \in [M],\ j \in [N], x^n \in \mathcal{X}^n: \ \Pr(d_i^j = x^n) = \prod_{l=1}^n P_X(x_l).$$

The proof of Lemma 3 (e.g., in [14, 34]), following the above random coding argument, shows that each random code $(d_i^j, \mathcal{D}_i^j)_{i=1}^M$ provides the security and reliability requirements of a secure block code, in expectation; hence, the existence of one instance of each code that is a secure block code. Therefore, we only need to show that a randomly selected typical X^n is in at least one of the codes with high probability.

For given ϵ, for large enough n, each of the above codewords are ϵ-typical with high probability. From AEP for P_X, a randomly selected X^n equals to a codeword in a secure block code with probability at least $2^{-n(H(X)+\epsilon)}$. There are $M.N$ i.i.d. generated codewords for all the secure block codes. So, the probability that X^n does not match any of those codewords is at most

$$\begin{aligned}(1-2^{-n(H(X)+\epsilon)})^{M.N} &= \left((1-2^{-n(H(X)+\epsilon)})^{n(H(X)+\epsilon)}\right)^{M.N-n(H(X)+\epsilon)}\\ &\le \left(e^{-1}\right)^{M.N-n(H(X)+\epsilon)}\\ &= e^{-n(R'+R_c-H(X)-\epsilon)} = e^{-\gamma}.\end{aligned}$$

E Proof of Lemma 6

We prove the lemma for $R_{se} = H(Y|XZ) - \epsilon'$. This obviously implies the existence of secure equipartition with smallest rates R_{se}. In the proof, we assume that $|\mathcal{C}| \le 2^{n(H(Y)-5\epsilon')}$.

Since $R_{se} < H(Y|XZ) \le H(Y|Z)$ and $\epsilon \ge \epsilon' = 2^{n(R_{se}-H(Y|XZ))} \ge 2^{nR_{se}-H(Y|X)}$, from Lemma 5, we already have that there exists a (M, ϵ)-equipartition with M and ϵ defined in the statement of Lemma 6 such that each part has size at most $2^{n\epsilon'}.|\mathcal{C}|/M$. It remains to prove that there exits such an equipartition with the function ψ_t that satisfies the secrecy requirement (23). The conditional entropy of T is calculated as

$$\begin{aligned}H(T|X^n=c,Z^n) &= H(Y^n,T|X^n=c,Z^n) - H(Y^n|X^n=c,Z^n,T)\\ &= H(Y^n|X^n=c,Z^n) - H(Y^n|X^n=c,Z^n,T)\\ &\overset{(a)}{\ge} n(H(Y|X,Z)-\epsilon') - H(Y^n|X^n=c,Z^n,T)\\ &\ge \log M - n\epsilon' - H(Y^n|X^n=c,Z^n,T)\\ &\overset{(b)}{\ge} \log M - n\epsilon'(I(Y;X,Z)+1) - h(\epsilon')\\ &> \log M(1-\epsilon), \end{aligned} \tag{88}$$

where

$$\epsilon = \frac{3nI(Y;X,Z)h(\epsilon')}{\log M} = \frac{3I(Y;X,Z)h(\epsilon')}{H(Y|XZ)-\epsilon'}.$$

Inequality (a) follows from joint AEP for (Y, X, Z) and inequality (b) is shown in the sequel.

Knowing T reveals the part $\mathcal{C}(T)$ which Y^n belongs to. Consider $\mathcal{C}(T)$ as a codebook of size at most $2^{n\epsilon'}|\mathcal{C}|/M$. From joint AEP [10, Chaoter 8], if $\log(2^{n\epsilon'}|\mathcal{C}|/M)$ is less than $nI(Y;X,Z)$, then there exists such a partition with the corresponding encoding and decoding functions such that the error probability of decoding (X^n, Z^n) to Y^n is arbitrarily close to zero. We calculate $\log(2^{n\epsilon'}|\mathcal{C}|/M)$ as

$$\begin{aligned}\log(2^{n\epsilon'}|\mathcal{C}|/M) &= \log(|\mathcal{C}|) - \log(M) + n\epsilon'\\ &\le n(H(Y)-5\epsilon') - nR_{se} + n\epsilon'\\ &\le n(H(Y)-5\epsilon') - n(H(Y|X,Z)-\epsilon') + n\epsilon'\\ &\le nI(Y;X,Z) - 3n\epsilon'. \end{aligned} \tag{89}$$

As a consequence, the error probability of the above decoding is less than ϵ', and Fano's inequality gives that

$$H(Y^n|X^n=c,Z^n,T) \le h(\epsilon') + n\epsilon' I(Y;X,Z). \qquad \square$$