IRWIN AND JOAN JACOBS
CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGIES

A Large Deviations Approach to Secure Lossy Compression

Nir Weinberger and Neri Merhav

CCIT Report #884
April 2015

DEPARTMENT OF ELECTRICAL ENGINEERING
TECHNION - ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA 32000, ISRAEL

CCIT Report #884 April 2015

A Large Deviations Approach to Secure Lossy Compression

Nir Weinberger and Neri Merhav
Dept. of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 3200004, Israel
{nirwein@tx, merhav@ee}.technion.ac.il

Abstract

We consider a Shannon cipher system for memoryless sources, in which distortion is allowed at the legitimate decoder. The source is compressed using a rate distortion code secured by a shared key, which satisfies a constraint on the compression rate, as well as a constraint on the exponential rate of the excess-distortion probability at the legitimate decoder. Secrecy is measured by the exponential rate of the exiguous-distortion probability at the eavesdropper, rather than by the traditional measure of equivocation. We define the perfect secrecy exponent as the maximal exiguous-distortion exponent achievable when the key rate is unlimited. Under limited key rate, we prove that the maximal achievable exiguous-distortion exponent is equal to the minimum between the average key rate and the perfect secrecy exponent, for a fairly general class of variable key rate codes.

Index Terms

Information-theoretic secrecy, Shannon cipher system, secret key, cryptography, lossy compression, rate-distortion theory, error exponent, large-deviations, covering lemmas.

I. Introduction

In his seminal paper [1], Shannon has introduced a mathematical framework for secret communication. The cipher system is considered *perfectly secure* if the cryptogram and the message are statistically independent, and so, an eavesdropper does not gain any information when he observes the cryptogram. To achieve secrecy, the sender and the legitimate recipient share a secret key, which is used to encipher and decipher the message. It is rather apparent from ordinary compression [2] that a necessary and sufficient condition for perfect secrecy is that the available key rate is larger than the information rate required to compress the source (the entropy or rate-distortion function of the source in case of lossless or lossy compression, respectively). Usually, the supply of key bits is a

This work was supported by the Israel Science Foundation (ISF), grant no. 412/12.

limited resource, as they need to be transferred to the intended recipient via a completely secure channel. When the key rate is less than the information rate, secrecy is traditionally measured in terms of *equivocation,* that is, the conditional entropy of the message given the cryptogram. The use of equivocation as a secrecy measure was advocated by other models of secrecy systems, which do not assume a shared key. Instead, secrecy is achieved by the fact that the message intercepted by the eavesdropper is of lower quality than the one received by the legitimate receiver. For example, in the ubiquitous wire-tap model [3], [4], the channel of the wiretapper is degraded (or more noisy) with respect to (w.r.t.) the channel of the legitimate receiver. In the model of [5], [6], [7] the legitimate recipient has better quality of side information than the eavesdropper.

The equivocation is indeed an unambiguous measure for statistical dependence when it is equal to either its minimal value of zero (the random variables are deterministic functions of each other), or its maximal value of the unconditional entropy (the two random variables are independent). Nonetheless, for *partial secrecy*, i.e., when the equivocation takes values strictly between these two extremes, its operational meaning is disputable. Thus, in [8], it was proposed to measure partial secrecy by the expected number of spurious messages that explain the given cryptogram (which is somewhat equivalent to the probability of correctly decrypting the message). Later, in [9], it was proposed to measure partial secrecy by the minimum average distortion that an eavesdropper can attain (this was also considered previously, to some extent, in [10]). In addition, in [9] the possibility that the legitimate recipient can tolerate a certain distortion level was also incorporated into the system model. In [9, Theorems 2 and 3], inner and outer bounds were obtained on the achievable trade-off between the coding rate, the key rate, and distortion levels at the legitimate recipient and eavesdropper. However, in [11], it was revealed that this trade-off is, in fact, degenerated. It was demonstrated there that in some cases, a negligible key rate can cause maximum distortion at the eavesdropper. The following simple example (from [12, Section I.A]) demonstrates this: Consider an memoryless source $\mathbf{X} = (X_1, \ldots, X_n) \in \{0,1\}^n$ where $\mathbb{P}(X_i = 1) = \frac{1}{2}$ for $i = 1, \ldots n$, and a *single* key bit U, shared by the two legitimate parties, where $\mathbb{P}(U = 1) = \frac{1}{2}$. Suppose that the distortion measure at the eavesdropper side is the Hamming distortion measure. Then, if the encrypted message is $\mathbf{Y} = (Y_1, \ldots, Y_n)$, where $Y_i = X_i \oplus U$, then the distortion at the eavesdropper attains its maximal possible value of $\frac{1}{2}$, regardless of the estimate of the eavesdropper. Nonetheless, such a secrecy is severely insecure. If the eavesdropper becomes aware of just a single bit of the source, then it can decrypt the entire message. It was therefore proposed to consider models which are more robust to assumptions concerning the eavesdropper. These models indeed lead to a non-degenerated trade-off, that requires a positive key rate. In [12], [13] it was assumed that the eavesdropper's estimation is performed sequentially, and at the time it estimates the i-th symbol, it has noiseless/noisy estimates of all the previous message symbols and the previous reproduced symbols (at the legitimate recipient), in addition to the public cryptogram. This model was termed *causal disclosure*. It was justified by the scenario in which the sender and legitimate recipient attempt to coordinate actions in a distributed system in order to maximize a certain payoff, and the eavesdropper acts in order to minimize the payoff. In a different line of work [14], the eavesdropper produces a fixed-size list (of exponential cardinality in the block-length), and the distortion is measured w.r.t. the reproduction word in the

list which attains the minimal distortion.

However, the fact that the trade-off in [9] is degenerated can be attributed to the way that the distortion is measured, rather than to the weakness of the eavesdropper. For a given strategy of the eavesdropper, the average distortion, as assumed in [9], [12], [14], may be large due to message and key-bit combinations that lead to a very large distortion, albeit with small probability. A more refined figure of merit would include the probability that the distortion is less than some level, rather than the average distortion. Such a performance criterion is customary in ordinary rate-distortion theory (e.g. the ϵ-fidelity criterion in [15, Chapter 7]). Indeed, in the above single key-bit example, the eavesdropper can estimate the message exactly with probability $\frac{1}{2}$, irrespective of its length. Thus, for any positive distortion level, the probability of an exiguous-distortion event is $\frac{1}{2}$, which is clearly unacceptable for most applications.

For most source models, good estimation of the message at the eavesdropper should be a rare event, and finding its exact probability is difficult. Instead, an asymptotic analysis can be carried in order to find the exponential decrease rate (i.e. the *exponent*) of the correct decryption probability. The results of [10] can be considered as a special case of this line of thought, for the restricted class of instantaneous encoders. In [10], the exponent of decrypting the message by the eavesdropper was found as a function of the exponent of exiguous-distortion of the estimation by the eavesdropper. For the same model, the exponent of the minimal probability of correct decryption by the eavesdropper was found in [16]. Later, in [17] secrecy was defined in a large-deviations sense: A system is considered secure if the exponent of the probability of the eavesdropper *correctly* decrypting the message is the same with and without the cryptogram. This, in turn, required the analysis of the correct decryption probability. In [10], [16], [17], it was assumed that the legitimate recipient must reproduce the message exactly (i.e., with zero distortion).

In this paper, we adopt a similar large-deviations approach to measuring secrecy, using a distortion measure, and generalize the results of [17]. For a memoryless source, we allow an imperfect reproduction at the legitimate recipient, and measure distortion both at the legitimate recipient and at the eavesdropper using a large-deviations measure. Specifically, we will define two exponents. First, for a given distortion level D_{L}, the *excess-distortion exponent* is defined in the usual way [15, Chapter 9], as the exponent of the probability that the distortion between the legitimate recipient reproduction and the source sequence is larger than D_{L}. Second, for a given distortion level D_{E}, we define the *exiguous-distortion exponent* as the exponent of the probability that the distortion between the eavesdropper estimate and the source sequence is *less* than D_{E}. We will derive the *perfect secrecy exponent* function $E_e^*(\mathsf{D}_{\mathrm{E}})$, which is the exiguous-distortion exponent of the eavesdropper when it estimates the message blindly, without the cryptogram (alternatively, for codes with unlimited key rate). It will be assumed that the secrecy system has a limited coding rate R_{L}, and that for a given distortion level D_{L}, the excess-distortion exponent must be larger than E_{L}. Our main result is that under mild conditions on the *compression constraints* $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$, the maximal achievable exiguous-distortion exponent is equal to the minimum between the key rate R, and $E_e^*(\mathsf{D}_{\mathrm{E}})$, calculated at distortion level required by the eavesdropper D_{E}. Since this maximal exiguous-distortion exponent

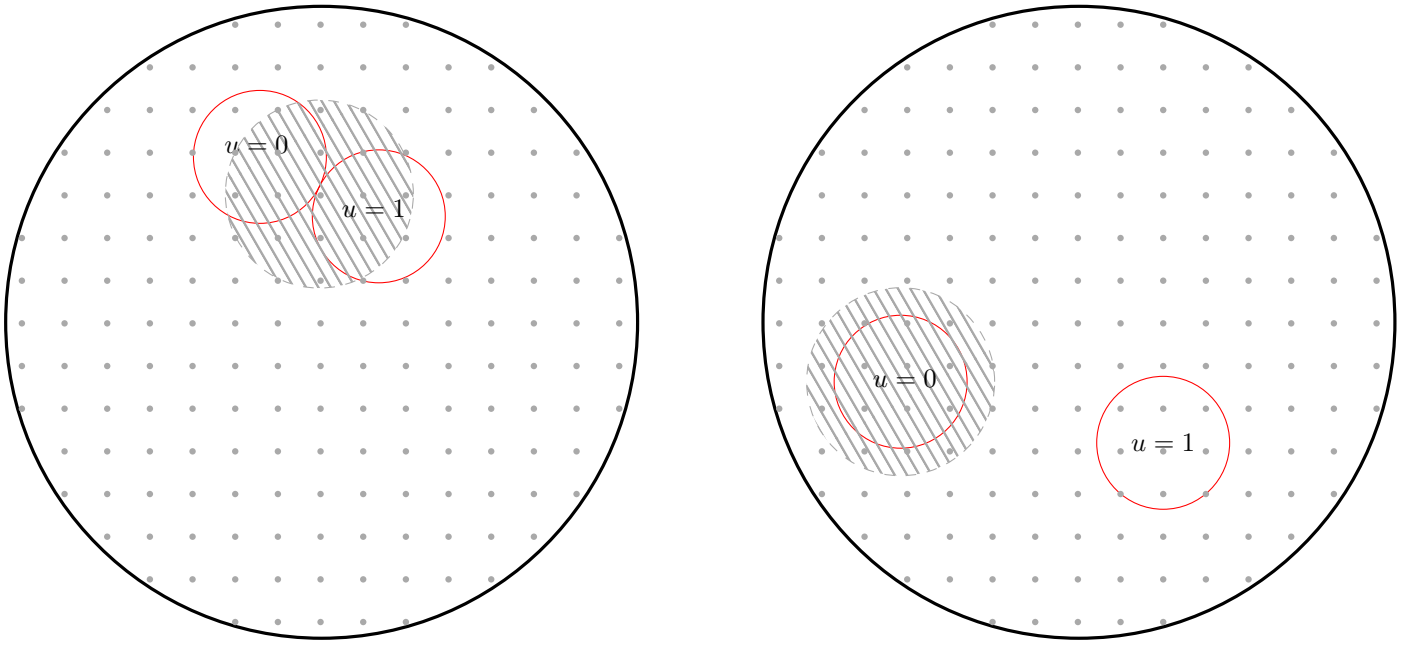

Figure 1. Two cases of ambiguity for the eavesdropper, for a single key bit code. Left side: Assume for simplicity that the source is distributed uniformly over the dots encapsulated by the outermost circle. The two small solid line circles represent two reproduction cells, which are mapped to the same cryptogram by the two possible values of the key bit u. The dashed larger circle represents all the source block for which the distortion between the source block and the best estimate of the eavesdropper is less than D_{E}. As can be seen, there is a large exiguous-distortion probability. Right side: Under the same assumptions, in this case the two reproduction cells are far apart. The best estimate of the eavesdropper can 'cover' at most one of the reproduction cells, and the exiguous-distortion probability is $\frac{1}{2}$.

does not depend on $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$ (in the interesting domain of these parameters), such a result implies that as far as performance trade-offs are concerned, the compression and secrecy problems are essentially decoupled: The fact that the message is required to be kept secret does not affect the compression performance. It should be stressed, however, that this result does not imply a separation theorem from the operational point of view. The rate-distortion code should be designed in a certain manner in order to provide secrecy, in contrast to, e.g., [9], [7], [18]. A concatenation of an arbitrary good rate-distortion code, followed by encryption using the available key bits, does not necessarily achieve a good exiguous-distortion exponent. For intuition, consider an ordinary rate-distortion code, assume that one key bit is available, and that the distortion measures of the legitimate decoder and eavesdropper are the same. The eavesdropper, in this case, knows that the reproduction of the legitimate decoder is one of two possible reproductions (of equal probability). If these two reproductions are close, then it can approximate them using a single reproduction, and achieve a distortion which may be only slightly larger than the distortion of the legitimate decoder. If, however, the rate-distortion code is designed in such a way that these two reproductions are sufficiently far apart, then the eavesdropper will have a poor compromise between them, and will achieve high distortion. This is illustrated in Figure 1. More generally, unlike ordinary rate-distortion codes, in which the performance is determined only by the reproduction cells, and the way in which the reproduction cells are mapped to transmitted bits is immaterial, here, the latter will be crucial for the security performance.

To show this result, we will prove both achievability (lower bound on the exiguous-distortion exponent) and a matching converse (upper bound). In the achievability part, we will demonstrate the existence of a secrecy system

in which the compression constraints are satisfied, and it has a fixed key rate R. For this secrecy system, the best strategy of the eavesdropper will be either to (1) guess the secret key and reproduce the message as a legitimate recipient (using the cryptogram), or (2) blindly estimate the message. The secrecy system constructed will also be *universal* in the following two senses. First, it does not require the knowledge of the source statistics, as long it is a memoryless source. Second, it is not designed for a specific value of D_{E}, yet the exiguous-distortion exponent $\min\{\mathsf{R}, E_e^*(\mathsf{D}_{\mathrm{E}})\}$ will be achieved for any value of D_{E}, *by the same sequence of codes*, as long as $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$. As a converse, we will show that even if *variable* key rate is allowed, yet with average key rate less than R, then the exiguous-distortion exponent cannot be larger than $\min\{\mathsf{R}, E_e^*(\mathsf{D}_{\mathrm{E}})\}$. The results of [17] are essentially recovered from our results, as a special case with $\mathsf{D}_{\mathrm{L}} = \mathsf{D}_{\mathrm{E}} = 0$. We also remark that in our model, the distortion measures of the legitimate recipient and the eavesdropper can be different, as long as they satisfy a certain relationship.

Finally, we briefly mention a related work in which large-deviations aspects were also incorporated. In [19], the *guessing* model of [20], [21] was relaxed to allow, after a maximum of possible guesses has passed, a small probability of large distortion for the eavesdropper. To analyze the asymptotic limits of the system, the excess-distortion exponent of the *eavesdropper* was restricted, and the maximal normalized logarithm of the number of guesses was found[1]. However, in our model, no testing mechanism is assumed to be available to the eavesdropper, which allows it to validate its estimate.

The outline of the rest of the paper is as follows. In Section II, we establish notation conventions, and in Section III, we formulate the problem. In Section IV, we present our main theorem, and discuss its implications. In Section V, we provide the outline and the main ideas of the proof. The proof of the main theorem appears in Section VI.

II. Notation Conventions

Throughout the paper, random variables will be denoted by capital letters, specific values they may take will be denoted by the corresponding lower case letters, and their alphabets will be denoted by calligraphic letters. Random vectors and their realizations will be denoted, respectively, by capital letters and the corresponding lower case letters, both in the bold face font. Their alphabets will be superscripted by their dimensions. For example, the random vector $\mathbf{X} = (X_1, \ldots, X_n)$ (n positive integer), may take a specific vector value $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathcal{X}^n$, the nth order Cartesian power of $\mathcal{X}$, which is the alphabet of each component of this vector. For any given vector $\mathbf{x}$, we will also denote $\mathbf{x}_i^j = (x_i, \ldots, x_j)$ for $1 \leq i \leq j \leq n$, and use the shorthand $\mathbf{x}_1^j = \mathbf{x}^j$.

We will follow the standard notation conventions for probability distributions, e.g., $P_X(x)$ will denote the probability of the letter $x \in \mathcal{X}$ under the distribution P_X. The arguments will be omitted when we address the entire distribution, e.g., P_X. Similarly, generic distributions will be denoted by Q, Q^*, and in other forms, subscripted by the relevant random variables/vectors/conditionings, e.g. Q_{XZ}, $Q_{X|Z}$. Whenever clear from context,

[1]Reference [19] is a one page abstract, and contains only a description of the problem. The results were not published, but a detailed version of [19] can be found in [22]. However, we believe that the achievability results provided in [22] are not actually proven. Specifically, in the achievability proof, no system is actually constructed, and the claims about the expected number of guesses of the eavesdropper are made on *any* given secrecy system. Obviously, there are, particularly bad, secrecy systems, in which a single guess suffices to find the message exactly.

these subscripts will be omitted. An exceptional case will be the 'hat' notation. For this notation, $\hat{Q}_{\mathbf{x}}$ will denote the empirical distribution of a vector $\mathbf{x} \in \mathcal{X}^n$, i.e., the vector of relative frequencies $\hat{Q}_{\mathbf{x}}(x)$ of each symbol $x \in \mathcal{X}$ in $\mathbf{x}$. The type class of $\mathbf{x} \in \mathcal{X}^n$, which will be denoted by $\mathcal{T}_n(\hat{Q}_{\mathbf{x}})$, is the set of all vectors $\mathbf{x}'$ with $\hat{Q}_{\mathbf{x}'} = \hat{Q}_{\mathbf{x}}$. The set of all type classes of vectors of length n over $\mathcal{X}$ will be denoted by $\mathcal{P}_n(\mathcal{X})$, and the set of all possible types over $\mathcal{X}$ will be denoted by $\mathcal{P}(\mathcal{X}) \triangleq \bigcup_{n=1}^{\infty} \mathcal{P}_n(\mathcal{X})$. Similar notation for type classes will also be used for generic types $Q_X \in \mathcal{P}(\mathcal{X})$, i.e., $\mathcal{T}_n(Q_X)$ will denote the set of all vectors $\mathbf{x}$ with $\hat{Q}_{\mathbf{x}} = Q_X$. In the same manner, the empirical distribution of a pair of vectors $(\mathbf{x}, \mathbf{z})$ will be denoted by $\hat{Q}_{\mathbf{xz}}$ and the joint type class will be denoted by $\mathcal{T}_n(\hat{Q}_{\mathbf{xz}})$. The joint type classes over the Cartesian product alphabet $\mathcal{X} \times \mathcal{Z}$ will be denoted by $\mathcal{P}_n(\mathcal{X} \times \mathcal{Z})$, and $\mathcal{P}(\mathcal{X} \times \mathcal{Z}) \triangleq \bigcup_{n=1}^{\infty} \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$. For a joint type $Q_{XZ} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z})$, $\mathcal{T}_n(Q_{XZ})$ will denote the set of all pairs of vectors $(\mathbf{x}, \mathbf{z})$ with $\hat{Q}_{\mathbf{xz}} = Q_{XZ}$. The conditional type class, namely, the set $\{\mathbf{x}' : \hat{Q}_{\mathbf{x}'\mathbf{z}} = \hat{Q}_{\mathbf{xz}}\}$, will be denoted by $\mathcal{T}_n(\hat{Q}_{\mathbf{x}|\mathbf{z}}, \mathbf{z})$, or more generally $\mathcal{T}_n(Q_{X|Z}, \mathbf{z})$ for a generic empirical conditional probability distribution $Q_{X|Z}$. The probability simplex for $\mathcal{X}$ will be denoted by $\mathcal{Q}(\mathcal{X})$, and the simplex for the alphabet $\mathcal{X} \times \mathcal{Z}$ will be denoted by $\mathcal{Q}(\mathcal{X} \times \mathcal{Z})$. Similar notations will be used for triplets of random variables.

For two distributions P_X, Q_X over the same finite alphabet $\mathcal{X}$, we will denote the variational distance ($\mathcal{L}_1$ norm) by

$$||P_X - Q_X|| \stackrel{\triangle}{=} \sum_{x \in \mathcal{X}} |P_X(x) - Q_X(x)|. \tag{1}$$

When optimizing a function of a distribution Q_X over the entire probability simplex $\mathcal{Q}(\mathcal{X})$, the explicit display of the constraint will be omitted. For example, for a function $f(Q)$, we will write $\min_Q f(Q)$ instead of $\min_{Q \in \mathcal{Q}(\mathcal{X})} f(Q)$. The same will hold for optimization of a function of a distribution Q_{XZ} over the probability simplex $\mathcal{Q}(\mathcal{X} \times \mathcal{Z})$, and for similar optimizations.

The expectation operator w.r.t. a given distribution, e.g., Q_{XZ}, will be denoted by $\mathbb{E}_Q[\cdot]$ where, the subscript Q_{XZ} will be omitted if the underlying probability distribution is clear from the context. In general, information-theoretic quantities will be denoted by the standard notation [23], with subscript indicating the distribution of the relevant random variables, e.g. $H_Q(X|Z), I_Q(X;Z), I_Q(X;Z|W)$, under $Q = Q_{XZW}$. For notational convenience, the entropy of X under Q will be denoted both by $H_Q(X)$ and $H(Q_X)$, depending on the context. The binary entropy function will be denoted by $h_{\mathrm{B}}(q)$ for $0 \leq q \leq 1$. The information divergence between two distributions, e.g. P_X and Q_X, will be denoted by $D(P_X||Q_X)$. In all information measures above, the distribution may also be an empirical distribution, for example, $H(\hat{Q}_{\mathbf{x}})$, $D(\hat{Q}_{\mathbf{x}}||P_X)$ and so on.

We will denote the Hamming distance between two vectors, $\mathbf{x} \in \mathcal{X}^n$ and $\mathbf{z} \in \mathcal{X}^n$, by $d_{\mathrm{H}}(\mathbf{x}, \mathbf{z})$. The length of a string b will be denoted by $|b|$, the concatenation of strings $b_1, b_2, \ldots$ will be denoted by $(b_1, b_2, \ldots)$, and the empty string will be denoted by ϕ. We will denote the complement of a set $\mathcal{A}$ by $\mathcal{A}^c$, and its interior by $\mathrm{int}(\mathcal{A})$. For a finite set $\mathcal{A}$, we will denote its cardinality by $|\mathcal{A}|$. The probability of the event $\mathcal{A}$ will be denoted by $\mathbb{P}(\mathcal{A})$, and $\mathbb{I}(\mathcal{A})$ will denote its indicator function.

For two positive sequences, $\{a_n\}$ and $\{b_n\}$ the notation $a_n \doteq b_n$, will mean asymptotic equivalence in the

exponential scale, that is, $\lim_{n\to\infty}\frac{1}{n}\log(\frac{a_n}{b_n})=0$. Similarly, $a_n\dot{\leq}b_n$ will mean $\limsup_{n\to\infty}\frac{1}{n}\log(\frac{a_n}{b_n})\leq 0$, and so on. The ceiling function will be denoted by $\lceil\cdot\rceil$. The notation $[t]_+$ will stand for $\max\{t,0\}$. For two integers, a,b, we denote by $a \bmod b$ the modulo of a w.r.t. b. Logarithms and exponents will be understood to be taken to the binary base.

Throughout, we will ignore integer code length constraints for the sake of simplicity, as they do not have any effect on the results. For example, instead of $\lceil n\mathsf{R}\rceil$ bits we will write $n\mathsf{R}$ bits. For a given finite ordered set, $\mathcal{A}=\{\mathbf{a}_1,\ldots,\mathbf{a}_{|\mathcal{A}|}\}$, we will denote by $\mathbb{B}[\mathbf{a};\log|\mathcal{A}|]$ the binary representation of the index of $\mathbf{a}$ in $\mathcal{A}$, i.e. $\mathbb{B}[\mathbf{a};\log|\mathcal{A}|]=i$ if $\mathbf{a}=\mathbf{a}_i$, for $i=1,\ldots|A|$.

In general, the subscript 'L' will be used for quantities related to the legitimate decoder, and the subscript 'E' will be used for eavesdropper-related quantities.

III. Problem Statement

Let the source vector $\mathbf{X}=(X_1,\ldots,X_n)$ be formed by n independent copies of a random variable $X\in\mathcal{X}$, where $\mathcal{X}$ is a finite alphabet, and X_i is distributed according to $P_X(x)=\mathbb{P}(X=x)$. Let $\mathcal{W}$ and $\mathcal{Z}$ be finite reproduction alphabets. In addition, let $\{U_i\}_{i=1}^{\infty}$ be a sequence of purely random bits (i.e. a Bernoulli process with $\mathbb{P}(U_i=1)=\frac{1}{2}$), independent of the source $\mathbf{X}$.

A *secure rate-distortion code* $\mathcal{S}_n=(f_n,\varphi_n)$ of block-length n is defined by a *key-length* function $k_n:\mathcal{X}^n\to\mathbb{Z}_+$, which assigns a key length $k_n(\mathbf{x})$ to every $\mathbf{x}\in\mathcal{X}^n$, an *encoder* $f_n:\mathcal{X}^n\times\{0,1\}^*\to\mathcal{Y}_n$, which generates a cryptogram, $y=f_n(\mathbf{x},\mathbf{u})$, where $\mathbf{u}=(u_1,\ldots,u_{k_n(\mathbf{x})})$, and where $\mathcal{Y}_n$ is a finite alphabet[2], and a *legitimate decoder* $\varphi_n:\mathcal{Y}_n\times\{0,1\}^*\to\mathcal{W}^n$, which generates a reproduction $\mathbf{w}=\varphi_n(y,\mathbf{u})$[3]. A sequence of codes $\{\mathcal{S}_n\}_{n\geq 1}$, indexed by the block-length n, is denoted by $\mathcal{S}$. The performance of the legitimate decoder is evaluated by a distortion measure $d_{\mathrm{L}}:\mathcal{X}\times\mathcal{W}\to\mathbb{R}_+$, where without loss of generality (w.l.o.g.), it is assumed that for every $x\in\mathcal{X}$, there exists $w\in\mathcal{W}$ such that $d_{\mathrm{L}}(x,w)=0$. Also, with a slight abuse of notation, the distortion between $\mathbf{x}$ and $\mathbf{w}$ is defined as the average,

$$d_{\mathrm{L}}(\mathbf{x},\mathbf{w})\triangleq\frac{1}{n}\sum_{i=1}^{n}d_{\mathrm{L}}(x_i,w_i). \tag{2}$$

We say that $\mathcal{S}$ satisfies a *compression constraint* $(\mathsf{R}_{\mathrm{L}},\mathsf{D}_{\mathrm{L}},\mathsf{E}_{\mathrm{L}})$, if the *coding rate* satisfies[4]

$$\limsup_{n\to\infty}\frac{1}{n}\log|\mathcal{Y}_n|\leq\mathsf{R}_{\mathrm{L}}, \tag{3}$$

and for any given $\{U_i\}_{i=1}^{\infty}=\{u_i\}_{i=1}^{\infty}$ the *excess-distortion exponent*, at distortion level D_{L}, is larger than E_{L} for

[2]This alphabet need not be the nth order Cartesian power of some alphabet $\mathcal{Y}$.

[3]It is implicit in the definition of the encoder and decoder that both are aware of the key-length $k_n(\mathbf{x})$. Specifically, one can define an *inverse-key length* function $l_n:\mathcal{Y}_n\times\{0,1\}^*\to\mathbb{Z}_+$, which reproduces the key-length at the decoder side, i.e. $k_n(\mathbf{x})=l_n(y,\{u_i\}_{i=1}^{\infty})$.

[4]This constraint can be weakened to a constraint on the normalized entropy of the cryptogram. See discussion in Section IV.

the legitimate decoder, i.e.[5]

$$\liminf_{n\to\infty} -\frac{1}{n}\mathbb{P}\left[d_{\mathsf{L}}(\mathbf{X},\varphi_n(f_n(\mathbf{X},\mathbf{u}),\mathbf{u}))\geq \mathsf{D}_{\mathsf{L}}\right]\geq \mathsf{E}_{\mathsf{L}}. \tag{4}$$

Note that for a zero excess-distortion exponent $\mathsf{E}_{\mathsf{L}}=0^+$, this requirement implies that an *average-distortion constraint*[6] $\mathbb{E}\left[d_{\mathsf{L}}(\mathbf{X},\mathbf{W})\right]\leq \mathsf{D}_{\mathsf{L}}$ is also satisfied. An *eavesdropper* decoder is a function $\sigma_n:\mathcal{Y}_n\to\mathcal{Z}^n$, where $\mathbf{z}=\sigma_n(y)$ is the *estimate* of the eavesdropper. It is assumed that the eavesdropper has full knowledge of all system properties: The source statistics, the encoder (f_n,k_n), and the legitimate decoder φ_n. The set of all eavesdropper decoders for a block-length n is denoted by Σ_n. In what follows, we also consider genie-aided eavesdropper decoders, which are aware of the type class of the source block, i.e., $\tilde{\sigma}_n:\mathcal{Y}_n\times\mathcal{P}_n\to\mathcal{X}^n$, and in this case, the estimate of the decoder is $\mathbf{z}=\tilde{\sigma}_n(y,\hat{Q}_{\mathbf{x}})$. The set of all genie-aided eavesdropper decoders of block-length n is denoted by $\tilde{\Sigma}_n$.

The performance of the eavesdropper is evaluated by a distortion measure $d_{\mathsf{E}}:\mathcal{X}\times\mathcal{Z}\to\mathbb{R}_+$, where again, it is assumed that for every $x\in\mathcal{X}$, there exists $z\in\mathcal{Z}$ such that $d_{\mathsf{E}}(x,z)=0$. As before, the distortion between $\mathbf{x}$ and $\mathbf{z}$ is defined as

$$d_{\mathsf{E}}(\mathbf{x},\mathbf{z})\triangleq\frac{1}{n}\sum_{i=1}^{n}d_{\mathsf{E}}(x_i,z_i). \tag{5}$$

For a given $\mathsf{D}_{\mathsf{E}}\geq 0$, the *exiguous-distortion probability*, for a given code $\mathcal{S}_n$, is denoted by

$$p_d(\mathcal{S}_n,\mathsf{D}_{\mathsf{E}})\triangleq\max_{\sigma_n\in\Sigma_n}\mathbb{P}\left[d_{\mathsf{E}}(\mathbf{X},\mathbf{Z})\leq\mathsf{D}_{\mathsf{E}}\right]. \tag{6}$$

The *limit inferior exiguous-distortion exponent*, achieved for a sequence of codes $\mathcal{S}$, is defined as

$$\mathcal{E}_d^-(\mathcal{S},\mathsf{D}_{\mathsf{E}})\triangleq\liminf_{n\to\infty}-\frac{1}{n}\log p_d(\mathcal{S}_n,\mathsf{D}_{\mathsf{E}}), \tag{7}$$

and the *limit superior exiguous-distortion exponent* achieved, $\mathcal{E}_d^+(\mathcal{S},\mathsf{D}_{\mathsf{E}})$, is defined analogously, with limit superior replacing the limit inferior. While, $\mathcal{E}_d^-(\mathcal{S},\mathsf{D}_{\mathsf{E}})\leq\mathcal{E}_d^+(\mathcal{S},\mathsf{D}_{\mathsf{E}})$, it is guaranteed that $p_d(s_n,\mathsf{D}_{\mathsf{E}})\dot{\geq}\exp\left[-n\mathcal{E}_d^-(\mathcal{S},\mathsf{D}_{\mathsf{E}})\right]$ for all sufficiently large block-lengths, while $p_d(s_n,\mathsf{D}_{\mathsf{E}})\doteq\exp\left[-n\mathcal{E}_d^+(\mathcal{S},\mathsf{D}_{\mathsf{E}})\right]$ may hold only for some sub-sequence of block-lengths. Thus, $\mathcal{E}_d^-(\mathcal{S},\mathsf{D}_{\mathsf{E}})$ is less sensitive to the choice of the block-length. For a given $Q_X\in\mathcal{P}(\mathcal{X})$, let $n_l=n_0l$, $l=1,2,\ldots$, be the sub-sequence of block-lengths such that $\mathcal{T}_n(Q_X)$ is non-empty, where n_0 is the minimal such block-length. We define, with a slight abuse of notation, the *conditional limit inferior*

[5]This constraint can be weakened to be only satisfied for an excess-distortion probability averaged over $\{U_i\}_{i=1}^{\infty}$. See discussion in Section IV.

[6]Indeed, suppose that $\mathbb{P}\left(d_{\mathsf{L}}(\mathbf{X},\varphi_n(f_n(\mathbf{X},\mathbf{u}),\mathbf{u}))\geq\mathsf{D}_{\mathsf{L}}\right)$ decays to zero for all $\{u_i\}_{i=1}^{\infty}$, but only sub-exponentially. Assuming $\overline{d}_{\mathsf{L}}\triangleq\min_{w\in\mathcal{W}}\max_{x\in\mathcal{X}}d_{\mathsf{L}}(x,w)<\infty$, for any $\delta>0$ and all n sufficiently large

$$\begin{aligned}\mathbb{E}\left[d_{\mathsf{L}}(\mathbf{X},\mathbf{W})\right]&\leq\mathsf{D}_{\mathsf{L}}\cdot\mathbb{P}\left[d_{\mathsf{L}}(\mathbf{X},\varphi_n(f_n(\mathbf{X},\mathbf{u}),\mathbf{u}))\leq\mathsf{D}_{\mathsf{L}}\right]+\overline{d}_{\mathsf{L}}\cdot\mathbb{P}\left[d_{\mathsf{L}}(\mathbf{X},\varphi_n(f_n(\mathbf{X},\mathbf{u}),\mathbf{u}))\leq\mathsf{D}_{\mathsf{L}}\right]\\&\leq\mathsf{D}_{\mathsf{L}}+\overline{d}_{\mathsf{L}}\cdot\mathbb{P}\left[d_{\mathsf{L}}(\mathbf{X},\varphi_n(f_n(\mathbf{X},\mathbf{u}),\mathbf{u}))\leq\mathsf{D}_{\mathsf{L}}\right]\\&\leq\mathsf{D}_{\mathsf{L}}+\delta.\end{aligned}$$

exiguous-distortion exponent as

$$\mathcal{E}_d^-(\mathcal{S},\mathsf{D}_{\mathrm{E}},Q_X)\triangleq\liminf_{l\to\infty}-\frac{1}{n_l}\log\max_{\sigma_{n_l}\in\Sigma_{n_l}}\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z})\le\mathsf{D}_{\mathrm{E}}|\mathbf{X}\in\mathcal{T}_{n_l}(Q_X)\right],\tag{8}$$

and $\mathcal{E}_d^+(\mathcal{S},\mathsf{D}_{\mathrm{E}},Q_X)$ is defined analogously.

The key rate of $\mathbf{x}\in\mathcal{X}^n$ is defined as $r_n(\mathbf{x})\triangleq\frac{1}{n}\left|k_n(\mathbf{x})\right|$. A code is termed a *fixed key rate* code of rate R_0 if $r_n(\mathbf{x})=\mathsf{R}_0$ for all $\mathbf{x}\in\mathcal{X}^n$, otherwise, it is called a *variable key rate* code, and it has an *average key rate* $\mathbb{E}[r_n(\mathbf{X})]$. We define the *conditional key rate* of $Q_X\in\mathcal{P}(\mathcal{X})$ as

$$\overline{R}(\mathcal{S},Q_X)\triangleq\lim_{l\to\infty}\mathbb{E}[r_{n_l}(\mathbf{X})|\mathbf{X}\in\mathcal{T}_{n_l}(Q_X)]\tag{9}$$

whenever the limit exist.

The rate-distortion function of a memoryless source Q_X, under the distortion measure $d_{\mathrm{L}}(\cdot,\cdot)$ is denoted by

$$R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})\triangleq\min_{Q_{W|X}:\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D}_{\mathrm{L}}}I_Q(X;W)\tag{10}$$

and, similarly, the rate-distortion function of Q_X under the distortion measure $d_{\mathrm{E}}(\cdot,\cdot)$ is denoted by $R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})$.

The main result of this paper, in Theorem 1, is a single-letter formula for the largest achievable exiguous-distortion exponent for codes under a compression constraint $(\mathsf{R}_{\mathrm{L}},\mathsf{D}_{\mathrm{L}},\mathsf{E}_{\mathrm{L}})$ and limited key rate.

IV. Main Result

The achievability part will be proved using fixed key rate codes, but in the converse part, we will allow also variable key rate codes, that satisfy the following assumptions:

1) *Upper bound on the key rate:* As $k_n(\mathbf{x})=n\log|\mathcal{X}|$ key-bits are always sufficient to perfectly encrypt the source, even without distortion, it will be assumed that $k_n(\mathbf{x})\le n\log|\mathcal{X}|$ for all $\mathbf{x}\in\mathcal{X}^n$.
2) *Uniform convergence of the conditional key rate:* We assume that for every $Q_X\in\mathcal{P}(\mathcal{X})$, conditioned on $\mathbf{X}\in\mathcal{T}_n(Q_X)$, the key rate $r_n(\mathbf{X})$ converges in probability to $\overline{R}(\mathcal{S},Q_X)$, and moreover, this convergence is uniform over $\mathcal{P}(\mathcal{X})$. Namely, for any $\delta>0$

$$\max_{Q_X\in\mathcal{P}_n(\mathcal{X})}\mathbb{P}\left[\left|r_n(\mathbf{X})-\overline{R}(\mathcal{S},Q_X)\right|>\delta|\mathbf{X}\in\mathcal{T}_n(Q_X)\right]\xrightarrow[n\to\infty]{}0.\tag{11}$$

   It is easy to prove that since $0\le r_n(\mathbf{X})\le\log|\mathcal{X}|$ with probability 1, then uniform convergence in the mean ($\mathcal{L}_1$ norm) is also satisfied, and the limit in (9) exists, uniformly over $Q_X\in\mathcal{P}(\mathcal{X})$.
3) *Admissible encoders:* An encoder f_n will be termed *admissible*, if $\mathbf{u}\neq\mathbf{u}'$ implies that $f_n(\mathbf{x},\mathbf{u})\neq f_n(\mathbf{x},\mathbf{u}')$ for all $\mathbf{x}\in\mathcal{X}^n$. We assume that f_n is an admissible encoder.

   In addition, we make two more assumptions. These assumptions are inessential, and are only made in order to simplify the exposition of our results.
4) *Upper bound on the legitimate excess-distortion exponent:* It is well known [15, Theorem 9.5],[24], that for

a given D_{L}, if

$$\liminf_{n\to\infty} \frac{1}{n} \log|\mathcal{Y}_n| \geq \mathsf{R}_{\mathrm{L}} \tag{12}$$

then there exist a sequence of codes $\mathcal{S}$ which satisfies the compression constraint $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$ iff

$$\mathsf{E}_{\mathrm{L}} \leq E_{\mathrm{L}}(P_X, \mathsf{D}_{\mathrm{L}}, \mathsf{R}_{\mathrm{L}}) \triangleq \inf_{Q_X : R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}}) > \mathsf{R}_{\mathrm{L}}} D(Q_X||P_X), \tag{13}$$

where $E_{\mathrm{L}}(P_X, \mathsf{D}_{\mathrm{L}}, \mathsf{R}_{\mathrm{L}})$ is known as *Marton's source coding exponent*. It will be assumed that the required excess-distortion exponent at the legitimate decoder is strictly positive and not larger than Marton's exponent, i.e., $0 < \mathsf{E}_{\mathrm{L}} \leq E_{\mathrm{L}}(P_X, \mathsf{D}_{\mathrm{L}}, \mathsf{R}_{\mathrm{L}})$.

5) *Partial ordering between distortion measures:* The distortion measure $d_{\mathrm{E}}(\cdot,\cdot)$ will be termed more *lenient* than $d_{\mathrm{L}}(\cdot,\cdot)$, if for every $\mathbf{w} \in \mathcal{W}^n$, there exists $\mathbf{z} \in \mathcal{Z}^n$ such that

$$\{\mathbf{x} \in \mathcal{X}^n : d_{\mathrm{L}}(\mathbf{x}, \mathbf{w}) \leq \mathsf{D}\} \subseteq \{\mathbf{x} \in \mathcal{X}^n : d_{\mathrm{E}}(\mathbf{x}, \mathbf{z}) \leq \mathsf{D}\}, \tag{14}$$

for every $\mathsf{D} \geq 0$. This corresponds to a worst case assumption regarding the distortion measure (and the reproduction alphabet $\mathcal{Z}$) used by the eavesdropper - it is at least not more demanding than the distortion measure used by the legitimate decoder. In addition, this also puts, in some sense, the distortion levels at the legitimate decoder and at the eavesdropper decoder, on the same scale. Therefore, it will be assumed that $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$, namely, the distortion level allowed by the eavesdropper is larger than the one allowed by the legitimate decoder. It is also easily verified that this assumption implies

$$R_{\mathrm{E}}(Q_X, \mathsf{D}) \leq R_{\mathrm{L}}(Q_X, \mathsf{D}) \tag{15}$$

for every $\mathsf{D} > 0$.

We denote by

$$E_e^*(\mathsf{D}_{\mathrm{E}}) \triangleq \min_{Q_X} \{D(Q_X||P_X) + R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}})\} \tag{16}$$

the *perfect-secrecy exponent.* Using standard method of types, it can be shown that this is the maximal exiguous-distortion exponent that can be achieved when the eavesdropper blindly estimates the source, i.e. without using the cryptogram. Alternatively, as evident from Theorem 1, this is the maximal exponent for unlimited key rate. We are now ready to state our main result.

Theorem 1. *Let $\delta > 0$ be given. Then, there exists a sequence of codes $\mathcal{S}$ of fixed key rate R, which satisfies a compression constraint $(\mathsf{R}_{\mathrm{L}} + \delta, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$ and properties 1-5 above,*

$$\mathcal{E}_d^-(\mathcal{S}, \mathsf{D}_{\mathrm{E}}) \geq \min\{\mathsf{R}, E_e^*(\mathsf{D}_{\mathrm{E}})\} - \delta \tag{17}$$

for all $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$. Conversely, for every sequence of codes $\mathcal{S}$ of average key rate $\mathbb{E}[r_n(\mathbf{x})] \leq \mathsf{R}$ for all n, which

satisfies a compression constraint $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$ *and properties 1-5 above,*

$$\mathcal{E}_d^+(\mathcal{S}, \mathsf{D}_{\mathrm{E}}) \leq \min\{\mathsf{R}, E_e^*(\mathsf{D}_{\mathrm{E}})\} \tag{18}$$

for all $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$.

Section VI is devoted to the proof of Theorem 1, and here we discuss its implications. The main implication of this theorem is that the performance of lossy compression and encryption are essentially decoupled. Note that in Theorem 1, the exiguous-distortion exponent of the eavesdropper is determined solely by the key rate and the distortion level D_{E} at the eavesdropper, and not by the compression constraint $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$ (as long as the assumptions hold). Specifically, it holds for $\mathsf{D}_{\mathrm{L}} = 0$, which means that increasing D_{L} does not increase D_{E}. In other words, reducing the amount of information sent to the legitimate decoder cannot improve secrecy. Nonetheless, on a positive note, as long as $\mathsf{R} \leq E_e^*(\mathsf{D}_{\mathrm{E}})$, the maximal secrecy can be attained, for every $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$, without affecting the compression performance. In addition, note that in Theorem 1, D_{E} has a special stature: A single sequence of codes $\mathcal{S}$ is *universal* for all $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$. This enables the construction of secure rate-distortion codes that are robust to the choice of D_{E}, which may be unspecified when designing the system.

As previously mentioned, the achievability part of Theorem 1 is proved using fixed rate codes. Since fixed rate codes clearly satisfy the second assumption above, the maximal exiguous-distortion exponent is fully characterized for fixed key rate coding. Furthermore, the theorem shows that variable key rate codes, from the class of codes which satisfy the above assumptions, offer no advantage over fixed key rate codes in terms of exiguous-distortion exponent. This is in contrast to similar problems (variable-rate channel coding with feedback [25], [26], variable-rate Slepian-Wolf coding [27]), where the more lenient average-rate constraint allows to increase the error exponent. It should be mentioned that while the class of variable key rate codes is restricted to satisfy uniform convergence in probability of the conditional key rate (see the second assumption above), the important class of *type dependent variable key rate* codes satisfy this assumption. In a type dependent variable key rate code, the key rate $r_n(\mathbf{x})$ depends on $\mathbf{x}$ only via its type, namely, $\hat{Q}_{\mathbf{x}} = \hat{Q}_{\tilde{\mathbf{x}}}$ implies $r_n(\mathbf{x}) = r_n(\tilde{\mathbf{x}}) = \rho(Q_X)$ for some *key rate function* $\rho(\cdot) : \mathcal{P}(\mathcal{X}) \to \mathbb{R}^+$. Due to the symmetry of source blocks from the same type class, such a key rate allocation is indeed plausible, and also practically motivated due to its simplicity. Such codes trivially satisfy the convergence requirement, and so the converse part of Theorem 1 is valid.

Theorem 1 essentially generalizes [17, Theorem 1]. In [17], it was assumed that all alphabets are identical $\mathcal{X} = \mathcal{W} = \mathcal{Z}$, and that $\mathsf{D}_{\mathrm{E}} = \mathsf{D}_{\mathrm{L}} = 0$. Thus, the legitimate decoder need to perfectly reproduce the source block, and the eavesdropper performance is measured by its probability of correct estimate, i.e.

$$p_d(\mathcal{S}_n, \mathsf{D}_{\mathrm{E}}) = \max_{\sigma_n \in \Sigma_n} \mathbb{P}(\mathbf{X} = \mathbf{Z}). \tag{19}$$

Note also that for this specific case, the perfect-secrecy exponent for this case is given by

$$E_e^*(\mathsf{D}_{\mathrm{E}}) = \min_{Q_X} \{D(Q_X||P_X) + H(Q_X)\} \tag{20}$$

$$= -\log \max_{x\in\mathcal{X}} P_X(x). \tag{21}$$

Indeed, even without using the cryptogram, the eavesdropper can choose $\mathbf{z} = (x^*, \ldots, x^*)$ where $x^* = \max_{x\in\mathcal{X}} P_X(x)$, and achieve $E_e^*(\mathsf{D}_{\mathrm{E}})$.

V. Outline of the Proof of Theorem 1

Since the proof of Theorem 1 is considerably involved, this section is devoted to an informal description of the structure and the main ideas in this proof. Hopefully, this will facilitate the reading of the formal proof, or at least give the reader an idea of the main highlights.

To begin, we observe, in Subsection VI-A, that the exiguous-distortion exponent remains unchanged even if the eavesdropper is aware of the type of the source block $\hat{Q}_{\mathbf{x}}$. This enables us to first, consider each type of the source separately, and only then incorporate all types simultaneously, both in the achievability and the converse parts. Next, in Subsection VI-B, we provide a technique which facilitates the construction of secure rate-distortion codes, such that in view of the eavesdropper the cryptograms are symmetric. The idea is to *cover* a type class $\mathcal{T}_n(Q_X)$ using an essentially minimal number of permutations of a constituent set $\mathcal{D}_n \subseteq \mathcal{T}_n(Q_X)$. To wit, if $\mathcal{D}_n \triangleq \{\mathbf{x}(0), \ldots, \mathbf{x}(|\mathcal{D}_n|-1)\}$ then for any permutation π over $\{1, \ldots, n\}$, we define

$$\pi(\mathcal{D}_n) \triangleq \{\pi(\mathbf{x}(0)), \ldots, \pi(\mathbf{x}(|\mathcal{D}_n|-1))\}, \tag{22}$$

and then find a set of permutations $\{\pi_{n,t}\}_{t=0}^{\kappa_n}$ such that

$$\bigcup_{t=0}^{\kappa_n} \pi_{n,t}(\mathcal{D}_n) = \mathcal{T}_n(Q_X), \tag{23}$$

where κ_n is asymptotically close to its minimal value of $\frac{|\mathcal{T}_n(Q_X)|}{|\mathcal{D}_n|}$. For ordinary rate-distortion, such covering lemma can be used to show the existence of a good rate-distortion code (e.g. instead of [15, Lemma 9.1]). Let us define, the *D-cover* of $\mathbf{w} \in \mathcal{W}^n$ as

$$\mathfrak{D}(\mathbf{w}, Q_X, \mathsf{D}_{\mathrm{L}}) \triangleq \{\mathbf{x} \in \mathcal{T}_n(Q_X) : d_{\mathrm{L}}(\mathbf{x}, \mathbf{w}) \leq \mathsf{D}_{\mathrm{L}}\}. \tag{24}$$

If we set $\mathcal{D}_n = \mathfrak{D}(\mathbf{w}, Q_X, \mathsf{D}_{\mathrm{L}})$ and find permutations $\{\pi_{n,t}\}_{t=0}^{\kappa_n}$ such that (23) holds, then the set $\hat{\mathcal{C}}_n \triangleq \{\pi_{n,t}(\mathbf{w})\}_{t=0}^{\kappa_n}$ is a rate-distortion code such that for every $\mathbf{x} \in \mathcal{T}_n(Q_X)$ there exists $\mathbf{w} \in \hat{\mathcal{C}}_n$ such that $d_{\mathrm{L}}(\mathbf{x}, \mathbf{w}) \leq \mathsf{D}_{\mathrm{L}}$. Such permutations can be found for all types of the source, and using the method of types, it can be verified that Marton's source coding exponent can be achieved by such a construction. For the construction of secure rate-distortion codes, we will use permutations of more complicated sets to cover the type.

The achievability part (lower bound) is proved in Subsection VI-C using codes of fixed key rate R. Let us

first focus on a single type Q_X. For the legitimate decoder, a source block $\mathbf{x} \in \mathcal{T}_n(Q_X)$ is reproduced by some $\mathbf{w} \in \overline{\mathcal{C}}_n \triangleq \left\{\varphi_n(y,\mathbf{u}) : y \in \mathcal{Y}_n, \mathbf{u} \in \{0,1\}^{n\mathsf{R}}\right\}$, which satisfies $d_{\mathrm{L}}(\mathbf{x},\mathbf{w}) \leq \mathsf{D}_{\mathrm{L}}$, unless no such $\mathbf{w}$ exists. The compression constraint $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$ ensures that large-distortion reproduction occurs with an exponentially decaying probability. The eavesdropper, on the other hand, reproduces using only the cryptogram y. With a slight abuse of notation of (24), let us define, for a given the D-cover of $\mathcal{C}_n \subseteq \mathcal{W}^n$ as

$$\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}) \triangleq \bigcup_{\mathbf{w}\in\mathcal{C}_n} \mathfrak{D}(\mathbf{w}, Q_X, \mathsf{D}_{\mathrm{L}}). \tag{25}$$

When the eavesdropper observes y, it knows that the legitimate decoder will reproduce $\mathbf{w}$ from the set $\mathcal{C}_n(y) = \left\{\varphi_n(y,\mathbf{u}) : \mathbf{u} \in \{0,1\}^{n\mathsf{R}}\right\}$ of size $|\mathcal{C}_n(y)| = 2^{n\mathsf{R}}$. Furthermore, conditioning on the cryptogram y and the type Q_X, the source block $\mathbf{X}$ is distributed uniformly over $\mathfrak{D}(\mathcal{C}_n(y), Q_X, \mathsf{D}_{\mathrm{L}})$. The proof of achievability is divided into three steps. In the first step (Lemma 7), we demonstrate the existence of a good and secure rate-distortion code conditioned on a single cryptogram, in the second step, we extend this code for an entire type class $\mathcal{T}_n(Q_X)$ (Lemma 9), and in the third step, we extend it to all types.

In more detail, the first step of the proof (Lemma 7) shows, by a random selection mechanism, that there exists a set $\mathcal{C}_n^*$ of size $2^{n\mathsf{R}}$ such that when $\mathbf{X}$ is distributed uniformly over $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$, the exiguous-distortion probability of any eavesdropper is asymptotically not larger than $2^{-n\cdot\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\}}$. Geometrically, this implies that the *D-covers* for $\mathbf{w} \in \mathcal{C}_n$ are distant from each other, under $d_{\mathrm{E}}(\cdot,\cdot)$. Thus, a secure rate-distortion code satisfying $\mathcal{C}_n(y) = \mathcal{C}_n^*$ for some cryptogram y, will have a good conditional exiguous-distortion probability given y.

In the second step, we define the code for all $\mathbf{x} \in \mathcal{T}_n(Q_X)$, using a symmetry argument. Observe that the distortion measures of both the legitimate and eavesdropper decoders are invariant to permutations (see (2) and (5)). Thus, $\mathfrak{D}(\pi(\mathcal{C}_n), Q_X, \mathsf{D}_{\mathrm{L}}) = \pi\left(\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})\right)$, and the exiguous-distortion probability for an eavesdropper when $\mathbf{X}$ is distributed uniformly over $\pi\left(\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})\right)$ is the same as for $\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})$. In Lemma 9, we use a minimal number of permutations (from Subsection VI-B) of a good D-cover $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$ to cover $\mathcal{T}_n(Q_X)$, and then obtain a good secure rate-distortion code for all $\mathcal{T}_n(Q_X)$. There is a certain subtlety in the proof of Lemma 9. For an ordinary rate-distortion code, there might be more than a single $\mathbf{w} \in \overline{\mathcal{C}}_n$ such that $d_{\mathrm{L}}(\mathbf{x},\mathbf{w}) \leq \mathsf{D}_{\mathrm{L}}$. From the excess-distortion probability point of view, there is no importance to which one of these $\{\mathbf{w}\}$ will reproduce $\mathbf{x}$. However, this might result in $\mathbf{w} \in \overline{\mathcal{C}}_n$ for which only a small portion of $\mathfrak{D}(\mathbf{w}, Q_X, \mathsf{D}_{\mathrm{L}})$ is actually reproduced by $\mathbf{w}$ (as $\mathbf{x} \in \mathfrak{D}(\mathbf{w}, Q_X, \mathsf{D}_{\mathrm{L}})$ might be reproduced by some $\mathbf{w}' \in \overline{\mathcal{C}}_n$ which also satisfies $d_{\mathrm{L}}(\mathbf{x}, \overline{\mathbf{w}}) \leq \mathsf{D}_{\mathrm{L}}$), which might be harmful for secrecy purposes. Indeed, the secure rate-distortion code is constructed in Lemma 9 with the will that conditioned on any cryptogram y, the source is distributed uniformly over $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$. But, since a source block must eventually be reproduced by a single $\mathbf{w}$, then conditioned on some of the cryptograms y, the source block will be distributed on a smaller set than $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$. For such cryptograms, the conditional exiguous-distortion probability of the eavesdropper might be large. Lemma 9 shows that if the efficient covering described above is utilized, then the total effect of such events is negligible.

Until this stage, we have constructed a code for $\mathcal{T}_n(Q_X)$ with appropriate conditional exiguous-distortion exponent. As we shall see, in the construction of Lemma 7 and Lemma 9, the convergence of probabilities to their asymptotic exponent is not necessarily uniform (cf. Remark 8). In the third step of the achievability proof, we prove that uniform convergence is possible, using an elaborated construction, built from the previous one. The idea is to consider a dense grid on the simplex $\mathcal{Q}(\mathcal{X})$, and construct a secure rate-distortion code, as in Lemma 9, for each of the types in the grid. Since the of number of types in the grid is *finite*, then uniform convergence is assured for types in the grid. If the type of the source block belongs to the grid, then one of the constructed codes is used, according to its type. Otherwise, the source block will be first modified, such that the modified source block does have type within the grid, which is not very far from the type of the original source block. The modified source block will then be encoded using one of the codes of the grid, and thus will have both low legitimate excess-distortion probability, and large exiguous-distortion probability for the eavesdropper. It will be shown that the overheads required for the legitimate decoder to reproduce the original source block, rather than the modified source block are negligible.

In Subsection VI-D, we prove the converse part in two steps. Recall that in general, for any given type $Q_X \in \mathcal{P}(\mathcal{X})$, we have defined the average rate $\overline{R}(\mathcal{S}, Q_X)$, but we allow each source block $\mathbf{x} \in \mathcal{T}_n(Q_X)$ to have a different key rate $r_n(\mathbf{x}) \in [0, \log_n|\mathcal{X}|]$. In addition, for a code satisfying the compression constraint $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$, and type Q_X such that $D(Q_X||P_X) \leq \mathsf{E}_{\mathrm{L}}$, the legitimate excess-distortion probability must decay to zero exponentially as $2^{-n[\mathsf{E}_{\mathrm{L}} - D(Q_X||P_X)]}$ but does not need to be strictly zero. In the first step of the proof of the converse, we prove a lemma that shows that the optimal limit superior exiguous-distortion exponent is not deteriorated, if we restrict $r_n(\mathbf{x})$ to be a constant within $\mathcal{T}_n(Q_X)$, which is less than $\overline{R}(\mathcal{S}, Q_X) + \delta$, and also restrict the legitimate excess-distortion probability to be exactly zero. It will be easier to prove a converse for codes with such properties, as will be done in the second step of the proof. In the second step, we assume the structure of the code from the first step, and evaluate the performance of an eavesdropper which adopts one of the following two simple strategies: (1) It can guess the secret key bits, and then decode using these bits just like the legitimate decoder. (2) It can ignore the cryptogram altogether and choose an estimate $\mathbf{z} \in \mathcal{Z}^n$, based on only $\hat{Q}_{\mathbf{x}}$. Clearly, in the first case, the probability of success is $2^{-n\mathsf{R}}$, and it is not difficult to show that the exiguous-distortion probability for the second strategy is asymptotically $2^{-nE_e^*(\mathsf{D}_{\mathrm{E}})}$. This implies the upper bound (18). We remark that the asymptotic optimality of these two simple strategies (sometimes called *key-attack* and *blind guessing*, respectively) can also be found to some extent in related problems [14], [21], [22].

We conclude the outline of the proof with the following comments:

- *Awareness of key-length:* Since the number of possible key-lengths is $n\log|\mathcal{X}|$, it can be compressed and fully encrypted using negligible coding rate and key rate of $\frac{1}{n}\log(n\log|\mathcal{X}|)$ bits, and it can be assumed that the exiguous-distortion exponent is not deteriorated if the eavesdropper is aware of the key-length (as in Subsection VI-A). Thus, in the converse proof, we could have found the exiguous-distortion exponent conditioned on both the type and the key-length, and then average over them. The main obstacle in this

approach is proving the second property (full type covering) assured in Lemma 13. To show this property using the methods of Lemma 13, would require showing that the subsets of the type classes of fixed key-length, i.e., $\tilde{\mathcal{T}}_n(Q_X, m) \triangleq \mathcal{T}_n(Q_X) \cap \{\mathbf{x} : k_n(\mathbf{x}) = m\}$ for some $0 \leq m \leq n \log|\mathcal{X}|$, can cover a type class by essentially a minimal number of permutations, as in Lemma 4 (Subsection VI-B). However, in turn, the proof of Lemma 4 is based on the fact that $\mathcal{T}_n(Q_X)$ is invariant to permutations, which may not hold for $\tilde{\mathcal{T}}_n(Q_X, m)$.

- *Full type covering:* Let $Q_X \in \mathcal{P}(\mathcal{X})$ be given such that $D(Q_X||P_X) < \mathsf{E}_{\mathrm{L}}$. The method of types and the expression (13) reveal that to satisfy the compression constraint $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$, the following condition should hold for any given $\{u_i\}_{i=1}^{\infty}$

$$\mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \varphi_n(f_n(\mathbf{X}, \mathbf{u}), \mathbf{u})) > \mathsf{D}_{\mathrm{L}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \doteq 2^{-n[\mathsf{E}_{\mathrm{L}} - D(Q_X||P_X)]}. \tag{26}$$

  For ordinary rate-distortion codes, it is well known[7] that if for a given $\epsilon \in (0, 1)$ and for all n sufficiently large

$$\mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \mathbf{W}) > \mathsf{D}_{\mathrm{L}}\right] \leq 1 - \epsilon \tag{27}$$

  then there exists a rate-distortion code with almost the same rate, such that

$$\mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \mathbf{W}) > \mathsf{D}_{\mathrm{L}}\right] = 0. \tag{28}$$

  Thus, to ensure an exponent constraint E_{L} for ordinary rate-distortion codebook, the type classes of types which are 'close' enough to P_X (in the divergence sense) should be almost covered by the reproduction set (26), but in fact, can be *fully* covered by the reproduction set (28). Then, the minimal rate required to satisfy (26) is the same as the minimal rate to satisfy (28), and the compression rate cannot be decreased due to the softer requirement in (26). By contrast, in the presence of the eavesdropper, it might happen that the softer requirement in (26) can lead to better exiguous-distortion exponent: Even if a type class can be fully covered using the available coding rate, perhaps the exiguous-distortion exponent can be improved if some of the source blocks are reproduced with distortion larger than D_{L}, but this occurs with sufficiently small probability, as in (26). Lemma 13 shows that this is *not* the case.
- *Compression constraint conditions:* The conditions required to satisfy the coding rate constraint (3), and the excess-distortion exponent constraint for the legitimate decoder (4) can be weakened without affecting Theorem 1. First, (3) can be weakened to

$$\limsup_{n\to\infty} \frac{1}{n} H(Y) \leq \mathsf{R}_{\mathrm{L}}, \tag{29}$$

  where $H(Y)$ is the entropy of the cryptogram. Second, the excess-distortion exponent can be weakened to

[7]This can also be easily verified using Lemma 4.

apply to the expectation constraint over the key-bits $\{U_i\}_{i=1}^{\infty}$, rather than for every given $\{u_i\}_{i=1}^{\infty}$, i.e.

$$\liminf_{n\to\infty} -\frac{1}{n}\mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X},\varphi_n(f_n(\mathbf{X},\mathbf{U}),\mathbf{U}))\geq \mathsf{D}_{\mathrm{L}}\right]\geq \mathsf{E}_{\mathrm{L}}. \tag{30}$$

Obviously, since the achievability part is proved using the stronger conditions (3) and (4), it also holds under the weaker conditions (29) and (30). For the converse, note that in Lemma 13 and in the proof of the converse, the coding rate is essentially not constrained. The excess-distortion exponent constraint is used in the converse proof only in eq. (255), which follows directly from the weaker condition (30). Therefore, the achievability part holds under the strong conditions, and the converse part holds under the weak conditions.

- *Legitimate excess-distortion exponent:* As is evident from Theorem 1, there is no improvement in the exiguous-distortion exponent even if E_{L} vanishes (to wit, the distortion D_{L} is achieved only on the average). Thus, the excess-distortion exponent can be set to its maximal value of $E_{\mathrm{L}}(P_X,\mathsf{D}_{\mathrm{L}},\mathsf{R}_{\mathrm{L}})$, as defined in (13).
- *Dependency on the source distribution:* From the proof of the achievability, it is evident that given $\hat{Q}_{\mathbf{x}}$, the operation of the encoder, the legitimate decoder and the eavesdropper decoder depend on P_X only on whether $\mathsf{R}_{\mathrm{L}} > R_{\mathrm{L}}(\hat{Q}_{\mathbf{x}},\mathsf{D}_{\mathrm{L}})$ or not (equivalently, from the previous comment, whether $D(Q_X||P_X)\leq \mathsf{E}_{\mathrm{L}}$ or not). Since it can be assumed that $\hat{Q}_{\mathbf{x}}$ is known to all parties, then prior knowledge of the source distribution P_X is not required to either party. Hence, the secure rate-distortion codes constructed are *universal*. Of course, the exponents achieved depend on P_X.

VI. Proof of the Theorem 1

We remind the reader the *reverse Markov inequality* [28, Section 9.3, p. 159], which is a useful tool for the proof.

Lemma 2. *Let X be a positive random variable which satisfies $\mathbb{P}(X\leq \alpha\mathbb{E}[X])=1$ for some $\alpha>1$. Then, for any $\beta<1$,*

$$\mathbb{P}\left(X>\beta\mathbb{E}[X]\right)\geq \frac{1-\beta}{\alpha-\beta}. \tag{31}$$

The proof is based on the ordinary Markov inequality for the positive random variable $\tilde{X}=\alpha\mathbb{E}[X]-X$.

A. Type Awareness of the Eavesdropper

Consider the following simple observation, which simplifies later derivations: The largest achievable exiguous-distortion exponent is not deteriorated if the eavesdropper is aware of the type of the source block, in addition to the cryptogram.

Proposition 3. *For any $Q_X\in\mathcal{P}(\mathcal{X})$*

$$\mathcal{E}_d^{-}(\mathcal{S},\mathsf{D}_{\mathrm{E}},Q_X)=\liminf_{n\to\infty}\left\{-\frac{1}{n}\max_{\sigma_n\in\tilde{\Sigma}_n}\log\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z})\leq \mathsf{D}_{\mathrm{E}}|\mathbf{X}\in\mathcal{T}_n(Q_X)\right]\right\}. \tag{32}$$

An analogous result holds for $\mathcal{E}_d^{+}(\mathcal{S},\mathsf{D}_{\mathrm{E}},Q_X)$.

Proof: Since $\Sigma_n \subset \tilde{\Sigma}_n$

$$\mathcal{E}_d^-(\mathcal{S}, \mathsf{D}_{\mathrm{E}}, Q_X) \geq \liminf_{n\to\infty} \left\{ -\frac{1}{n} \log \max_{\sigma_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X) \right] \right\}. \tag{33}$$

To show equality, let $\{\tilde{\sigma}_n^* \in \tilde{\Sigma}_n\}$ be the sequence of decoders which achieve the maximum in the right hand side of (33). Let us define a sequence of decoders $\{\sigma_n \in \Sigma_n\}$ as follows. First, σ_n produces a random guess $Q \in \mathcal{P}_n$ of the type of the source, with the uniform distribution over $\mathcal{P}_n$, and second, it decodes

$$\sigma_n(y) = \tilde{\sigma}_n^*(y, Q). \tag{34}$$

Given $Q_X \in \mathcal{P}$, the resulting conditional exiguous-distortion probability is given by

$$\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \sigma_n(Y)) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X) \right] \tag{35}$$

$$\geq \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \tilde{\sigma}_n^*(Y, Q)) \leq \mathsf{D}_{\mathrm{E}} | Q = \hat{Q}_{\mathbf{x}}, \mathbf{X} \in \mathcal{T}_n(Q_X) \right] \cdot \mathbb{P}\left[Q = \hat{Q}_{\mathbf{x}} | \mathbf{X} \in \mathcal{T}_n(Q_X) \right] \tag{36}$$

$$= \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \tilde{\sigma}_n^*(Y, \hat{Q}_{\mathbf{x}})) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X) \right] \cdot \frac{1}{|\mathcal{P}_n|} \tag{37}$$

and as $|\mathcal{P}_n| \leq (n+1)^{|\mathcal{X}|}$, equality is achieved in (33). ■

B. Covering a Type Class via Permutations

In this subsection, we discuss the possibility to cover a type class by means of permutations of a constituent subset. The fact that the distortion measure of the eavesdropper is invariant to permutations of both arguments hints on the usefulness of such a covering in the construction of good secure rate-distortion codes.

Given a type $Q_X \in \mathcal{P}(\mathcal{X})$ and $\delta > 0$, the method of types implies that for $n > n_0(\delta, |\mathcal{X}|)$

$$2^{n[H(Q_X)-\delta]} \leq |\mathcal{T}_n(Q_X)| \leq 2^{nH(Q_X)}. \tag{38}$$

Now, consider the subset $\mathcal{D}_n \subset \mathcal{T}_n(Q_X)$, where the elements of $\mathcal{D}_n$ are distinct. We say that a set of permutations $\{\pi_{n,t}\}_{t=0}^{\kappa_n}$ *cover* $\mathcal{T}_n(Q_X)$ if

$$\bigcup_{t=0}^{\kappa_n} \pi_{n,t}(\mathcal{D}_n) = \mathcal{T}_n(Q_X), \tag{39}$$

where $\pi_{n,t}(\mathcal{D}_n)$ means that the same permutation $\pi_{n,t}(\cdot)$ operates on all $\mathbf{x} \in \mathcal{D}_n$, as defined in (22). Let κ_n^* be the minimal number of permutations of $\mathcal{D}_n$ required to cover $\mathcal{T}_n(Q_X)$. By a simple counting argument, we must have

$$\kappa_n^* \geq \frac{|\mathcal{T}_n(Q_X)|}{|\mathcal{D}_n|}. \tag{40}$$

The following lemma guaranteed the existence of a cover which essentially achieves the lower bound.

Lemma 4 ([29, Section 6, Covering Lemma 2])**.** *For every $\mathcal{D}_n \subset \mathcal{T}_n(Q_X)$, $Q_X \in \mathcal{P}_n(\mathcal{X})$*

$$\kappa_n^* \leq \frac{|\mathcal{T}_n(Q_X)|}{|\mathcal{D}_n|} \cdot \log |\mathcal{T}_n(Q_X)|. \tag{41}$$

The main application of this lemma is for a sequence of sets $\{\mathcal{D}_n\}_{n=1}^{\infty}$. Let n_l be the sequence of block-lengths such that $\mathcal{T}_{n_l}(Q_X)$ is non-empty, and let $\mathcal{D}_{n_l} \subset \mathcal{T}_{n_l}(Q_X)$ such that

$$|\mathcal{D}_{n_l}| \doteq 2^{n_l \tilde{\mathsf{R}}}. \tag{42}$$

Then, Lemma 4 implies that for every $\delta > 0$ and $l \geq l_0(\delta, |\mathcal{X}|)$ both

$$\kappa^*_{n_l} \geq \frac{2^{n_l[H(Q_X)-\delta]}}{2^{n_l(\tilde{\mathsf{R}}+\delta)}} \tag{43}$$

$$= 2^{n_l\left[H(Q_X)-\tilde{\mathsf{R}}-2\delta\right]} \tag{44}$$

from (40) and

$$\kappa^*_{n_l} \leq \frac{2^{n_l H(Q_X)}}{2^{n_l(\tilde{\mathsf{R}}-\delta)}} n_l \left[H(Q_X)+\delta\right] \tag{45}$$

$$\leq 2^{n_l\left[H(Q_X)-\tilde{\mathsf{R}}+2\delta\right]} \tag{46}$$

from Lemma 4. Thus, the cover is asymptotically efficient, and this implies that the permuted sets cannot overlap too much. To further explore this property, let $\{\pi_{n_l,t}\}_{t=0}^{\kappa^*_{n_l}}$ be the permutations constructed in Lemma 4 for block-length n_l, and define the *exclusive permutations sets* as

$$\mathcal{G}_{n_l,t} \triangleq \pi_{n_l,t}(\mathcal{D}_{n_l}) \backslash \left\{ \bigcup_{s=0}^{t-1} \pi_{n_l,s}(\mathcal{D}_{n_l}) \right\}. \tag{47}$$

Note that $\mathcal{T}_{n_l}(Q_X)$ is a disjoint union $\mathcal{G}_{n_l,t}$, and for any $\overline{\mathsf{R}} < \tilde{\mathsf{R}}$, consider the union of exclusive permutations sets of small cardinality, namely

$$\mathcal{H}(\overline{\mathsf{R}}) \triangleq \bigcup_{t:|\mathcal{G}_{n_l,t}| \leq 2^{n\overline{\mathsf{R}}}} \mathcal{G}_{n_l,t}. \tag{48}$$

A simple aspect of the asymptotic efficiency of the covering is that under the uniform distribution on the type class, the probability that the source block belongs to a small exclusive permutations set is also small.

Lemma 5. *For any* $\overline{\mathsf{R}} \leq \tilde{\mathsf{R}}$

$$\mathbb{P}\left[\mathbf{X} \in \mathcal{H}(\overline{\mathsf{R}}) | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \stackrel{\cdot}{\leq} 2^{-n(\tilde{\mathsf{R}}-\overline{\mathsf{R}})} \tag{49}$$

Proof: Let an arbitrary $\delta > 0$ be given. For all n sufficiently large, if $\mathcal{T}_n(Q_X)$ is empty then the statement of the lemma is satisfied by convention. Otherwise,

$$\mathbb{P}\left[\mathbf{X} \in \mathcal{H}(\overline{\mathsf{R}}) | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \leq \frac{\kappa^*_n \cdot e^{n\overline{\mathsf{R}}}}{|\mathcal{T}_n(Q_X)|} \tag{50}$$

$$\leq \frac{2^{n\left[H(Q_X)-\tilde{\mathsf{R}}+2\delta\right]} \cdot e^{n\overline{\mathsf{R}}}}{2^{n[H(Q_X)-\delta]}} \tag{51}$$

$$= 2^{n(\overline{\mathsf{R}}-\tilde{\mathsf{R}}+3\delta)}. \tag{52}$$

■

C. Proof of Achievability Part of Theorem 1

We follow the three steps outlined in Section V. In the first step of the proof, we focus on a single cryptogram, $\mathcal{C}_n(y) = \left\{\varphi_n(y, \mathbf{u}) : \mathbf{u} \in \{0,1\}^{n\mathsf{R}}\right\}$, which we generically denote by the set $\mathcal{C}_n = \{\mathbf{w}(0), \ldots, \mathbf{w}(2^{n\mathsf{R}} - 1)\} \subset \mathcal{W}^n$. We begin with some definitions and simple properties. For a given $(\mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}})$ and $Q_X \in \mathcal{P}_n(\mathcal{X})$, let $\tilde{\mathbf{X}}$ be uniformly distributed over $\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}})$ (defined in (25)). The exiguous-distortion probability for the set $\mathcal{C}_n$ is defined as[8]

$$p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}}) \triangleq \max_{\mathbf{z} \in \mathcal{Z}^n} \mathbb{P}\left[d_{\text{E}}(\tilde{\mathbf{X}}, \mathbf{z}) \le \mathsf{D}_{\text{E}}\right]. \tag{53}$$

We have the following simple properties for $p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}})$.

Proposition 6. *Let $\mathcal{C}_n \subset \mathcal{W}^n$ and $Q_X \in \mathcal{P}_n(\mathcal{X})$ be given. Then:*

1) *For every permutation π*

$$p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}}) = p_d(\pi(\mathcal{C}_n), Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}}), \tag{54}$$

where $\pi(\mathcal{C}_n)$ is as defined in (22)*.*

2) *Let $\overline{\mathbf{X}}$ be uniformly distributed over $\mathcal{D}_n \subseteq \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}})$. Then,*

$$\max_{\mathbf{z} \in \mathcal{Z}^n} \mathbb{P}\left[d_{\text{E}}(\overline{\mathbf{X}}, \mathbf{z}) \le \mathsf{D}_{\text{E}}\right] \le \frac{|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}})|}{|\mathcal{D}_n|} \cdot p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}}). \tag{55}$$

Proof:

1) Let $\mathbf{z}^*$ be the maximizer of (53). Since $d_{\text{L}}(\mathbf{x}, \mathbf{w}) = d_{\text{L}}(\pi(\mathbf{x}), \pi(\mathbf{w}))$ then $\mathfrak{D}(\pi(\mathcal{C}_n), Q_X, \mathsf{D}_{\text{L}}) = \pi\left(\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}})\right)$. Since also $d_{\text{E}}(\mathbf{x}, \mathbf{z}) = d_{\text{E}}(\pi(\mathbf{x}), \pi(\mathbf{z}))$ then

$$\begin{aligned} p_d\left[\pi(\mathcal{C}_n), Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}}\right] &= \max_{\mathbf{z} \in \mathcal{Z}^n} \mathbb{P}\left[d_{\text{E}}(\pi(\tilde{\mathbf{X}}), \mathbf{z}) \le \mathsf{D}_{\text{E}}\right] && (56)\\ &\ge \mathbb{P}\left[d_{\text{E}}(\pi(\tilde{\mathbf{X}}), \pi(\mathbf{z}^*)) \le \mathsf{D}_{\text{E}}\right] && (57)\\ &= p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}}), && (58) \end{aligned}$$

and the reverse inequality can be obtained similarly, by considering the inverse permutation π^{-1}.

2) For every $\mathbf{z} \in \mathcal{Z}^n$

$$\begin{aligned} \mathbb{P}\left[d_{\text{E}}(\overline{\mathbf{X}}, \mathbf{z}) \le \mathsf{D}_{\text{E}}\right] &= \frac{|\overline{\mathbf{x}} \in \mathcal{D}_n : d_{\text{E}}(\overline{\mathbf{x}}, \mathbf{z}) \le \mathsf{D}_{\text{E}}|}{|\mathcal{D}_n|} && (59)\\ &\le \frac{|\overline{\mathbf{x}} \in \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}) : d_{\text{E}}(\overline{\mathbf{x}}, \mathbf{z}) \le \mathsf{D}_{\text{E}}|}{|\mathcal{D}_n|} && (60)\\ &= \frac{|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}})|}{|\mathcal{D}_n|} \cdot \frac{|\overline{\mathbf{x}} \in \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}) : d_{\text{E}}(\overline{\mathbf{x}}, \mathbf{z}) \le \mathsf{D}_{\text{E}}|}{|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}})|} && (61)\\ &\le \frac{|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}})|}{|\mathcal{D}_n|} \cdot p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\text{L}}, \mathsf{D}_{\text{E}}). && (62) \end{aligned}$$

[8]With a slight abuse of notation, we also use here the notation $p_d(\cdot)$.

■

The next lemma is the first step in the proof, in which we prove the existence of a good set $\mathcal{C}_n^*$ by a random selection.

Lemma 7. *Let $\delta > 0$ and $Q_X \in \mathcal{P}(\mathcal{X})$ be given, and let n_l be the sequence of block-lengths such that $\mathcal{T}_{n_l}(Q_X)$ is non-empty. There exists a sequence of sets $\mathcal{C}^* = \{\mathcal{C}_{n_l}^*\}$ of size $|\mathcal{C}_{n_l}^*| = 2^{n_l \mathsf{R}}$ such that for all l sufficiently large*

$$\frac{1}{n_l} \log |\mathfrak{D}(\mathcal{C}_{n_l}^*, Q_X, \mathsf{D}_{\mathrm{L}})| \geq H(Q_X) + \mathsf{R} - R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}}) - \delta, \tag{63}$$

and

$$-\frac{1}{n_l} \log \max_{\mathbf{z} \in \mathcal{Z}^{n_l}} \mathbb{P}\left[d_{\mathrm{E}}(\tilde{\mathbf{X}}, \mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}\right] \geq \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}})\right\} - \delta, \tag{64}$$

for all $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$, where $\tilde{\mathbf{X}}$ is distributed uniformly over $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$.

Proof: Let n be given such that $\mathcal{T}_n(Q_X)$ is non-empty. Also, let D_{E} be given, choose any $Q_W \in \mathcal{P}_n(\mathcal{W})$, and consider an ensemble of randomly chosen sets $\mathcal{C}_n$, where each member is selected independently at random, uniformly within a type class $\mathcal{T}_n(Q_W)$. By definition, for any given $\mathcal{C}_n$

$$p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}, \mathsf{D}_{\mathrm{E}}) = \frac{\max_{\mathbf{z} \in \mathcal{Z}^n} |\{\mathbf{x} \in \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}) : d_{\mathrm{E}}(\mathbf{x}, \mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}\}|}{|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})|}. \tag{65}$$

It should be noticed, that unlike the situation in standard random coding bounds, here the denominator of (65) is also a random variable. Nonetheless, we will show that there exists a set $\mathcal{C}_n$ such that both the numerator and denominator of (65) are close to their expected values. To begin, let us analyze the expected value of the size of the D-cover in the denominator of (65). We first consider the case $\mathsf{R} \leq R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})$. For a given $\mathcal{C}_n$ and Q_{XW}, define the *type class enumerator*

$$N(Q_{XW}|\mathbf{x}) \triangleq \left|\left\{\mathbf{w} \in \mathcal{C}_n : \hat{Q}_{\mathbf{xw}} = Q_{XW}\right\}\right|, \tag{66}$$

and let

$$E_0 \triangleq H(Q_X) + \mathsf{R} - R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}}). \tag{67}$$

Note that in the last equation the X-marginal (W-marginal) of Q is constrained to the given type Q_X (respectively, Q_W). For brevity, here and throughout the sequel, such constraints will be omitted. Then,

$$\mathbb{E}[|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})|] = \mathbb{E}\left[\sum_{\mathbf{x} \in \mathcal{T}_n(Q_X)} \mathbb{I}\left\{\exists \mathbf{w} \in \mathcal{C}_n : d_{\mathrm{L}}(\mathbf{x}, \mathbf{w}) \leq \mathsf{D}_{\mathrm{L}}\right\}\right] \tag{68}$$

$$= \mathbb{E}\left[\sum_{\mathbf{x} \in \mathcal{T}_n(Q_X)} \mathbb{I}\left\{\bigcup_{Q_{XW} : \mathbb{E}_Q[d_{\mathrm{L}}(X,W)] \leq \mathsf{D}_{\mathrm{L}}} \{N(Q_{XW}|\mathbf{x}) \geq 1\}\right\}\right] \tag{69}$$

$$\doteq \mathbb{E}\left[\sum_{\mathbf{x} \in \mathcal{T}_n(Q_X)} \sum_{Q_{XW} : \mathbb{E}_Q[d_{\mathrm{L}}(X,W)] \leq \mathsf{D}_{\mathrm{L}}} \mathbb{I}\{N(Q_{XW}|\mathbf{x}) \geq 1\}\right] \tag{70}$$

$$= \sum_{\mathbf{x}\in\mathcal{T}_n(Q_X)} \sum_{Q_{XW}:\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}} \mathbb{P}\left\{N(Q_{XW}|\mathbf{x})\geq 1\right\} \tag{71}$$

$$\stackrel{(a)}{=} \sum_{\mathbf{x}\in\mathcal{T}_n(Q_X)} \sum_{Q_{XW}:\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}, I_Q(X;W)>\mathsf{R}} \mathbb{P}\left\{N(Q_{XW}|\mathbf{x})\geq 1\right\} \tag{72}$$

$$\stackrel{(b)}{\doteq} \sum_{\mathbf{x}\in\mathcal{T}_n(Q_X)} \sum_{Q_{XW}:\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}, I_Q(X;W)>\mathsf{R}} 2^{n[\mathsf{R}-I_Q(X;W)]} \tag{73}$$

$$\doteq 2^{nH_Q(X)} \max_{Q_{XW}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{W}):\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}, I_Q(X;W)>\mathsf{R}} 2^{n[\mathsf{R}-I_Q(X;W)]} \tag{74}$$

$$\stackrel{(c)}{=} \exp\left\{n\cdot\left[H_Q(X)+\mathsf{R}-\min_{Q_{XW}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{W}):\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}} I_Q(X;W)\right]\right\} \tag{75}$$

$$\stackrel{(d)}{=} 2^{nE_0}, \tag{76}$$

where in (a) and (c) we have used the assumption $\mathsf{R}\leq R_{\mathsf{L}}(Q_X,\mathsf{D}_{\mathsf{L}})$, and so, the set $\{Q_{XW}:\mathbb{E}_Q\left[d_{\mathsf{L}}(X,W)\right]\leq \mathsf{D}_{\mathsf{L}}, I_Q(X;W)\leq \mathsf{R}\}$ is empty. In (b), we have used the fact that $N(Q_{XW}|\mathbf{x})$ is a binomial random variable pertaining to $2^{n\mathsf{R}}$ trials and probability of success of exponential order $\exp\left[-nI_Q(X;W)\right]$. Passage (d) follows from the fact that $\mathcal{P}(\mathcal{X}\times\mathcal{W})$ is dense in $\mathcal{Q}(\mathcal{X}\times\mathcal{W})$ and $I_Q(X;W)$ is continuous. In addition, using the union bound, with probability 1,

$$|\mathfrak{D}(\mathcal{C}_n,Q_X,\mathsf{D}_{\mathsf{L}})| \leq \sum_{\mathbf{w}\in\mathcal{C}_n} |\{\mathbf{x}\in\mathcal{T}_n(Q_X): d_{\mathsf{L}}(\mathbf{x},\mathbf{w})\leq \mathsf{D}_{\mathsf{L}}\}| \tag{77}$$

$$\stackrel{\cdot}{\leq} 2^{n\mathsf{R}}\cdot\exp\left[n\cdot\max_{Q_{XW}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{W}):\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}} H_Q(X|W)\right] \tag{78}$$

$$= 2^{nE_0}. \tag{79}$$

Next, we upper bound the numerator of (65). For a given $\mathcal{C}_n$ and $\mathbf{z}\in\mathcal{Z}^n$, define now the type class enumerator

$$N(Q_{ZW}|\mathbf{z}) \triangleq \left|\left\{\mathbf{w}\in\mathcal{C}_n : \hat{Q}_{\mathbf{zw}} = Q_{ZW}\right\}\right|. \tag{80}$$

Then,

$$|\{\mathbf{x}\in\mathfrak{D}(\mathcal{C}_n,Q_X,\mathsf{D}_{\mathsf{L}}): d_{\mathsf{E}}(\mathbf{x},\mathbf{z})\leq \mathsf{D}_{\mathsf{E}}\}| \tag{81}$$

$$= \left|\bigcup_{\mathbf{w}\in\mathcal{C}_n}\{\mathbf{x}\in\mathcal{T}_n(Q_X): d_{\mathsf{E}}(\mathbf{x},\mathbf{z})\leq \mathsf{D}_{\mathsf{E}}, d_{\mathsf{L}}(\mathbf{x},\mathbf{w})\leq \mathsf{D}_{\mathsf{L}}\}\right| \tag{82}$$

$$= \left|\bigcup_{Q_{ZW}}\bigcup_{\mathbf{w}\in\mathcal{T}_n(Q_{W|Z},\mathbf{z})\cap\mathcal{C}_n}\bigcup_{Q_{X|ZW}:\mathbb{E}_Q[d_{\mathsf{E}}(X,Z)]\leq \mathsf{D}_{\mathsf{E}},\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}}\left\{\mathbf{x}\in\mathcal{T}_n(Q_{X|ZW},\mathbf{z},\mathbf{w})\right\}\right| \tag{83}$$

$$\stackrel{(a)}{\leq} \sum_{Q_{ZW}}\sum_{\mathbf{w}\in\mathcal{T}_n(Q_{W|Z},\mathbf{z})\cap\mathcal{C}_n}\sum_{Q_{X|ZW}:\mathbb{E}_Q[d_{\mathsf{E}}(X,Z)]\leq \mathsf{D}_{\mathsf{E}},\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}}\left|\left\{\mathbf{x}\in\mathcal{T}_n(Q_{X|ZW},\mathbf{z},\mathbf{w})\right\}\right| \tag{84}$$

$$\doteq \sum_{Q_{ZW}}\sum_{\mathbf{w}\in\mathcal{T}_n(Q_{W|Z},\mathbf{z})\cap\mathcal{C}_n}\sum_{Q_{X|ZW}:\mathbb{E}_Q[d_{\mathsf{E}}(X,Z)]\leq \mathsf{D}_{\mathsf{E}},\mathbb{E}_Q[d_{\mathsf{L}}(X,W)]\leq \mathsf{D}_{\mathsf{L}}} 2^{nH_Q(X|ZW)} \tag{85}$$

$$\doteq \sum_{Q_{ZW}} \sum_{\mathbf{w}\in\mathcal{T}_n(Q_{W|Z},\mathbf{z})\cap\mathcal{C}_n} \max_{Q_{X|ZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\le\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D}_{\mathrm{L}}} 2^{nH_Q(X|ZW)} \tag{86}$$

$$= \sum_{Q_{ZW}} N(Q_{ZW}|\mathbf{z}) \max_{Q_{X|ZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\le\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D}_{\mathrm{L}}} 2^{nH_Q(X|ZW)} \tag{87}$$

$$\doteq \max_{Q_{ZW}} \max_{Q_{X|ZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\le\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D}_{\mathrm{L}}} N(Q_{ZW}|\mathbf{z})2^{nH_Q(X|ZW)} \tag{88}$$

$$\doteq \sum_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\le\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D}_{\mathrm{L}}} N(Q_{ZW}|\mathbf{z})2^{nH_Q(X|ZW)} \tag{89}$$

where (a) is the union bound, and in all the above equations, $Q_{XZW} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Z}\times\mathcal{W})$. Let

$$\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}}) \triangleq \{Q_{XZW}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{Z}\times\mathcal{W}) : \mathbb{E}_Q\left[d_{\mathrm{E}}(X,Z)\right]\le\mathsf{D}_{\mathrm{E}}, \mathbb{E}_Q\left[d_{\mathrm{L}}(X,W)\right]\le\mathsf{D}_{\mathrm{L}}\}. \tag{90}$$

Taking expectation, and using the fact that $|\mathcal{P}_n(\mathcal{X}\times\mathcal{Z}\times\mathcal{W})| \le (n+1)^{|\mathcal{X}||\mathcal{Z}||\mathcal{W}|}$ i.e., increases with n only polynomially,

$$\mathbb{E}\left[\max_{\mathbf{z}\in\mathcal{Z}^n} |\{\mathbf{x}\in\mathfrak{D}(\mathcal{C}_n,Q_X,\mathsf{D}_{\mathrm{L}}) : d_{\mathrm{E}}(\mathbf{x},\mathbf{z})\le\mathsf{D}_{\mathrm{E}}\}|\right] \tag{91}$$

$$\dot{\le} \mathbb{E}\left[\max_{\mathbf{z}\in\mathcal{Z}^n} \sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}})} N(Q_{ZW}|\mathbf{z})2^{nH_Q(X|ZW)}\right] \tag{92}$$

$$= \mathbb{E}\left[\lim_{\beta\to\infty}\left\{\sum_{\mathbf{z}\in\mathcal{Z}^n}\left(\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}})} N(Q_{ZW}|\mathbf{z})2^{nH_Q(X|ZW)}\right)^{\beta}\right\}^{1/\beta}\right] \tag{93}$$

$$\overset{(a)}{=} \lim_{\beta\to\infty}\mathbb{E}\left[\left\{\sum_{\mathbf{z}\in\mathcal{Z}^n}\left(\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}})} N(Q_{ZW}|\mathbf{z})2^{nH_Q(X|ZW)}\right)^{\beta}\right\}^{1/\beta}\right] \tag{94}$$

$$\doteq \lim_{\beta\to\infty}\mathbb{E}\left[\left\{\sum_{\mathbf{z}\in\mathcal{Z}^n}\left(\max_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}})} N(Q_{ZW}|\mathbf{z})2^{nH_Q(X|ZW)}\right)^{\beta}\right\}^{1/\beta}\right] \tag{95}$$

$$= \lim_{\beta\to\infty}\mathbb{E}\left[\left(\sum_{\mathbf{z}\in\mathcal{Z}^n}\max_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}})} N(Q_{ZW}|\mathbf{z})^{\beta}2^{n\beta H_Q(X|ZW)}\right)^{1/\beta}\right] \tag{96}$$

$$\doteq \lim_{\beta\to\infty}\mathbb{E}\left[\left(\sum_{\mathbf{z}\in\mathcal{Z}^n}\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}})} N(Q_{ZW}|\mathbf{z})^{\beta}2^{n\beta H_Q(X|ZW)}\right)^{1/\beta}\right] \tag{97}$$

$$\overset{(b)}{\le} \lim_{\beta\to\infty}\left(\sum_{\mathbf{z}\in\mathcal{Z}^n}\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}})} \mathbb{E}\left[N(Q_{ZW}|\mathbf{z})^{\beta}\right]2^{n\beta H_Q(X|ZW)}\right)^{1/\beta} \tag{98}$$

$$= \lim_{\beta\to\infty}\Bigg(\sum_{\mathbf{z}\in\mathcal{Z}^n}\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}}):I_Q(Z;W)\le\mathsf{R}} \mathbb{E}\left[N(Q_{ZW}|\mathbf{z})^{\beta}\right]2^{n\beta H_Q(X|ZW)}$$
$$+ \sum_{\mathbf{z}\in\mathcal{Z}^n}\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D}_{\mathrm{L}},\mathsf{D}_{\mathrm{E}}):I_Q(Z;W)>\mathsf{R}} \mathbb{E}\left[N(Q_{ZW}|\mathbf{z})^{\beta}\right]2^{n\beta H_Q(X|ZW)}\Bigg)^{1/\beta} \tag{99}$$

$$\stackrel{(c)}{\doteq} \lim_{\beta\to\infty}\Bigg(\sum_{\mathbf{z}\in\mathcal{Z}^n}\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):Q_Z=\hat{Q}_{\mathbf{z}},I_Q(Z;W)\le\mathsf{R}} 2^{n\beta[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)} + \sum_{\mathbf{z}\in\mathcal{Z}^n}\sum_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):Q_Z=\hat{Q}_{\mathbf{z}},I_Q(Z;W)>\mathsf{R}} 2^{n[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)}\Bigg)^{1/\beta} \tag{100}$$

$$\doteq \lim_{\beta\to\infty}\Bigg(\sum_{Q_Z}2^{nH_Q(Z)}\sum_{Q_{XW|Z}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\le\mathsf{D_E},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D_L},I_Q(Z;W)\le\mathsf{R}} 2^{n\beta[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)} + \sum_{Q_Z}2^{nH_Q(Z)}\sum_{Q_{XW|Z}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\le\mathsf{D_E},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D_L},I_Q(Z;W)>\mathsf{R}} 2^{n[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)}\Bigg)^{1/\beta} \tag{101}$$

$$\doteq \lim_{\beta\to\infty}\Bigg(\max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)\le\mathsf{R}} 2^{nH_Q(Z)}2^{n\beta[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)} + \max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)>\mathsf{R}} 2^{nH_Q(Z)}2^{n[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)}\Bigg)^{1/\beta} \tag{102}$$

$$\doteq \lim_{\beta\to\infty}\Bigg(\max\Bigg\{\max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)\le\mathsf{R}} 2^{nH_Q(Z)}2^{n\beta[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)}, \max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)>\mathsf{R}} 2^{nH_Q(Z)}2^{n[\mathsf{R}-I_Q(Z;W)]}2^{n\beta H_Q(X|ZW)}\Bigg\}\Bigg)^{1/\beta} \tag{103}$$

$$= \lim_{\beta\to\infty}\max\Bigg\{\max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)\le\mathsf{R}} 2^{n\frac{1}{\beta}H_Q(Z)}2^{n[\mathsf{R}-I_Q(Z;W)]}2^{nH_Q(X|ZW)}, \max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)>\mathsf{R}} 2^{n\frac{1}{\beta}H_Q(Z)}2^{n\frac{1}{\beta}[\mathsf{R}-I_Q(Z;W)]}2^{nH_Q(X|ZW)}\Bigg\} \tag{104}$$

$$= \max\Bigg\{\max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)\le\mathsf{R}} 2^{n[\mathsf{R}-I_Q(Z;W)]}2^{nH_Q(X|ZW)}, \max_{Q_{XZW}\in\mathcal{J}(\mathsf{D_L},\mathsf{D_E}):I_Q(Z;W)>\mathsf{R}} 2^{nH_Q(X|ZW)}\Bigg\} \tag{105}$$

where (a) is by the Lebesgue monotone convergence theorem [30, Theorem 11.28] and the monotonicity of the argument inside the expectation operator in β, and (b) is by the Jensen inequality. In (c), we have used the analysis in [31, Subsection 6.3] of the moments of $N(Q_{ZW}|\mathbf{z})$, which is a binomial random variable with $2^{n\mathsf{R}}$ trials and probability of success of the exponential order of $\exp\left[-nI_Q(Z;W)\right]$. Also, note that in all the above equations, $Q_{XZW}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{Z}\times\mathcal{W})$ but since $\mathcal{P}(\mathcal{X}\times\mathcal{Z}\times\mathcal{W})$ is dense in $\mathcal{Q}(\mathcal{X}\times\mathcal{Z}\times\mathcal{W})$ and the arguments of the maximization are continuous functions of Q_{XZW}, we can change the maximization to be over $\mathcal{Q}(\mathcal{X}\times\mathcal{Z}\times\mathcal{W})$. Thus,

$$\mathbb{E}\left[\max_{\mathbf{z}\in\mathcal{Z}^n}\left|\{\mathbf{x}\in\mathfrak{D}(\mathcal{C}_n,Q_X,\mathsf{D_L}):d_{\mathrm{E}}(\mathbf{x},\mathbf{z})\le\mathsf{D_E}\}\right|\right]\stackrel{\cdot}{\le}2^{nE_1(\mathsf{D_E})} \tag{106}$$

where

$$E_1(\mathsf{D_E})\triangleq\max_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\le\mathsf{D_E},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\le\mathsf{D_L}}\left\{H_Q(X|ZW)+\left[\mathsf{R}-I_Q(Z;W)\right]_+\right\}. \tag{107}$$

Now, let $\delta>0$ be given. There exists $n_0(Q_X)$ such that for all $n\ge n_0(Q_X)$, we have from (76)

$$\mathbb{E}\left(\left|\mathfrak{D}(\mathcal{C}_n,Q_X,\mathsf{D_L})\right|\right)\ge 2^{n(E_0-\frac{\delta}{2})}, \tag{108}$$

and from (79)

$$|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D_L})| \leq 2^{n(E_0+\frac{\delta}{2})}. \tag{109}$$

Define, for the given ensemble of the random sets

$$\mathcal{A}_0 \triangleq \left\{\mathcal{C}_n : |\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D_L})| > 2^{-n\frac{\delta}{2}} \mathbb{E}[|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D_L})|]\right\}. \tag{110}$$

The reverse Markov lemma (Lemma 2) implies

$$\mathbb{P}(\mathcal{A}_0) \geq \frac{1 - 2^{-n\frac{\delta}{2}}}{2^{n\delta} - 2^{-n\frac{\delta}{2}}} \geq 2^{-2n\delta} \tag{111}$$

where the second inequality is satisfied for all $n \geq n_0'$ for some $n_0' \geq n_0(Q_X)$.

Now, note that we need to prove that a single set $\mathcal{C}_n^*$ satisfies (64) for all $\mathsf{D_E} \geq \mathsf{D_L}$. To show this, we consider a quantization of the possible values of $\mathsf{D_E}$. To this end, let an arbitrary $\eta > 0$ be given, such that $J = \frac{R_{\mathrm{E}}(Q_X, \mathsf{D_L})}{\eta}$ is integer, and find $\overline{\mathsf{D}}_{\mathrm{E}}$ sufficiently large such that[9]

$$R_{\mathrm{E}}(Q_X, \overline{\mathsf{D}}_{\mathrm{E}}) \leq \lim_{\mathsf{D_E}\to\infty} R_{\mathrm{E}}(Q_X, \mathsf{D_E}) + \eta. \tag{112}$$

Let us quantize the interval $[R_{\mathrm{E}}(Q_X, \overline{\mathsf{D}}_{\mathrm{E}}), R_{\mathrm{E}}(Q_X, \mathsf{D_L})]$ to values $\{\mathsf{R}(0), \ldots, \mathsf{R}(J)\}$, where $\mathsf{R}(j) = j\eta$ and let $\mathsf{D_E}(j) = R_{\mathrm{E}}^{-1}(Q_X, \mathsf{R}(j))$, where $R_{\mathrm{E}}^{-1}(Q_X, \mathsf{R})$ is the inverse function of $R_{\mathrm{E}}(Q_X, \mathsf{D_E})$. By (105), there exists $n_1(j, Q_X)$ such that for all $n \geq n_1(j, Q_X)$

$$\mathbb{E}\left[\max_{\mathbf{z}\in\mathcal{Z}^n} |\{\mathbf{x} \in \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D_L}) : d_{\mathrm{E}}(\mathbf{x}, \mathbf{z}) \leq \mathsf{D_E}(j)\}|\right] \leq 2^{n[E_1(\mathsf{D_E}(j))+\delta]}, \tag{113}$$

where the expectation is over the random ensemble of sets $\mathcal{C}_n$. By defining

$$\mathcal{A}_{1j} \triangleq \left\{\mathcal{C}_n : \max_{\mathbf{z}\in\mathcal{Z}^n} |\{\mathbf{x} \in \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D_L}) : d_{\mathrm{E}}(\mathbf{x}, \mathbf{z}) \leq \mathsf{D_E}(j)\}| \leq 2^{n[E_1(\mathsf{D_E}(j))+4\delta]}\right\} \tag{114}$$

the ordinary Markov lemma implies

$$\mathbb{P}(\mathcal{A}_{1j}) \geq 1 - \frac{\mathbb{E}\left[\max_{\mathbf{z}\in\mathcal{Z}^n} |\{\mathbf{x} \in \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D_L}) : d_{\mathrm{E}}(\mathbf{x}, \mathbf{z}) \leq \mathsf{D_E}(j)\}|\right]}{2^{n[E_1(\mathsf{D_E}(j))+4\delta]}} \tag{115}$$

$$\geq 1 - 2^{-3n\delta}. \tag{116}$$

Defining $\mathcal{A}_1 \triangleq \bigcap_{j=0}^{J} \mathcal{A}_{1j}$ we get

$$\mathbb{P}(\mathcal{A}_1) = \mathbb{P}\left(\bigcap_{j=0}^{J} \mathcal{A}_{1j}\right) \tag{117}$$

$$= 1 - \mathbb{P}\left(\bigcup_{j=0}^{J} \mathcal{A}_{1j}^c\right) \tag{118}$$

[9]Note that if $d_{\mathrm{E}}(x, z) < \infty$ for all $x \in \mathcal{X}, z \in \mathcal{Z}$, then $\lim_{\mathsf{D_E}\to\infty} R_{\mathrm{E}}(Q_X, \mathsf{D_E}) = 0$.

$$\geq 1 - \sum_{j=0}^{J} \mathbb{P}\left(\mathcal{A}_{1j}^c\right) \tag{119}$$

$$\geq 1 - J \cdot 2^{-3n\delta}. \tag{120}$$

Thus, since J does not depend on n, there exists $n_1' \geq \max_{0\leq j\leq J} n_1(j, Q_X)$ such that for all $n \geq n_1'$

$$\mathbb{P}\left(\mathcal{A}_0 \cap \mathcal{A}_1\right) = 1 - \mathbb{P}\left(\mathcal{A}_0^c \cup \mathcal{A}_1^c\right) \tag{121}$$

$$\geq 1 - \mathbb{P}\left(\mathcal{A}_0^c\right) - \mathbb{P}\left(\mathcal{A}_1^c\right) \tag{122}$$

$$\geq 1 - (1 - 2^{-2n\delta}) - J2^{-n\frac{5\delta}{2}} \tag{123}$$

$$= 2^{-2n\delta} - J \cdot 2^{-3n\delta} \tag{124}$$

$$> 0. \tag{125}$$

Therefore, for all sufficiently large $n > \max\{n_0', n_1'\}$, there exists $\mathcal{C}_n \in \mathcal{A}_0 \cap \{\bigcap_{j=0}^{J} \mathcal{A}_{1j}\}$, i.e., $\mathcal{C}_n$ which satisfies both

$$|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})| > 2^{-n\frac{\delta}{2}} \mathbb{E}[|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})|] \tag{126}$$

and

$$\max_{\mathbf{z}\in\mathcal{Z}^n} |\{\mathbf{x} \in \mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}) : d_{\mathrm{E}}(\mathbf{x}, \mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}(j)\}| \leq 2^{4n\delta} 2^{nE_1(\mathsf{D}_{\mathrm{E}}(j))} \tag{127}$$

for all $0 \leq j \leq J$. Thus we get

$$p_d\left[\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}, \mathsf{D}_{\mathrm{E}}(j)\right] \leq \frac{2^{4n\delta} 2^{nE_1(\mathsf{D}_{\mathrm{E}}(j))}}{2^{-n\frac{\delta}{2}} 2^{n(E_0 - n\frac{\delta}{2})}} = 2^{5n\delta} \cdot 2^{n[E_1(\mathsf{D}_{\mathrm{E}}(j)) - E_0]}. \tag{128}$$

If we now define $E(\mathsf{D}_{\mathrm{E}}) \triangleq E_1(\mathsf{D}_{\mathrm{E}}) - E_0$, then for any given $Q_W \in \mathcal{P}_n(\mathcal{W})$

$$\liminf_{n\to\infty} -\frac{1}{n} \log p_d\left[\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}, \mathsf{D}_{\mathrm{E}}(j)\right] \geq E(\mathsf{D}_{\mathrm{E}}). \tag{129}$$

Now, choose let Q_W be the W-marginal of Q_{XW} which achieves $R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})$. Then,

$$E(\mathsf{D}_{\mathrm{E}}) \geq \min_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq\mathsf{D}_{\mathrm{L}},I_Q(Z;W)\leq\mathsf{R}} \{I_Q(Z;W) + I_Q(X;Z,W)\}$$
$$- \min_{Q_{XW}:\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq\mathsf{D}_{\mathrm{L}}} I_Q(X;W) \tag{130}$$

$$\overset{(a)}{\geq} \min_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq\mathsf{D}_{\mathrm{L}},I_Q(Z;W)\leq\mathsf{R}} \{I_Q(Z;W) + I_Q(X;Z,W) - I_Q(X;W)\} \tag{131}$$

$$= \min_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq\mathsf{D}_{\mathrm{L}},I_Q(Z;W)\leq\mathsf{R}} \{I_Q(Z;W) + I_Q(X;Z|W)\} \tag{132}$$

$$= \min_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq\mathsf{D}_{\mathrm{E}},\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq\mathsf{D}_{\mathrm{L}},I_Q(Z;W)\leq\mathsf{R}} I_Q(X,W;Z) \tag{133}$$

$$\overset{(b)}{\geq} \min_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq\mathsf{D}_{\mathrm{E}}} I_Q(X;Z) \tag{134}$$

$$= R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}}) \tag{135}$$

where (a) is by restricting Q_{XW} to be the same in both minimizations of (130), and (b) is by the data processing property of the mutual information. Similarly,

$$\begin{aligned} E(\mathsf{D}_{\mathrm{E}}) &\geq \mathsf{R} + \min_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}, \mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq \mathsf{D}_{\mathrm{L}}, I_Q(Z;W)>\mathsf{R}} I_Q(X;Z,W) \\ &\quad - \min_{Q_{XW}:\mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq \mathsf{D}_{\mathrm{L}}} I_Q(X;W) && (136) \\ &\geq \mathsf{R} + \min_{Q_{XZW}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}, \mathbb{E}_Q[d_{\mathrm{L}}(X,W)]\leq \mathsf{D}_{\mathrm{L}}, I_Q(Z;W)>\mathsf{R}} I_Q(X;Z|W) && (137) \\ &\geq \mathsf{R}. && (138) \end{aligned}$$

by restricting Q_{XW} to be the same in both minimizations of (136).

Therefore, (129), (135) and (136) imply that

$$\liminf_{n\to\infty} -\frac{1}{n}\log p_d\left[\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}, \mathsf{D}_{\mathrm{E}}(j)\right] \geq \min\left\{R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}}(j)), \mathsf{R}\right\} \tag{139}$$

for all $0 \leq j \leq J$. By taking $\eta \downarrow 0$, continuity of $R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}})$ in D_{E} provides the lower bound (64) for all $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$. Then, (63) is obtained from (126) and (108).

To complete the proof of the lemma, we consider the case of $\mathsf{R} \geq R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})$. Denote by $Q_{XW}^{(n)}$ a sequence of distributions such that $Q_{XW}^{(n)} \to Q_{XW}^*$ as $n \to \infty$, where Q_{XW}^* achieves the rate-distortion function $R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})$. For a given $\mathcal{C}_n$, let $\tilde{\mathcal{C}}_n$ be a subset formed by the first $e^{nR_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})}$ members of $\mathcal{C}_n$. The same analysis as before shows that when randomly drawing a set $\mathcal{C}_n$ uniformly over the W-marginal of $Q_{XW}^{(n)}$, there exists a sequence of sets $\{\mathcal{C}_n\}$ such that

$$|\mathfrak{D}(\tilde{\mathcal{C}}_n, Q_X, \mathsf{D}_{\mathrm{L}})| \geq 2^{n(E_0-\delta)} \geq 2^{n[H(Q_X)-\delta]}. \tag{140}$$

Then, for $\mathcal{C}_n$

$$\begin{aligned} p_d(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}, \mathsf{D}_{\mathrm{E}}) &= \frac{\max_{\mathbf{z}\in\mathcal{Z}^n}\left|\{\mathbf{x}\in\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}) : d_{\mathrm{E}}(\mathbf{x},\mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}\}\right|}{|\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}})|} && (141) \\ &\leq \frac{\max_{\mathbf{z}\in\mathcal{Z}^n}\left|\{\mathbf{x}\in\mathfrak{D}(\mathcal{C}_n, Q_X, \mathsf{D}_{\mathrm{L}}) : d_{\mathrm{E}}(\mathbf{x},\mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}\}\right|}{|\mathfrak{D}(\tilde{\mathcal{C}}_n, Q_X, \mathsf{D}_{\mathrm{L}})|} && (142) \\ &\leq \frac{\max_{\mathbf{z}\in\mathcal{Z}^n}\left|\{\mathbf{x}\in\mathcal{T}_n(Q_X) : d_{\mathrm{E}}(\mathbf{x},\mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}\}\right|}{|\mathfrak{D}(\tilde{\mathcal{C}}_n, Q_X, \mathsf{D}_{\mathrm{L}})|} && (143) \\ &\leq \frac{\max_{\mathbf{z}\in\mathcal{Z}^n}\left|\{\mathbf{x}\in\mathcal{T}_n(Q_X) : d_{\mathrm{E}}(\mathbf{x},\mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}\}\right|}{2^{n[H(Q_X)-\delta]}} && (144) \\ &= 2^{-n[H(Q_X)-\delta]} \max_{\mathbf{z}\in\mathcal{Z}^n} \sum_{Q_{X|Z}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}} \left|\mathcal{T}_n(Q_{X|Z}, \mathbf{z})\right| && (145) \\ &\leq 2^{-n[H(Q_X)-\delta]} \max_{Q_Z} \sum_{Q_{X|Z}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}} 2^{nH_Q(X|Z)} && (146) \\ &\doteq \exp\left(-n\left[H_Q(X) - \delta - \max_{Q_{XZ}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}} H_Q(X|Z)\right]\right) && (147) \end{aligned}$$

$$\leq 2^{-n[R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})-\delta]} \tag{148}$$

and the proof of the lemma is complete, as δ is arbitrary. ■

Remark 8. As mentioned in Section V, to show achievability of an exiguous-distortion exponent using the method of types, *uniform* convergence of $-\frac{1}{n}\log p_d(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}}, \mathsf{D}_{\mathrm{E}})$ to the exponent $\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}})\}$ is required (cf. eq. (233)). However, the proof of Lemma 7 is not sufficient to show this. Specifically, the convergence in the asymptotic analysis of the type class enumerators, i.e. the relations

$$\mathbb{P}\left\{N(Q_{XW}|\mathbf{x}) \geq 1\right\} \doteq 2^{n[\mathsf{R}-I_Q(X;W)]} \tag{149}$$

used in (73) and

$$\mathbb{E}\left[N(Q_{ZW}|\mathbf{z})^{\beta}\right] \doteq \begin{cases} 2^{n[\mathsf{R}-I_Q(Z;W)]}, & I_Q(Z;W) \leq \mathsf{R} \\ 2^{n\beta[\mathsf{R}-I_Q(Z;W)]}, & I_Q(Z;W) > \mathsf{R} \end{cases} \tag{150}$$

used in (100), are not uniform in Q_X.

We continue with the second step of the proof, which constructs from the set $\mathcal{C}_n^*$ a secure rate-distortion code for all $\mathbf{x} \in \mathcal{T}_n(Q_X)$. The proof of the next lemma is based on the permutations technique described in Subsection VI-B.

Lemma 9. *For any given $Q_X \in \mathcal{P}(\mathcal{X}) \cap \operatorname{int}\mathcal{Q}(\mathcal{X})$ and $\delta > 0$, there exists a sequence of secure rate-distortion codes $\mathcal{S}^*$ of fixed key rate R such that*

$$\lim_{n\to\infty}\frac{1}{n}\log|\mathcal{Y}_n| \leq R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}}) + \delta, \tag{151}$$

and,

$$\mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \varphi_n^*(f_n^*(\mathbf{X}, \mathbf{u}))) \geq \mathsf{D}_{\mathrm{L}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] = 0 \tag{152}$$

for every $\mathbf{u} \in \{0,1\}^{n\mathsf{R}}$, as well as

$$\mathcal{E}_d^-(\mathcal{S}^*, \mathsf{D}_{\mathrm{E}}, Q_X) \geq \min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}})\} - \delta \tag{153}$$

for all $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$.

Proof: Assume that $Q_X \in [\operatorname{int}\mathcal{Q}(\mathcal{X})] \cap \mathcal{P}_{n_0}(\mathcal{X})$ for some minimal $n_0 \in \mathbb{N}$. Since the statements in the lemma are only about conditional events given the type Q_X, it is clear that the secure rate-distortion codes constructed $\mathcal{S}_n^*$, may only encode $\mathbf{x} \in \mathcal{T}_n(Q_X)$, and so only block-lengths $n \bmod n_0 = 0$ should be considered, as otherwise $\mathcal{T}_n(Q_X)$ is empty.

Let $\mathcal{C}^* = \{\mathcal{C}_n^*\}$ be a sequence of sets of size $2^{n\mathsf{R}}$ constructed according to Lemma 7. So for all n sufficiently

large

$$p_d(\mathcal{C}_n^*, Q_X, \mathsf{D_L}, \mathsf{D_E}) \leq 2^{-n[\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta]}, \tag{154}$$

and

$$|\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D_L})| \geq 2^{n(A-\delta)}, \tag{155}$$

where

$$A \triangleq \min\left\{H(Q_X) + \mathsf{R} - R_{\mathrm{L}}(Q_X, \mathsf{D_L}), H(Q_X)\right\}. \tag{156}$$

Now, let $\{\pi_{n,t}\}_{t=0}^{\kappa_n}$ be a set of permutations constructed according to Lemma 4, such that

$$\bigcup_{t=0}^{\kappa_n} \pi_{n,t}(\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D_L})) = \mathcal{T}_n(Q_X), \tag{157}$$

where $\kappa_n \leq 2^{n[H(Q_X)-A+2\delta)]}$, and let $\{\mathcal{G}_{n,t}\}$ be the resulting exclusive permutation sets, as defined in (47). We construct the following secure rate-distortion codes $\mathcal{S}_n^* = (f_n^*, \varphi_n^*)$ of fixed key rate R, which only encode $\mathbf{x} \in \mathcal{T}_n(Q_X)$. We utilize the covering of the type class $\mathcal{T}_n(Q_X)$ by permutations of a D-cover of the set $\mathcal{C}_n^*$ to encode the source block in the following way. Assume that the elements of $\mathcal{C}_n^*$ are arbitrarily ordered, i.e. $\mathcal{C}_n^* = \{\mathbf{w}(0), \ldots, \mathbf{w}(2^{n\mathsf{R}} - 1)\}$. For a given $\mathbf{x} \in \mathcal{T}_n(Q_X)$, let

$$t^*(\mathbf{x}) \triangleq \min\left\{t : \mathbf{x} \in \mathcal{G}_{n,t}\right\}, \tag{158}$$

and

$$i^*(\mathbf{x}) \triangleq \min\{i : \mathbf{w}(i) \in \mathcal{G}_{n,t^*(\mathbf{x})}, d_{\mathrm{L}}(\mathbf{x}, \mathbf{w}(i)) \leq \mathsf{D_L}\} \tag{159}$$

The encoding is a concatenation of the following two parts $y = f_n^*(\mathbf{x}, \mathbf{u}) = (t_y, i_y)$:

- A description of the permutation set, defined as $t_y \triangleq \mathbb{B}[t^*(\mathbf{x}); n(H(Q_X) - A + 2\delta)]$.
- An encrypted description of the distortion covering codeword, defined as $i_y \triangleq \mathbb{B}[i^*(\mathbf{x}); n\mathsf{R}] \oplus \mathbf{u}$.

It is easily verified that given $\mathbf{u}$, the legitimate decoder can reproduce $\mathbf{w} = \varphi_n(y, \mathbf{u})$ such that $d_{\mathrm{L}}(\mathbf{x}, \mathbf{w}) \leq \mathsf{D_L}$, for all $\mathbf{x} \in \mathcal{T}_n(Q_X)$, and so (152) is satisfied. Regarding the coding rate, note that

$$\frac{1}{n}\log|\mathcal{Y}_n| = H(Q_X) - A + 2\delta + \mathsf{R} \tag{160}$$

$$\leq R_{\mathrm{L}}(Q_X, \mathsf{D_L}) + 3\delta \tag{161}$$

for all n sufficiently large, which results in (151).

It remains to prove that for any eavesdropper σ_n, the conditional exiguous-distortion exponent, given that $\mathbf{X} \in \mathcal{T}_n(Q_X)$, is larger than $\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta$. From Proposition 3, it may be assumed that the eavesdropper is aware of the type Q_X. Moreover, given the cryptogram $Y = y$, the source block $\mathbf{X}$ is distributed uniformly over $\mathcal{G}_{n,t_y}$, and independent of i_y. Thus, the optimal eavesdropper has the same estimate for cryptograms with the same

t_y, and we may denote its estimate as $\mathbf{z} = \sigma_n(y) \triangleq \mathbf{z}(t_y)$. Since $\mathcal{G}_{n,0} = \mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D_L})$, then conditioned on the event $\{t^*(\mathbf{X}) = 0\}$, for any $\mathbf{z} \in \mathcal{Z}^n$, Lemma 7 implies

$$\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{T}_n(Q_X), t^*(\mathbf{X}) = 0\right] = \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{G}_{n,0}\right] \tag{162}$$

$$\leq 2^{-n[\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta]} \tag{163}$$

for all n sufficiently large. It then follows that for $0 < t \leq \kappa_n$,

$$\begin{aligned}\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{T}_n(Q_X), t^*(\mathbf{X}) = t\right] &= \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{G}_{n,t}\right] \\ &\overset{(a)}{\leq} \frac{|\mathcal{G}_{n,0}|}{|\mathcal{G}_{n,t}|} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{G}_{n,0}\right] \\ &\leq \frac{|\mathcal{G}_{n,0}|}{|\mathcal{G}_{n,t}|} 2^{-n(\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta)},\end{aligned} \tag{164}$$

where (a) follows from the fact that for any $0 < t \leq \kappa_n$, there exists a permutation π such that $\pi\left(\mathcal{G}_{n,t}\right) \subset \mathcal{G}_{n,0} = \mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D_L})$ and Proposition 6. Thus, the exiguous-distortion probability conditioned on $t^*(\mathbf{X}) = t$ can be larger than the same probability conditioned on $t^*(\mathbf{X}) = 0$, but only up to a factor of $\frac{|\mathcal{G}_{n,0}|}{|\mathcal{G}_{n,t}|}$, which is large if $|\mathcal{G}_{n,t}|$ is small. Next, we show that the contribution to the exiguous-distortion probability of these small sets does not impact its exponential behavior. To this end, for any fixed $0 < \eta < A + \delta$ such that $J = \frac{A+\delta}{\eta}$ is an integer, let us quantize the interval $[0, A + \delta]$ to values $\{A_0, \ldots, A_J\}$, where $A_j = j\eta$. We will treat separately sets such that $2^{nA_j} \leq |\mathcal{G}_{n,t}| \leq 2^{nA_{j+1}}$. For all n sufficiently large

$$\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{165}$$

$$= \sum_{t=0}^{\kappa_n} \mathbb{P}\left[\mathbf{X} \in \mathcal{G}_{n,t} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}(t)) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{G}_{n,t}, \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{166}$$

$$= \sum_{j=0}^{J-1} \sum_{t: 2^{nA_j} \leq |\mathcal{G}_{n,t}| \leq 2^{nA_{j+1}}} \mathbb{P}\left[\mathbf{X} \in \mathcal{G}_{n,t} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}(t)) \leq \mathsf{D_E} | \mathbf{X} \in \mathcal{G}_{n,t}, \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{167}$$

$$\overset{(a)}{\leq} \sum_{j=0}^{J-1} \sum_{t: 2^{nA_j} \leq |\mathcal{G}_{n,t}| \leq 2^{nA_{j+1}}} \mathbb{P}\left[\mathbf{X} \in \mathcal{G}_{n,t} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \frac{|\mathcal{G}_{n,0}|}{|\mathcal{G}_{n,t}|} 2^{-n(\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta)} \tag{168}$$

$$\leq \sum_{j=0}^{J-1} \sum_{t: 2^{nA_j} \leq |\mathcal{G}_{n,t}| \leq 2^{nA_{j+1}}} \mathbb{P}\left[\mathbf{X} \in \mathcal{G}_{n,t} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \frac{2^{n(A+\delta)}}{2^{nA_j}} 2^{-n(\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta)} \tag{169}$$

$$= \sum_{j=0}^{J-1} \frac{2^{n(A+\delta)}}{2^{nA_j}} 2^{-n(\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta)} \sum_{t: 2^{nA_j} \leq |\mathcal{G}_{n,t}| \leq 2^{nA_{j+1}}} \mathbb{P}\left[\mathbf{X} \in \mathcal{G}_{n,t} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{170}$$

$$\overset{(b)}{\leq} \sum_{j=0}^{J-1} \frac{2^{n(A+\delta)}}{2^{nA_j}} 2^{-n(\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta)} \mathbb{P}\left[\mathbf{X} \in \mathcal{H}(A_{j+1}) | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{171}$$

$$\overset{(c)}{\leq} \sum_{j=0}^{J-1} \frac{2^{n(A+\delta)}}{2^{nA_j}} 2^{-n(\min\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D_E})\} - \delta)} 2^{-n(A - A_{j+1} - \delta)} \tag{172}$$

$$\leq J \cdot \max_{0\leq j\leq J-1} \frac{2^{n(A_{j+1}+2\delta)}}{2^{nA_j}} 2^{-n(\min\{\mathsf{R},R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\}-\delta)} \tag{173}$$

$$\leq 2^{n(\eta+3\delta)} 2^{-n\cdot\min\{\mathsf{R},R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\}} \tag{174}$$

$$\overset{(d)}{\leq} 2^{n(\eta+4\delta)} 2^{-n\cdot\min\{\mathsf{R},R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\}} \tag{175}$$

where (a) is using (164), (b) is using the definition in (48), (c) is using Lemma 5, and (d) is since $J \doteq 1$. The result follows by taking $\eta \downarrow 0$. ■

Remark 10. Note that only the properties (154)-(155) of $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$ were used in order to prove Lemma 9. The same proof of Lemma 9 can be used to show that if some other set $\mathcal{D}_n \subset \mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$ satisfies similar properties, i.e. if for some $E > 0$

$$\max_{\mathbf{z}\in\mathcal{Z}^n} \mathbb{P}\left[d_{\mathrm{E}}(\tilde{\mathbf{X}},\mathbf{z}) \leq \mathsf{D}_{\mathrm{E}}\right] \leq 2^{-nE}, \tag{176}$$

where here $\tilde{\mathbf{X}}$ is distributed uniformly over $\mathcal{D}_n$, and

$$|\mathcal{D}_n| \geq 2^{n(A-\delta)} \tag{177}$$

then a secure rate-distortion code can be constructed, with conditional exiguous-distortion exponent E. In this case, the code is constructed such that only source blocks in $\mathcal{D}_n$ are mapped to the permutation index $t^*(\mathbf{x}) = 0$, but not source blocks from $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})\backslash\mathcal{D}_n$. In addition, if the coding rate is unconstrained, then the condition (177) is not required. This fact will be utilized in the sequel in the proof of Lemma 13.

In the third step of the achievability proof, we construct the secure rate-distortion code for all types in $\mathcal{P}(\mathcal{X})$. We will need the following two lemmas.

Lemma 11. *Let $Q_X, Q'_X \in \mathcal{P}_n(\mathcal{X})$ and assume that*[10] $||Q_X - Q'_X|| = \frac{2d^*}{n}$ *where $d^* > 0$. If $\mathbf{x} \in \mathcal{T}_n(Q_X)$ then*

$$\min_{\mathbf{x}'\in\mathcal{T}_n(Q'_X)} d_{\mathrm{H}}(\mathbf{x},\mathbf{x}') \leq d^*. \tag{178}$$

Proof: See the extended version of [27, Lemma 20]. ■

Lemma 12. *Let $Q_X \in \mathcal{P}_n(\mathcal{X})$ and $\mathbf{x} \in \mathcal{T}_n(Q_X)$. For any given $1 \leq k < n$ let $\mathbf{x}' = \mathbf{x}_1^{n-k}$. Then*

$$||\hat{Q}_{\mathbf{x}} - \hat{Q}_{\mathbf{x}'}|| < |\mathcal{X}|\cdot\frac{k}{n-k}. \tag{179}$$

Proof: See the extended version of [27, Lemma 21]. ■

We are now ready for the third and final step of the proof of the achievability part of Theorem 1.

Proof of achievability part of Theorem 1: Let $0 < \epsilon < 1$ be given, and find n_0 sufficiently large such that for any $Q'_X \in \mathcal{P}(\mathcal{X})$ there exists $Q_X \in \mathcal{P}_{n_0}(\mathcal{X}) \cap \mathrm{int}\,\mathcal{Q}(\mathcal{X})$ such that $||Q_X - Q'_X|| \leq \frac{\epsilon}{2}$. We will term $\mathcal{P}_{n_0}(\mathcal{X}) \cap \mathrm{int}\,\mathcal{Q}(\mathcal{X})$ as the *grid*. Also let $n_1 = n_0\epsilon + 2n_0|\mathcal{X}|$. We construct the following sequence of secure rate-distortion codes $\mathcal{S}$ for

[10]For two different types in $\mathcal{P}_n(\mathcal{X})$, the minimal variation distance is $\frac{2}{n}$.

all $n > \max\{n_0, n_1\}$. We will use the following definitions and constructions:

- Let $\tilde{n} = \left\lfloor \frac{n}{n_0} \right\rfloor \cdot n_0$.
- Enumerate the types of the source $\mathcal{P}_n(\mathcal{X})$.
- Assume, w.l.o.g., that $\mathcal{X} = \{1, \ldots, |\mathcal{X}|\}$ and let $\overline{\mathcal{X}} \triangleq \{0\} \cup \mathcal{X}$.
- Let

$$\mathcal{B}_{\mathrm{H}}^n(\epsilon) \triangleq \left\{ \mathbf{x} \in \overline{\mathcal{X}}^n : d_{\mathrm{H}}(\mathbf{x}, \mathbf{0}) \leq \frac{n\epsilon}{2} \right\}, \tag{180}$$

  i.e., an Hamming ball of radius $\frac{n\epsilon}{2}$ and dimension n.
- Construct the codes $\mathcal{S}^*_{\tilde{n},Q_X} = (f^*_{\tilde{n},Q_X}, \varphi^*_{\tilde{n},Q_X})$ of key rate R as in Lemma 9, for all $Q_X \in \mathcal{P}_{n_0}(\mathcal{X}) \cap \operatorname{int} \mathcal{Q}(\mathcal{X})$.
- For every given $Q_X \in \mathcal{P}_n(\mathcal{X})$ find

$$\Phi_\epsilon(Q_X) \triangleq \operatorname*{arg\,min}_{Q'_X \in \mathcal{P}_{n_0}(\mathcal{X}) \cap \operatorname{int} \mathcal{Q}(\mathcal{X})} ||Q_X - Q'_X||. \tag{181}$$

- For any given $\mathbf{x} \in \mathcal{X}^n$ and $\overline{\mathbf{x}} \in \overline{\mathcal{X}}^n$, define the *replacement operator* $\Psi : \mathcal{X}^n \times \overline{\mathcal{X}}^n \to \overline{\mathcal{X}}^n$ which for $\tilde{\mathbf{x}} = \Psi(\mathbf{x}, \overline{\mathbf{x}})$ satisfies

$$\tilde{x}_i = \begin{cases} x_i, & \overline{x}_i = 0 \\ \overline{x}_i, & \overline{x}_i \neq 0 \end{cases} \tag{182}$$

- For a given $\mathbf{x} \in \mathcal{X}^n$, define the *replacement set*

$$\mathcal{K}(\mathbf{x}, \epsilon) \triangleq \left\{ \overline{\mathbf{x}} \in \mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon) : \Psi(\mathbf{x}_1^{\tilde{n}}, \overline{\mathbf{x}}) \in \mathcal{T}_{\tilde{n}}(\Phi_\epsilon(\hat{Q}_{\mathbf{x}})) \right\}. \tag{183}$$

  Note that the size of $\mathcal{K}(\mathbf{x}, \epsilon)$ depends on $\mathbf{x}$ only via its type $\hat{Q}_{\mathbf{x}}$.

The above type enumeration and the codes constructed are revealed to both the encoder and the decoder off-line. Before we provide the details of the encoding and the legitimate decoding, we outline the main ideas. Using the construction of Lemma 9, we construct secure rate distortion codes for each type in the *grid* $\mathcal{P}_{n_0}(\mathcal{X}) \cap \operatorname{int} \mathcal{Q}(\mathcal{X})$. Since this grid has a *finite* number of types, then for all sufficiently large n, the normalized logarithm of the conditional exiguous-distortion probability is close to the exponent (153) *uniformly* over all types in the grid. As mentioned in the outline of the proof in Section IV, we will modify any given source block so that it can be encoded using one of the codes in the grid. In order to allow the legitimate decoder to be able to reproduce with the desired distortion D_{L}, the cryptogram will be comprised of (at most) four parts, each one of them being encrypted using key bits $\mathbf{u}^{(i)}$ for $1 \leq i \leq 4$. First, the type of the source $\hat{Q}_{\mathbf{x}}$ is conveyed to the legitimate decoder, and, in accordance with Proposition 3, the type information is not encrypted, and so $\mathbf{u}^{(1)}$ is the empty string. This type will be modified to the type $\Phi_\epsilon(\hat{Q}_{\mathbf{x}})$, which is also known to the legitimate decoder and the eavesdropper. Second, since if $n \bmod n_0 \neq 0$ then $\hat{Q}_{\mathbf{x}}$ may not belong to the grid, we first truncate the source block to the length $\tilde{n}$. The truncated part $\mathbf{x}_{\tilde{n}+1}^n$ will be sent to the legitimate decoder losslessly, and fully encrypted using $\mathbf{u}^{(2)}$. Third, we will modify $\mathbf{x}_1^{\tilde{n}}$ to the *modified vector* $\mathbf{v}$, such that $\hat{Q}_{\mathbf{v}} = \Phi_\epsilon(\hat{Q}_{\mathbf{x}})$. This will be done by replacing a small number of the

symbols of $\mathbf{x}$. The symbols of $\mathbf{x}$ which were replaced in order to create $\mathbf{v}$ will be sent to the legitimate decoder losslessly, and fully encrypted using $\mathbf{u}^{(3)}$. Note, that there might be more than one way to replace the symbols of $\mathbf{x}$, and in fact, any $\overline{\mathbf{x}} \in \mathcal{K}(\mathbf{x}, \epsilon)$ can be used for this purpose if we define $\mathbf{v} \triangleq \Psi(\mathbf{x}_1^{\tilde{n}}, \overline{\mathbf{x}})$ using (182) and (183). For the sake of the analysis, it will be convenient to choose a replacement vector randomly from $\mathcal{K}(\mathbf{x}, \epsilon)$. This will be achieved using key bits $\overline{\mathbf{u}}$, which in this case, function as common randomness rather than for encryption. Fourth, the code $s^*_{\tilde{n},\Phi_\epsilon(\hat{Q}_\mathbf{x})}$ will be used to encode the modified vector $\mathbf{v}$ using the key bits $\mathbf{u}^{(4)}$. As we will prove, the whole modification procedure incurs a negligible cost on the compression and secrecy performance, which we analyze after formally defining the encoder and legitimate decoder.

Encoding: Let $\mathbf{u} = (\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{u}^{(3)}, \mathbf{u}^{(4)}, \overline{\mathbf{u}})$. The following cryptogram parts are generated:

- Source block type: Find the type index $0 \leq j^* \leq |\mathcal{P}_n(\mathcal{X})|-1$ of the source block type in the enumeration of the types, and let

$$y_1 \triangleq \mathbb{B}[j^*; \log|\mathcal{P}_n(\mathcal{X})|]. \tag{184}$$

  Set $\mathbf{u}^{(1)} = \phi$, namely, the type information is not encrypted, in accordance with Proposition 3.

- Fully encrypted source block tail:

$$y_2 \triangleq \mathbb{B}[\mathbf{x}_{\tilde{n}+1}^n; (n-\tilde{n})\log|\mathcal{X}|] \oplus \mathbf{u}^{(2)} \tag{185}$$

- Modification vector: Let $\overline{\mathbf{x}}$ be the $K_{\overline{\mathbf{u}}}$-th vector in $\mathcal{K}(\mathbf{x}, \epsilon)$, where $\overline{\mathbf{u}}$ is of length $\log|\mathcal{K}(\mathbf{x}, \epsilon)|$ bits, and $K_{\overline{\mathbf{u}}}$ is integer corresponding to $\mathbf{u}$, i.e.

$$K_{\overline{\mathbf{u}}} \triangleq \sum_{l=1}^{\log|\mathcal{K}(\mathbf{x},\epsilon)|} \overline{\mathbf{u}}_l \cdot 2^{(l-1)} + 1. \tag{186}$$

  Also, let

$$\mathbf{v} \triangleq \Psi(\mathbf{x}_1^{\tilde{n}}, \overline{\mathbf{x}}) \tag{187}$$

  and let $\mathbf{x}''' \in \overline{\mathcal{X}}^n$ where

$$x_i''' = \begin{cases} 0, & \overline{x}_i = 0 \\ x_i, & \overline{x}_i \neq 0 \end{cases}. \tag{188}$$

  As clearly $\mathbf{x}''' \in \mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)$, let i^* be the index of $\mathbf{x}'''$ in $\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)$ and

$$y_3 \triangleq \mathbb{B}[i^*; \log|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)|] \oplus \mathbf{u}^{(3)}. \tag{189}$$

- Cryptogram of modified vector: Let

$$y_4 \triangleq s^*_{\tilde{n},\Phi_\epsilon(\hat{Q}_\mathbf{x})}(\mathbf{v}, \mathbf{u}^{(4)}) \tag{190}$$

  where $\mathbf{u}^{(4)}$ is of length $n\mathsf{R}$ bits.

The encoding of the source block is separated into two cases, depending on its type $\hat{Q}_{\mathbf{x}}$. If $\mathsf{R}_{\text{L}} < R_{\text{L}}(\hat{Q}_{\mathbf{x}}, \mathsf{D}_{\text{L}})$ then

$$y = f_n^*(\mathbf{x}, \mathbf{u}) = y_1. \tag{191}$$

Otherwise, if $\mathsf{R}_{\text{L}} \geq R_{\text{L}}(\hat{Q}_{\mathbf{x}}, \mathsf{D}_{\text{L}})$ then

$$y = f_n^*(\mathbf{x}, \mathbf{u}) = (y_1, y_2, y_3, y_4). \tag{192}$$

To verify that such coding is possible, notice that from Lemma 12 and the fact that $n > n_1$, we have

$$||\hat{Q}_{\mathbf{x}_1^{\tilde{n}}} - \hat{Q}_{\mathbf{x}}|| \leq \frac{\epsilon}{2} \tag{193}$$

and by the triangle inequality

$$||\hat{Q}_{\mathbf{x}_1^{\tilde{n}}} - \hat{Q}_{\mathbf{v}}|| \leq ||\hat{Q}_{\mathbf{x}_1^{\tilde{n}}} - \hat{Q}_{\mathbf{x}}|| + ||\hat{Q}_{\mathbf{x}} - \hat{Q}_{\mathbf{v}}|| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \tag{194}$$

Thus, the definition (180), and Lemma 11 imply that $\mathcal{K}(\mathbf{x}, \epsilon)$ is indeed non-empty, and an appropriate $\overline{\mathbf{x}}$ can always be found.

Decoding by the legitimate decoder: Upon observing $y = f_n^*(\mathbf{x}, \mathbf{u})$:

- Recover the type $\hat{Q}_{\mathbf{x}}$ from y_1, and determine $\Phi_\epsilon(\hat{Q}_{\mathbf{x}})$ and $|\mathcal{K}(\mathbf{x}, \epsilon)|$.
- If $\mathsf{R}_{\text{L}} < R_{\text{L}}(\hat{Q}_{\mathbf{x}}, \mathsf{D}_{\text{L}})$ then arbitrarily choose a vector from $\tilde{\mathbf{w}} \in \mathcal{W}^n$, and reproduce

$$\mathbf{w} \triangleq \varphi_n^*(y, \mathbf{u}) = \tilde{\mathbf{w}}. \tag{195}$$

  Otherwise, if $\mathsf{R}_{\text{L}} \geq R_{\text{L}}(\hat{Q}_{\mathbf{x}}, \mathsf{D}_{\text{L}})$ then:

  - Recover $\mathbf{x}_{\tilde{n}+1}^n$ from y_2 and $\mathbf{u}^{(2)}$. Let $\mathbf{w}'' \in \mathcal{W}^{n-\tilde{n}}$ be such that $d_{\text{L}}(\mathbf{x}_{\tilde{n}+1}^n, \mathbf{w}'') = 0$.
  - Recover $\mathbf{x}'''$ from y_3 and $\mathbf{u}^{(3)}$, and let $\mathbf{w}''' \in \mathcal{W}^{\tilde{n}}$ be such that $d_{\text{L}}(\mathbf{x}''', \mathbf{w}''') = 0$.
  - Reproduce $\mathbf{v}$ from y_4 and $\mathbf{u}^{(4)}$ as

$$\mathbf{w}'''' \triangleq \varphi_{\tilde{n}, \hat{Q}_{\mathbf{x}}}^*(y_4, \mathbf{u}^{(4)}) \tag{196}$$

  - Reproduce the source block as

$$\mathbf{w} \triangleq \varphi_n^*(y, \mathbf{u}) = (\Psi(\mathbf{w}'''', \mathbf{w}'''), \mathbf{w}''). \tag{197}$$

Note that the decoder knows $|\mathcal{K}(\mathbf{x}, \epsilon)|$ and thus can compute the total length of $\mathbf{u}$. So, if multiple source blocks are encoded in succession, the legitimate decoder can stay synchronized with the encoder and use the correct key bits when deciphering the message.

For the sequence of codes $\mathcal{S}^*$ constructed, we need to verify that the compression constraint is satisfied, and to find the achievable exiguous-distortion exponent for any (type aware) eavesdropper, as well as the key rate. First, consider the compression constraint. For the rate, recall that the cryptogram is composed of at most four parts (192). Let $\mathcal{Y}_{nj}$ be the alphabet of the j-th part, for $1 \leq j \leq 4$, such that $|\mathcal{Y}_n| = \prod_{j=1}^4 |\mathcal{Y}_{nj}|$. We have,

$$|\mathcal{Y}_{n1}|= |\mathcal{P}_n(\mathcal{X})|\leq (n+1)^{|\mathcal{X}|}, \tag{198}$$

and

$$|\mathcal{Y}_{n2}|= |\mathcal{X}|^{n-\tilde{n}}. \tag{199}$$

For $\mathcal{Y}_{n3}$,

$$|\mathcal{Y}_{n3}|= \left|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)\right| = \sum_{k=0}^{\frac{\tilde{n}\epsilon}{2}} \binom{\tilde{n}}{k} |\mathcal{X}|^k \tag{200}$$

$$\leq \frac{\tilde{n}\epsilon}{2} \cdot \binom{\tilde{n}}{\lceil \frac{\tilde{n}\epsilon}{2} \rceil} |\mathcal{X}|^{\frac{\tilde{n}\epsilon}{2}} \tag{201}$$

$$\leq 2^{\tilde{n}\left[h_{\mathrm{B}}(\frac{\epsilon}{2})+\frac{\epsilon}{2}\log|\mathcal{X}|\right]} \tag{202}$$

$$\triangleq 2^{\tilde{n}g(\epsilon)} \tag{203}$$

where $g(\epsilon)$ was implicitly defined, and $g(\epsilon)\downarrow 0$ as $\epsilon \downarrow 0$. For $\mathcal{Y}_{n4}$, notice that the cryptogram part y_4 is only used for types Q_X which satisfy $\mathsf{R}_{\mathrm{L}} \geq R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})$. Thus,

$$|\mathcal{Y}_{n4}| \leq \sum_{Q_X \in \mathcal{P}_n(Q_X):\mathsf{R}_{\mathrm{L}} \geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} 2^{nR_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \tag{204}$$

$$\leq |\mathcal{P}_n(\mathcal{X})| \cdot 2^{n\mathsf{R}_{\mathrm{L}}} \tag{205}$$

Therefore, for all n sufficiently large

$$\limsup_{n\to\infty} \frac{1}{n}\log|\mathcal{Y}_n| \leq \limsup_{n\to\infty} \sum_{j=1}^{4} \frac{1}{n}\log|\mathcal{Y}_{nj}| \tag{206}$$

$$\leq \mathsf{R}_{\mathrm{L}} + g(\epsilon) + 3\delta. \tag{207}$$

Now, as the codes $\mathcal{S}^*_{\tilde{n},Q_X}$ are constructed according to Lemma 9, it is easily verified that if $\mathsf{R}_{\mathrm{L}} \geq R_{\mathrm{L}}(\hat{Q}_{\mathbf{x}}, \mathsf{D}_{\mathrm{L}})$ then for any $\mathbf{u}$

$$d_{\mathrm{L}}(\mathbf{x}, \varphi_n^*(f_n^*(\mathbf{x},\mathbf{u}),\mathbf{u})) \leq \mathsf{D}_{\mathrm{L}} \tag{208}$$

(see (152)). Thus, as $|\mathcal{P}_n(\mathcal{X})|\leq (n+1)^{|\mathcal{X}|}$, for all n sufficiently large

$$\mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \varphi_n^*(f_n^*(\mathbf{X},\mathbf{u}),\mathbf{u})) \geq \mathsf{D}_{\mathrm{L}}\right] \tag{209}$$

$$= \sum_{Q_X \in \mathcal{P}_n(\mathcal{X})} \mathbb{P}\left[\mathbf{X} \in \mathcal{T}_n(Q_X)\right] \mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \varphi_n^*(f_n^*(\mathbf{X},\mathbf{u}),\mathbf{u})) \geq \mathsf{D}_{\mathrm{L}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{210}$$

$$\leq \sum_{Q_X \in \mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}} < R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \mathbb{P}\left[\mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{211}$$

$$\leq \sum_{Q_X \in \mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}} < R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} 2^{-nD(Q_X||P_X)} \tag{212}$$

$$\leq 2^{-n[E_{\mathrm{L}}(P_X,\mathsf{D}_{\mathrm{L}},\mathsf{R}_{\mathrm{L}})-\delta]} \tag{213}$$

$$\leq 2^{-n(\mathsf{E}_{\mathrm{L}}-\delta)}. \tag{214}$$

Second, let us analyze the exiguous-distortion exponent of $\mathcal{S}$ for an arbitrary eavesdropper. Let $\hat{\mathbf{v}}^*$ be the eavesdropper which maximizes the exiguous-distortion probability for the modified source block $\mathbf{v}$, given the cryptogram y. Then,

$$\mathcal{E}_d^-(\mathcal{S},\mathsf{D}_{\mathrm{E}}) \overset{(a)}{=} \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) - \frac{1}{n}\log \max_{\tilde{\sigma}_n\in\tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z})\leq \mathsf{D}_{\mathrm{E}}|\mathbf{X}\in\mathcal{T}_n(Q_X)\right]\right\} \tag{215}$$

$$\overset{(b)}{\geq} \liminf_{n\to\infty} \min\Bigg\{ \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) - \frac{1}{n}\log\left(|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)|\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{V},\hat{\mathbf{V}}^*)\leq \mathsf{D}_{\mathrm{E}}|\mathbf{V}\in\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right]\right)\Bigg\},$$

$$\min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}<R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \left\{ D\left(Q_X||P_X\right) - \frac{1}{n}\log \max_{\tilde{\sigma}_n\in\tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z})\leq \mathsf{D}_{\mathrm{E}}|\mathbf{X}\in\mathcal{T}_n(Q_X)\right]\right\}\Bigg\} \tag{216}$$

$$\overset{(c)}{\geq} \liminf_{n\to\infty} \min\Bigg\{ \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) - \frac{1}{n}\log\left[|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)|\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{V},\hat{\mathbf{V}}^*)\leq \mathsf{D}_{\mathrm{E}}|\mathbf{V}\in\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right]\right]\Bigg\},$$

$$\min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}<R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \left\{ D\left(Q_X||P_X\right) + R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}}) - \delta\right\}\Bigg\} \tag{217}$$

$$= \min\Bigg\{ \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) - \frac{1}{n}\log\left[|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)|\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{V},\hat{\mathbf{V}}^*)\leq \mathsf{D}_{\mathrm{E}}|\mathbf{V}\in\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right]\right]\Bigg\},$$

$$\liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}<R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \left\{ D\left(Q_X||P_X\right) + R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}}) - \delta\right\}\Bigg\}, \tag{218}$$

where the passages are explained as follows:

- Equality (a) is standard method of types, (as, e.g., in (214)). Notice that the exiguous-distortion event $\{d_{\mathrm{E}}(\mathbf{X},\mathbf{Z})\leq \mathsf{D}_{\mathrm{E}}\}$ in this equation is for the code $\mathcal{S}_n$.
- Equality (b) is verified by establishing the following properties:
  - *Property 1:* Due to the permutation invariance of type classes and Hamming spheres, given the event $\mathbf{X}\in\mathcal{T}_n(Q_X)$, $\mathbf{V}$ is distributed uniformly over $\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))$. Indeed, let $\mathbf{v}',\mathbf{v}''\in\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))$, where $\mathbf{v}'=\pi(\mathbf{v}'')$ for some permutation π. Then, if for some $\mathbf{x}\in\mathcal{T}_n(Q_X)$ and $\overline{\mathbf{x}}\in\mathcal{K}(\mathbf{x},\epsilon)$

$$\mathbf{v}'=\Psi(\mathbf{x}_1^{\tilde{n}},\overline{\mathbf{x}}) \tag{219}$$

then

$$\mathbf{v}'' = \Psi(\pi(\mathbf{x}_1^{\tilde{n}}), \pi(\overline{\mathbf{x}})) \tag{220}$$

where $(\pi(\mathbf{x}_1^{\tilde{n}}), \mathbf{x}_{\tilde{n}+1}^n) \in \mathcal{T}_n(Q_X)$ and $\pi(\overline{\mathbf{x}}) \in \mathcal{K}((\pi(\mathbf{x}_1^{\tilde{n}}), \mathbf{x}_{\tilde{n}+1}^n), \epsilon)$[11]. The property then follows from the fact that $|\mathcal{K}(\mathbf{x}, \epsilon))|$ depends on $\mathbf{x}$ only via its type, which is identical for both $\mathbf{x}$ and $(\pi(\mathbf{x}_1^{\tilde{n}}), \mathbf{x}_{\tilde{n}+1}^n)$.

– *Property 2:* An eavesdropper for $\mathbf{v}$ is aware of its type (as $\hat{Q}_{\mathbf{v}} = \Phi_\epsilon(\hat{Q}_{\mathbf{x}})$)[12], and the cryptogram y_2 is not relevant for its estimate. Also, since y_3 is fully encrypted (pure random bits) then it is also useless. Thus, an eavesdropper for $\mathbf{v}$ uses only the type information in y_1 and y_4.

– *Property 3:* Consider the case $\mathsf{R}_{\text{L}} \geq R_{\text{L}}(Q_X, \mathsf{D}_{\text{L}})$. The source block $\mathbf{X}$ is distributed uniformly over $\mathcal{T}_n(Q_X)$ and $\mathbf{V}$ is distributed uniformly over $\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))$. Let $\hat{\mathbf{V}}^*$ be the eavesdropper which achieves the maximal exiguous-distortion probability for $\mathbf{V}$, given y_4. Then, for any eavesdropper decoder $\tilde{\sigma}_n$ which estimates $\mathbf{z}$

$$\frac{1}{|\mathcal{B}_{\text{H}}^{\tilde{n}}(\epsilon)|}\mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{Z}) \leq \mathsf{D}_{\text{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \leq \mathbb{P}\left[d_{\text{E}}(\mathbf{V}, \hat{\mathbf{V}}^*) \leq \mathsf{D}_{\text{E}} | \mathbf{V} \in \mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right]. \tag{221}$$

Indeed, since $\mathbf{X}_{\tilde{n}+1}^n$ is fully encrypted then it is easy to verify that

$$\mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{Z}) \leq \mathsf{D}_{\text{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \leq \mathbb{P}\left[d_{\text{E}}(\mathbf{X}_1^{\tilde{n}}, \mathbf{Z}_1^{\tilde{n}}) \leq \mathsf{D}_{\text{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right]. \tag{222}$$

Now, any eavesdropper $\mathbf{Z}_1^{\tilde{n}}$ for $\mathbf{X}_1^{\tilde{n}}$ can be transformed into an eavesdropper $\hat{\mathbf{V}}$ for $\mathbf{V}$, by a uniformly distributed guess of $\overline{\mathbf{X}}$ over $\mathcal{B}_{\text{H}}^{\tilde{n}}(b)$ (see (187)) and then setting

$$\hat{\mathbf{v}} = \begin{cases} \arg\min_{z \in \mathcal{Z}} d_{\text{E}}(\overline{\mathbf{x}}_i, z), & \overline{\mathbf{x}}_i \neq 0 \\ \mathbf{z}_i, & \overline{\mathbf{x}}_i = 0 \end{cases} \tag{223}$$

where by assumption, $\min_{z \in \mathcal{Z}} d_{\text{E}}(\overline{\mathbf{x}}_i, z) = 0$. If the guess of $\overline{\mathbf{x}}$ is correct (according to the relation (187)) then

$$d_{\text{E}}(\mathbf{v}, \hat{\mathbf{v}}) \leq d_{\text{E}}(\mathbf{x}, \mathbf{z}). \tag{224}$$

Since this happens with probability larger than $\left[|\mathcal{B}_{\text{H}}^{\tilde{n}}(\epsilon)|\right]^{-1}$, then (222) implies (221).

Equality (b) then follows from the above considerations.

- Inequality (c) is because in case $\mathsf{R}_{\text{L}} < R_{\text{L}}(Q_X, \mathsf{D}_{\text{L}})$ the eavesdropper has no knowledge beyond the type of the source block, and so given such y, $\mathbf{x}$ is distributed uniformly over $\mathcal{T}_n(Q_X)$. For any given $\mathbf{z} \in \mathcal{Z}^n$, using standard method of types

$$\mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D}_{\text{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] = \sum_{\mathbf{x} \in \mathcal{T}_n(Q_X) : d_{\text{E}}(\mathbf{x}, \mathbf{z}) \leq \mathsf{D}_{\text{E}}} \frac{1}{|\mathcal{T}_n(Q_X)|} \tag{225}$$

[11]Notice that $\mathcal{K}(\mathbf{x})$ depends on $\mathbf{x}$ only via its first $\tilde{n}$ components.

[12]Which is in fact not even required, using Proposition 3.

$$= \sum_{Q_{X|Z}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}} \sum_{\mathbf{x}\in\mathcal{T}_n(Q_{X|Z},\mathbf{z})} \frac{1}{|\mathcal{T}_n(Q_X)|} \tag{226}$$

$$= \frac{1}{|\mathcal{T}_n(Q_X)|} \sum_{Q_{X|Z}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}} \left|\mathcal{T}_n(Q_{X|Z},\mathbf{z})\right| \tag{227}$$

$$\doteq \exp\left\{-n\cdot \min_{Q_{X|Z}:\mathbb{E}_Q[d_{\mathrm{E}}(X,Z)]\leq \mathsf{D}_{\mathrm{E}}} \left[-H_Q(X|Z)+H(Q_X)\right]\right\} \tag{228}$$

Then,

$$\max_{\mathbf{z}} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{z})\leq \mathsf{D}_{\mathrm{E}}|\mathbf{X}\in\mathcal{T}_n(Q_X)\right] \leq 2^{-n[R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})-\delta]}. \tag{229}$$

Next, we further bound the first term in the minimization of (218) as follows

$$\liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) - \frac{1}{n}\log\left[|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)|\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{V},\hat{\mathbf{V}}^*)\leq \mathsf{D}_{\mathrm{E}}|\mathbf{V}\in\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right]\right] \Bigg\} \tag{230}$$

$$\overset{(a)}{\geq} \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) - \frac{1}{n}\log\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{V},\hat{\mathbf{V}}^*)\leq \mathsf{D}_{\mathrm{E}}|\mathbf{V}\in\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right] - g(\epsilon) \Bigg\} \tag{231}$$

$$\overset{(b)}{=} \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) - \frac{1}{\tilde{n}}\log\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{V},\hat{\mathbf{V}}^*)\leq \mathsf{D}_{\mathrm{E}}|\mathbf{V}\in\mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right] - g(\epsilon) \Bigg\} \tag{232}$$

$$\overset{(c)}{\geq} \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Big\{ D\left(Q_X||P_X\right) + \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\right\} - \delta - g(\epsilon) \Big\} \tag{233}$$

$$\overset{(d)}{\geq} \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Big\{ D\left(\Phi_\epsilon(Q_X)||P_X\right) + \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\right\} - \delta - \delta_1(\epsilon) - g(\epsilon) \Big\} \tag{234}$$

$$\overset{(e)}{=} \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_{n_0}(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Big\{ D\left(\Phi_\epsilon(Q_X)||P_X\right) + \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\right\} - \delta - \delta_1(\epsilon) - g(\epsilon) \Big\} \tag{235}$$

$$\overset{(f)}{=} \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_{n_0}(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Big\{ D\left(Q_X||P_X\right) + \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\right\} - \delta - \delta_1(\epsilon) - g(\epsilon) \Big\} \tag{236}$$

$$= \min_{Q_X\in\mathcal{P}_{n_0}(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Big\{ D\left(Q_X||P_X\right) + \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\right\} - \delta - \delta_1(\epsilon) - g(\epsilon) \Big\}, \tag{237}$$

$$\geq \liminf_{n\to\infty} \min_{Q_X\in\mathcal{P}_n(\mathcal{X}):\mathsf{R}_{\mathrm{L}}\geq R_{\mathrm{L}}(Q_X,\mathsf{D}_{\mathrm{L}})} \Big\{ D\left(Q_X||P_X\right) + \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X,\mathsf{D}_{\mathrm{E}})\right\} - \delta - \delta_1(\epsilon) - g(\epsilon) \Big\} \tag{238}$$

- Inequality (a) follows from the fact that since $0<\epsilon<1$, for all n sufficiently large $\left|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)\right| \leq 2^{\tilde{n}g(\epsilon)}$ as in (203).
- Equality (b) is because $\frac{\tilde{n}}{n}\to 1$ as $n\to\infty$.

- Inequality (c) is because there exists n_2 sufficiently large, such that for all $n > n_2$ the error probability of the any eavesdropper decoder $\sigma^*_{\tilde{n},\Phi_\epsilon(Q_X)}$ satisfies

$$-\frac{1}{\tilde{n}}\log\mathbb{P}\left[\hat{\mathbf{V}} \neq \mathbf{V}|\mathbf{V} \in \mathcal{T}_{\tilde{n}}(\Phi_\epsilon(Q_X))\right] \geq \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}})\right\} - \delta \tag{239}$$

  uniformly for all $Q_X \in \mathcal{P}_{n_0}(Q_X)$.

- Inequality (d) is by defining

$$\delta_1(\epsilon) \triangleq \max_{Q_X} \left| D\left(\Phi_\epsilon(Q_X)||P_X\right) - D(Q_X||P_X)\right|. \tag{240}$$

  Note that since $D(Q_X||P_X)$ is a continuous function of Q_X in $\mathcal{Q}(\mathcal{X})$ (as the support of P_X is assumed to be $\mathcal{X}$), it is also uniformly continuous. So, $\delta_1(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$.

- Equalities (e) and (f) are because $\Phi_\epsilon(Q_X) \in \mathcal{P}_{n_0}(\mathcal{X})$ for all $Q_X \in \mathcal{P}_n(\mathcal{X})$.

Substituting (238) into (218), and using the fact $\mathcal{P}_n(\mathcal{X}) \subset \mathcal{Q}(\mathcal{X})$ we obtain

$$\begin{aligned}
\mathcal{E}_d^-(\mathcal{S}, \mathsf{D}_{\mathrm{E}}) \geq \min\Bigg\{ & \min_{Q_X \in \mathcal{Q}(\mathcal{X}):\mathsf{R}_{\mathrm{L}} \geq R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) + \min\left\{\mathsf{R}, R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}})\right\} - g(\epsilon) - \delta_1(\epsilon), \\
& \min_{Q_X \in \mathcal{Q}(\mathcal{X}):\mathsf{R}_{\mathrm{L}} < R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})} \left\{ D\left(Q_X||P_X\right) + R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}}) \right\} \Bigg\} \Bigg\} - \delta \qquad (241) \\
\geq \min\Bigg\{ & \min_{Q_X \in \mathcal{Q}(\mathcal{X}):\mathsf{R}_{\mathrm{L}} \geq R_{\mathrm{L}}(Q_X, \mathsf{D}_{\mathrm{L}})} \Bigg\{ D\left(Q_X||P_X\right) + \mathsf{R}, \\
& \min_{Q_X \in \mathcal{P}(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) + R_{\mathrm{E}}(Q_X, \mathsf{D}_{\mathrm{E}}) \right\} \Bigg\} \Bigg\} - \delta - \delta_1(\epsilon) - g(\epsilon) \qquad (242) \\
\overset{(a)}{\geq} & \min\left\{\mathsf{R}, E_e^*(\mathsf{D}_{\mathrm{E}})\right\} - \delta - \delta_1(\epsilon) - g(\epsilon) \qquad (243)
\end{aligned}$$

where in (a) we have used the definition in (16), and the fact that the assumption $\mathsf{E}_{\mathrm{L}} > 0$ implies that $\mathsf{R}_{\mathrm{L}} \geq R_{\mathrm{L}}(P_X, \mathsf{D}_{\mathrm{L}})$.

Next, we analyze the required key rate. If $\mathsf{R}_{\mathrm{L}} < R_{\mathrm{L}}(\hat{Q}_{\mathbf{x}}, \mathsf{D}_{\mathrm{L}})$ then the required key rate is zero. Otherwise, if $\mathsf{R}_{\mathrm{L}} \geq R_{\mathrm{L}}(\hat{Q}_{\mathbf{x}}, \mathsf{D}_{\mathrm{L}})$ then the total key rate required to encode $\mathbf{x} \in Q_X$ is given by

$$\frac{1}{n}\left[(n-\tilde{n})\log|\mathcal{X}| + \log|\mathcal{K}(\mathbf{x},\epsilon)| + \log\left|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)\right| + n\mathsf{R}\right]. \tag{244}$$

Now, for all n sufficiently large

$$\frac{1}{n}(n-\tilde{n})\log|\mathcal{X}| \leq \frac{n_0\log|\mathcal{X}|+1}{n} \leq \delta, \tag{245}$$

$$\frac{1}{n}\log|\mathcal{K}(\mathbf{x},\epsilon)| \leq \frac{1}{n}\log\left|\mathcal{B}_{\mathrm{H}}^{\tilde{n}}(\epsilon)\right| \leq g(\epsilon), \tag{246}$$

Thus, the required key rate is less than

$$\mathsf{R} + 2g(\epsilon) + 2\delta. \tag{247}$$

By taking $\epsilon \downarrow 0$ we obtain $g(\epsilon) \downarrow 0$ and $\delta_1(\epsilon) \downarrow 0$, and so we obtain the achievability part of Theorem 1. ■

D. Proof of Converse Part of Theorem 1

Following the outline of the converse, we begin with a lemma which constructs from a given sequence of codes $\mathcal{S}$ a new sequence $\mathcal{S}^*$, with constant key rate, which is less than $\overline{R}(\mathcal{S}, Q_X) + \delta$, and a zero excess-distortion probability at the legitimate receiver.

Lemma 13. *Let $\mathcal{S}$ be an arbitrary sequence of secure rate-distortion codes, which satisfies a compression constraint $(\mathsf{R}_{\mathrm{L}}, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$. Also, let $Q_X \in \mathcal{P}(\mathcal{X})$ be given such that $D(Q_X||P_X) < \mathsf{E}_{\mathrm{L}}$. Then, for every $\delta > 0$, there exists a sequence of secure rate-distortion codes $\mathcal{S}^*$ such that:*

1) *For all n and all $\mathbf{x} \in \mathcal{T}_n(Q_X)$, $\mathcal{S}_n^*$ has fixed key rate $r^*(\mathbf{x}) = \mathsf{R}^*$ where $\mathsf{R}^* \leq \overline{R}(\mathcal{S}, Q_X) + \delta$.*
2) *For all n and $\{u_i\}_{i=1}^{\infty}$, $\mathcal{S}_n^* = (f_n^*, \varphi_n^*)$ satisfies*

$$\mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \varphi_n^*(f_n^*(\mathbf{X}, \mathbf{u}), \mathbf{u})) > \mathsf{D}_{\mathrm{L}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] = 0, \tag{248}$$

and in addition, $\mathcal{S}^*$ satisfies a compression constraint $(\mathsf{R}_{\mathrm{L}}^*, \mathsf{D}_{\mathrm{L}}, \mathsf{E}_{\mathrm{L}})$ for $\mathsf{R}_{\mathrm{L}}^* = \log|\mathcal{X}|$.

3) *For every $\mathsf{D}_{\mathrm{E}} \geq \mathsf{D}_{\mathrm{L}}$.*

$$\mathcal{E}_d^+(\mathcal{S}, \mathsf{D}_{\mathrm{E}}, Q_X) \leq \mathcal{E}_d^+(\mathcal{S}^*, \mathsf{D}_{\mathrm{E}}, Q_X) + \delta. \tag{249}$$

Proof: We will prove this lemma by modifying the sequence of codes $\mathcal{S}$ into the new sequence $\mathcal{S}^*$. Assume that $Q_X \in \operatorname{int} \mathcal{Q}(\mathcal{X})$, and $Q_X \in \mathcal{P}_{n_0}(\mathcal{X})$ for some minimal $n_0 \in \mathbb{N}$. Since the statements in the lemma are only about conditional events given the type Q_X, it is clear that the new secure rate-distortion codes constructed $\mathcal{S}_n^*$ need only be different from $\mathcal{S}_n$ for $\mathbf{x} \in \mathcal{T}_n(Q_X)$, and so only block-lengths $n \bmod n_0 = 0$ should be considered, as otherwise $\mathcal{T}_n(Q_X)$ is empty. To wit, the limit $n \to \infty$ should be read as limit $l \to \infty$ for $n = n_0 l$, but this will not be explicitly written, for the sake of brevity. Throughout the proof, quantities that are related to $\mathcal{S}^*$ will be superscripted by $*$. For brevity, we will denote the conditional key rate by $\overline{R}(Q_X)$ and $\overline{R}^*(Q_X)$ for $\mathcal{S}$ and $\mathcal{S}^*$, respectively .

Let $\delta > 0$ be given. For any length $0 \leq m \leq n \log|\mathcal{X}|$ and $y \in \mathcal{Y}_n$ define the *ambiguity sets for a given key-length as*

$$\mathcal{A}_n(y, m) \triangleq \left\{\mathbf{x} \in \mathcal{T}_n(Q_X) : k_n(\mathbf{x}) = m, f_n(\mathbf{x}, \mathbf{u}) = y \text{ for some } \mathbf{u} \in \{0,1\}^m\right\}, \tag{250}$$

and with a slight abuse of notation define the *ambiguity set*[13] as

$$\mathcal{A}_n(y) \triangleq \bigcup_{m=0}^{n \log|\mathcal{X}|} \mathcal{A}_n(y, m). \tag{251}$$

For any given y and $\mathbf{x} \in \mathcal{A}_n(y)$, let us denote the reproduction $\mathbf{w}(\mathbf{x}, y) \triangleq \varphi(y, \mathbf{u})$, where $\mathbf{u}$ satisfies $f_n(\mathbf{x}, \mathbf{u}) = y$,

[13]Called *residue class* in the terminology of [1].

and the *ambiguity set without excess-distortion*

$$\mathcal{D}_n(y) \triangleq \{\mathbf{x} \in \mathcal{A}_n(y) : d_{\mathrm{L}}(\mathbf{x}, \mathbf{w}(\mathbf{x}, y)) \leq \mathsf{D}_{\mathrm{L}}\}. \tag{252}$$

Also, consider the *modified ambiguity set*

$$\mathcal{A}_n^*(y) \triangleq \left\{ \mathcal{A}_n(y) \backslash \bigcup_{m=0}^{n(\overline{R}(Q_X)-\delta)} \mathcal{A}_n(y,m) \backslash \bigcup_{m=n(\overline{R}(Q_X)+\delta)}^{n\log|\mathcal{X}|} \mathcal{A}_n(y,m) \right\} \bigcap \mathcal{D}_n(y). \tag{253}$$

For a given y, the eavesdropper knows that $\mathbf{x} \in \mathcal{A}_n(y)$ and chooses its estimate accordingly. However, conditioned on y, the probability of $\mathbf{X}$ is not uniform over $\mathcal{A}_n(y)$, since $k_n(\mathbf{x})$ is not the same for all $\mathbf{x} \in \mathcal{A}_n(y)$. The proof of the lemma is divided into two steps and its outline is as follows. In the first step, we will identify a sequence of cryptograms $\{y_n^*\}$ which simultaneously satisfies the following properties:

1) The conditional exiguous-distortion exponent of the eavesdropper when $\mathbf{X}$ is distributed *uniformly* over $\mathcal{A}_n^*(y_n^*)$ is larger than the one for $\mathbf{X}$ distributed over $\mathcal{A}_n(y_n^*)$ according to the distribution induced by $\mathcal{S}_n$.
2) The conditional exiguous-distortion exponent conditioned on $Y = y_n^*$ equals the same exponent without this conditioning.

In the second step of the proof, we utilize the set $\mathcal{A}_n^*(y_n^*)$ to construct the new sequence of codes $\mathcal{S}^*$. This is done by the same technique used in the achievability proof of Lemma 9 - by an efficient covering of the type class using permutations of one good set $\mathcal{A}_n^*(y_n^*)$. The two properties above of y_n^* will be used to show that the exiguous-distortion exponent of $\mathcal{S}^*$ may be only slightly less than that of $\mathcal{S}$.

We begin with the first step. For brevity, let us assume that $\mathbf{X}$ is distributed uniformly over the type class $\mathcal{T}_n(Q_X)$, and probabilities, expectations and entropies will be calculated w.r.t. this probability distribution. So, we only consider y such that $\mathcal{A}_n(y)$ is non-empty. If we let

$$A(y) \triangleq \mathbb{P}\left[\overline{R}(Q_X) - \delta \leq r_n(\mathbf{X}) \leq \overline{R}(Q_X) + \delta, d_{\mathrm{L}}(\mathbf{X}, \mathbf{W}) \leq \mathsf{D}_{\mathrm{L}} | Y = y\right] \tag{254}$$

then for n sufficiently large

$$\begin{aligned}
\mathbb{E}\left[A(Y)\right] &= \mathbb{P}\left[\overline{R}(Q_X) - \delta \leq r_n(\mathbf{X}) \leq \overline{R}(Q_X) + \delta, d_{\mathrm{L}}(\mathbf{X}, \mathbf{W}) \leq \mathsf{D}_{\mathrm{L}}\right] \\
&\geq \mathbb{P}\left[\overline{R}(Q_X) - \delta \leq r_n(\mathbf{X}) \leq \overline{R}(Q_X) + \delta\right] - \mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \mathbf{W}) > \mathsf{D}_{\mathrm{L}}\right] \\
&\overset{(a)}{\geq} \delta - \mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X}, \mathbf{W}) > \mathsf{D}_{\mathrm{L}}\right] \\
&\overset{(b)}{\geq} \delta - 2^{-n[\mathsf{E}_{\mathrm{L}} - D(Q_X||P_X) - \delta]} \qquad (255) \\
&\triangleq \frac{\delta}{2} \qquad (256)
\end{aligned}$$

where (a) is using the convergence in probability of $r_n(\mathbf{X})$ to $\overline{R}(Q_X)$ (see (11)), and (b) is since $\mathcal{S}$ satisfies a

compression constraint $(\mathsf{R}_{\text{L}}, \mathsf{D}_{\text{L}}, \mathsf{E}_{\text{L}})$ and the assumption $D(Q_X||P_X) < \mathsf{E}_{\text{L}}$. Defining for any $0 < \beta < 1$

$$\mathcal{V}_n^{(1)} \triangleq \left\{ y \in \mathcal{Y}_n : A(y) \geq \beta \cdot \frac{\delta}{2} \right\}, \tag{257}$$

then, since from the definition (254) and (256)

$$0 \leq \frac{A(y)}{\mathbb{E}\left[A(Y)\right]} \leq \frac{2}{\delta} \tag{258}$$

for all $y \in \mathcal{Y}_n$, the reverse Markov inequity (Lemma 2) implies that

$$\mathbb{P}\left(Y \in \mathcal{V}_n^{(1)}\right) \geq \frac{1-\beta}{\frac{2}{\delta} - \beta} \triangleq \zeta(\delta, \beta), \tag{259}$$

and choosing some $\beta^* < \min\{1, \frac{2}{\delta}\}$, we obtain $\zeta^*(\delta) \triangleq \zeta(\delta, \beta^*) > 0$. Now, for $\gamma > 1$, let

$$\mathcal{V}_n^{(2)} \triangleq \left\{ y \in \mathcal{Y}_n : \max_{\mathbf{z}} \mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D}_{\text{E}} | Y = y\right] < \gamma \cdot \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{Z}) \leq \mathsf{D}_{\text{E}}\right] \right\}. \tag{260}$$

Then the Markov inequality implies

$$\mathbb{P}(Y \notin \mathcal{V}_n^{(2)}) = \mathbb{P}\left[\max_{\mathbf{z}} \mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D}_{\text{E}} | Y\right] \geq \gamma \cdot \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{Z}) \leq \mathsf{D}_{\text{E}}\right] \right] \tag{261}$$

$$\overset{(a)}{\leq} \frac{\mathbb{E}\left[\max_{\mathbf{z}} \mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D}_{\text{E}} | Y\right]\right]}{\gamma \cdot \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\text{E}}(\mathbf{X}, \mathbf{Z}) \leq \mathsf{D}_{\text{E}}\right]} \tag{262}$$

$$= \frac{1}{\gamma} \tag{263}$$

where in (a) is should be recalled that $\mathbf{z}$ is chosen as a function of Y. Hence, by the union bound

$$\mathbb{P}\left(Y \in \mathcal{V}_n^{(1)} \cap \mathcal{V}_n^{(2)}\right) \geq 1 - \mathbb{P}\left(Y \notin \mathcal{V}_n^{(1)}\right) - \mathbb{P}\left(Y \notin \mathcal{V}_n^{(2)}\right) \tag{264}$$

$$\geq \zeta^*(\delta) - \frac{1}{\gamma}. \tag{265}$$

Thus, for any given δ, there exists $\gamma^* > 1$ sufficiently large (but independent of n) such that

$$\mathbb{P}\left(Y \in \mathcal{V}_n^{(1)} \cap \mathcal{V}_n^{(2)}\right) > 0. \tag{266}$$

Therefore, there exists a sequence $\{y_n^*\}$ such that for all n sufficiently large, $y_n^* \in \mathcal{V}_n^{(1)} \cap \mathcal{V}_n^{(2)}$.

In the second step of the proof, we describe the construction of $\mathcal{S}_n^*$. Note that by letting

$$\mathcal{U}^* \triangleq \{\mathbf{u} : \exists \mathbf{x} \in \mathcal{A}_n^*(y_n^*) \text{ such that } f_n(\mathbf{x}, \mathbf{u}) = y_n^*\} \tag{267}$$

and

$$\mathcal{C}_n^* \triangleq \{\varphi_n(y_n^*, \mathbf{u}) : \mathbf{u} \in \mathcal{U}^*\} \tag{268}$$

we have that $\mathcal{A}_n^*(y_n^*) \subseteq \mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\text{L}})$. Now, recall that in Lemma 9 of the achievability proof, we have utilized

permutations of a D-cover $\mathfrak{D}(\mathcal{C}_n^*, Q_X, \mathsf{D}_{\mathrm{L}})$ (of a set $\mathcal{C}_n^*$) which cover the type class $\mathcal{T}_n(Q_X)$, to construct a secure rate-distortion code. Following remark 10, the set $\mathcal{A}_n^*(y_n)$ can also be used as a constituent set in the construction of a secure rate-distortion code, and the conditional exiguous-distortion exponent equal to the exponent achieved when the source block $\mathbf{X}$ is distributed uniformly over $\mathcal{A}_n^*(y_n^*)$, as in (176). Let us find the exponent achieved when $\mathbf{X}$ is distributed uniformly over $\mathcal{A}_n^*(y_n^*)$. To this end, denote

$$\mathcal{M}(\delta) \triangleq \left[n\left(\overline{R}(Q_X) - \delta\right), n\left(\overline{R}(Q_X) + \delta\right)\right]. \tag{269}$$

and observe that for an arbitrary eavesdropper $\overline{\mathbf{z}}$, and all n sufficiently large,

$$\max_{\mathbf{z}} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \mathbf{z}) \leq \mathsf{D}_{\mathrm{E}} | Y = y_n^*\right] \tag{270}$$

$$\geq \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}} | Y = y_n^*\right] \tag{271}$$

$$= \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left[\mathbf{X} = \mathbf{x} | Y = y_n^*\right] \tag{272}$$

$$= \sum_{m=0}^{n \log |\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left[\mathbf{X} = \mathbf{x} | Y = y_n^*\right] \tag{273}$$

$$= \frac{\sum_{m=0}^{n \log |\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)}{\mathbb{P}\left(Y = y_n^*\right)} \tag{274}$$

$$\geq \frac{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)}{\mathbb{P}\left(Y = y_n^*\right)} \tag{275}$$

$$\geq \frac{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)}{\mathbb{P}\left(Y = y_n^*\right)} \tag{276}$$

$$\overset{(a)}{\geq} \beta \frac{\delta}{2} \cdot \frac{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)}{\mathbb{P}\left[\overline{R}(Q_X) - \delta \leq r_n(\mathbf{X}) \leq \overline{R}(Q_X) + \delta, d_{\mathrm{L}}(\mathbf{X}, \mathbf{W}) \leq \mathsf{D}_{\mathrm{L}}, Y = y_n^*\right]} \tag{277}$$

$$= \beta \frac{\delta}{2} \cdot \frac{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)}{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*)} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)} \tag{278}$$

$$= \beta \frac{\delta}{2} \cdot \frac{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)}{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*)} \mathbb{P}\left(\mathbf{X} = \mathbf{x}, Y = y_n^*\right)} \tag{279}$$

$$= \beta \frac{\delta}{2} \cdot \frac{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*): d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}} \mathbb{P}\left(Y = y_n^* | \mathbf{X} = \mathbf{x}\right)}{\sum_{m \in \mathcal{M}(\delta)} \sum_{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*)} \mathbb{P}\left(Y = y_n^* | \mathbf{X} = \mathbf{x}\right)} \tag{280}$$

$$\overset{(b)}{=} \beta \frac{\delta}{2} \cdot \frac{\sum_{m \in \mathcal{M}(\delta)} 2^{-m} \cdot \left|\{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*) : d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}\}\right|}{\sum_{m \in \mathcal{M}(\delta)} 2^{-m} \cdot \left|\mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*)\right|} \tag{281}$$

$$\geq \beta \cdot \frac{\delta}{2} \frac{2^{-n\left(\overline{R}(Q_X) + \delta\right)} \cdot \sum_{m \in \mathcal{M}(\delta)} \left|\{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*) : d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}\}\right|}{2^{-n\left(\overline{R}(Q_X) - \delta\right)} \cdot \sum_{m \in \mathcal{M}(\delta)} \left|\mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*)\right|} \tag{282}$$

$$= \beta \frac{\delta}{2} \cdot 2^{-2n\delta} \frac{\sum_{m \in \mathcal{M}(\delta)} \left|\{\mathbf{x} \in \mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*) : d_{\mathrm{E}}(\mathbf{x}, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}\}\right|}{\sum_{m \in \mathcal{M}(\delta)} \left|\mathcal{A}_n(y_n^*, m) \cap \mathcal{D}_n(y_n^*)\right|} \tag{283}$$

$$\triangleq \beta \frac{\delta}{2} \cdot 2^{-2n\delta} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X}^*, \overline{\mathbf{z}}) \leq \mathsf{D}_{\mathrm{E}}\right], \tag{284}$$

where (a) is because as $y_n^* \in \mathcal{V}_n^{(1)}$ implies that

$$\frac{\mathbb{P}\left[\overline{R}(Q_X) - \delta \le r_n(\mathbf{X}) \le \overline{R}(Q_X) + \delta, d_{\mathsf{L}}(\mathbf{X},\mathbf{W}) \le \mathsf{D}_{\mathsf{L}}, Y = y_n^*\right]}{\beta_{\frac{\delta}{2}}} \ge \mathbb{P}(Y = y_n^*), \tag{285}$$

and (b) is because for admissible encoders and $\mathbf{x} \in \mathcal{A}_n(y_n^*, m)$

$$\mathbb{P}\left(Y = y_n^* | \mathbf{X} = \mathbf{x}\right) = 2^{-m}. \tag{286}$$

Thus,

$$\limsup_{n\to\infty} -\frac{1}{n} \log \max_{\mathbf{z}} \mathbb{P}\left[d_{\mathsf{E}}(\mathbf{X}^*, \mathbf{z}) \le \mathsf{D}_{\mathsf{E}}\right] \ge \limsup_{n\to\infty} -\frac{1}{n} \max_{\mathbf{z}} \log \mathbb{P}\left[d_{\mathsf{E}}(\mathbf{X}, \mathbf{z}) \le \mathsf{D}_{\mathsf{E}} | Y = y_n^*\right] - 3\delta \tag{287}$$

$$\stackrel{(a)}{=} \mathcal{E}_d^+(\mathcal{S}, \mathsf{D}_{\mathsf{E}}, Q_X) - 3\delta \tag{288}$$

where (a) is because $y_n^* \in \mathcal{V}_n^{(2)}$. So, by choosing δ sufficiently small, we can achieve (249) by the permutation construction of Lemma 9.

Finally, as the legitimate reconstruction $\mathbf{w}(\mathbf{x}, y_n^*)$ of any $\mathbf{x} \in \mathcal{A}_n^*(y_n^*)$ satisfies $d_{\mathsf{L}}(\mathbf{x}, \mathbf{w}(\mathbf{x}, y_n^*)) \le \mathsf{D}_{\mathsf{L}}$, the permutation construction assures this property for all $\mathbf{x} \in \mathcal{T}_n(Q_X)$. So, it is easy to verify that if $\mathcal{S}$ has excess-distortion exponent E_{L} at distortion level D_{L}, then $\mathcal{S}^*$ has an even larger exponent. As $\mathsf{R}_{\mathsf{L}}^* = \log|\mathcal{X}|$, the compression constraint $(\mathsf{R}_{\mathsf{L}}^*, \mathsf{D}_{\mathsf{L}}, \mathsf{E}_{\mathsf{L}})$ is satisfied by $\mathcal{S}^*$. ■

We are now ready for the second and final step of the proof of the converse part of Theorem 1.

Proof of converse part of Theorem 1: Let a sequence of secure rate-distortion codes $\mathcal{S}$ be given, which satisfies the compression constraint $(\mathsf{R}_{\mathsf{L}}, \mathsf{D}_{\mathsf{L}}, \mathsf{E}_{\mathsf{L}})$, and let $\delta > 0$ be given. From Proposition 3, it may be assumed that the eavesdropper is aware of the type of the source block Q_X. Moreover, from Lemma 13, it may be assumed that $\mathcal{S}_n$ satisfies the three properties in Lemma 13 for all Q_X such that $D(Q_X||P_X) < \mathsf{E}_{\mathsf{L}}$. Specifically, the first property implies that for some *rate-function* $\rho : \mathcal{P}(\mathcal{X}) \to \mathbb{R}_+$ the code $\mathcal{S}_n$ has a fixed rate $r_n(\mathbf{x}) = \rho(Q_X)$ for all $\mathbf{x} \in \mathcal{T}_n(Q_X)$, and $\rho(Q_X) \le \overline{R}(\mathcal{S}, Q_X) + \delta$, as long as $D(Q_X||P_X) < \mathsf{E}_{\mathsf{L}}$.

Let us first focus on a type Q_X that satisfies $D(Q_X||P_X) < \mathsf{E}_{\mathsf{L}}$, and a specific (type-aware) eavesdropper for $\mathcal{S}_n$. The eavesdropper first produces a guess $\hat{\mathbf{u}}$ of the key-bits $\mathbf{u}$ (with a uniform probability over $\{0,1\}^{n\rho(Q_X)}$, and then decodes $\hat{\mathbf{w}} = \varphi_n(y, \hat{\mathbf{u}})$. Since $d_{\mathsf{E}}(\cdot,\cdot)$ is more lenient than $d_{\mathsf{L}}(\cdot,\cdot)$, and $\mathsf{D}_{\mathsf{E}} \ge \mathsf{D}_{\mathsf{L}}$, there exists a $\hat{\mathbf{z}} \in \mathcal{Z}^n$ such that

$$\{\mathbf{x} \in \mathcal{X}^n : d_{\mathsf{L}}(\mathbf{x}, \hat{\mathbf{w}}) \le \mathsf{D}_{\mathsf{L}}\} \subseteq \{\mathbf{x} \in \mathcal{X}^n : d_{\mathsf{E}}(\mathbf{x}, \hat{\mathbf{z}}) \le \mathsf{D}_{\mathsf{L}}\} \tag{289}$$

$$\subseteq \{\mathbf{x} \in \mathcal{X}^n : d_{\mathsf{E}}(\mathbf{x}, \hat{\mathbf{z}}) \le \mathsf{D}_{\mathsf{E}}\}, \tag{290}$$

and so the final eavesdropper estimate is $\mathbf{z} = \hat{\mathbf{z}}$. For any n, let us bound the resulting conditional exiguous-distortion probability.

$$\mathbb{P}\left[d_{\mathsf{E}}(\mathbf{X}, \hat{\mathbf{Z}}) \le \mathsf{D}_{\mathsf{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \ge \mathbb{P}\left[\hat{\mathbf{U}} = \mathbf{U} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \times$$

$$\mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\hat{\mathbf{Z}}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X), \hat{\mathbf{U}} = \mathbf{U}\right] \tag{291}$$

$$\geq 2^{-n\rho(Q_X)} \cdot \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\hat{\mathbf{Z}}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X), \hat{\mathbf{U}} = \mathbf{U}\right] \tag{292}$$

$$\geq 2^{-n\rho(Q_X)} \cdot \mathbb{P}\left[d_{\mathrm{L}}(\mathbf{X},\mathbf{W}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{293}$$

$$\stackrel{(a)}{=} 2^{-n\rho(Q_X)} \tag{294}$$

where (a) is from the second property assured for $\mathcal{S}$ in Lemma 13.

We now analyze the exiguous-distortion probability of $\mathcal{S}$. Since $|\mathcal{P}_n(\mathcal{X})| \leq (n+1)^{|\mathcal{X}|}$

$$p_d(\mathcal{S}_n, \mathsf{D}_{\mathrm{E}}) = \sum_{Q_X \in \mathcal{P}_n(\mathcal{X})} \mathbb{P}\left[\mathbf{X} \in \mathcal{T}_n(Q_X)\right] \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{295}$$

$$\doteq \max_{Q_X \in \mathcal{P}_n(\mathcal{X})} e^{-nD(Q_X||P_X)} \cdot \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{296}$$

$$= \exp\left(-n \cdot \min_{Q_X \in \mathcal{P}_n(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) - \right.\right. \tag{297}$$

$$\left.\left. \frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\}\right) \tag{298}$$

Now, let $0 < \epsilon < \mathsf{E}_{\mathrm{L}}$ be given, and let $Q_X^* \in \mathcal{P}(\mathcal{X})$ be such that

$$D\left(Q_X^*||P_X\right) + \limsup_{n\to\infty} \left\{ -\frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X^*)\right] \right\} \leq$$

$$\inf_{Q_X \in \mathcal{P}(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) + \limsup_{n\to\infty} \left\{ -\frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\} \right\} + \epsilon \tag{299}$$

and let m_0 be sufficiently large so that

$$\sup_{n>m_0} \left\{ -\frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X^*)\right] \right\}$$

$$\leq \limsup_{n\to\infty} \left\{ -\frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X^*)\right] \right\} + \epsilon. \tag{300}$$

Then,

$$\mathcal{E}_d^+(\mathcal{S}, \mathsf{D}_{\mathrm{E}}) = \limsup_{n\to\infty} \min_{Q_X \in \mathcal{P}_n(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) - \frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\} \tag{301}$$

$$= \lim_{m\to\infty} \sup_{n\geq m} \min_{Q_X \in \mathcal{P}_n(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) - \frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\} \tag{302}$$

$$\stackrel{(a)}{=} \lim_{m\to\infty} \sup_{n\geq m} \inf_{Q_X \in \mathcal{P}(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) - \frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\} \tag{303}$$

$$\leq \sup_{n\geq m_0} \inf_{Q_X \in \mathcal{P}(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) - \frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\} \tag{304}$$

$$\leq \inf_{Q_X \in \mathcal{P}(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) + \sup_{n\geq m_0} \left\{ -\frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\} \right\} \tag{305}$$

$$\leq \left\{ D\left(Q_X^*||P_X\right) + \sup_{n>m_0} \left\{ -\frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathrm{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathrm{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X^*)\right] \right\} \right\} \tag{306}$$

$$\overset{(b)}{\leq} \inf_{Q_X \in \mathcal{P}(\mathcal{X})} \Bigg\{ D\left(Q_X||P_X\right) + \tag{307}$$

$$\limsup_{n\to\infty} \left\{ -\frac{1}{n} \log \max_{\tilde{\sigma}_n \in \tilde{\Sigma}_n} \mathbb{P}\left[d_{\mathsf{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathsf{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] \right\} \Bigg\} + 2\epsilon \tag{308}$$

$$= \inf_{Q_X \in \mathcal{P}(\mathcal{X})} \left\{ D\left(Q_X||P_X\right) + \mathcal{E}_d^+(\mathcal{S}, \mathsf{D}_{\mathsf{E}}, Q_X) \right\} + 2\epsilon \tag{309}$$

$$\leq \inf_{Q_X \in \mathcal{P}(\mathcal{X}): D(Q_X||P_X) < \mathsf{E}_{\mathsf{L}}} \left\{ D\left(Q_X||P_X\right) + \mathcal{E}_d^+(\mathcal{S}, \mathsf{D}_{\mathsf{E}}, Q_X) \right\} + 2\epsilon \tag{310}$$

$$\overset{(c)}{\leq} \inf_{Q_X \in \mathcal{P}(\mathcal{X}): D(Q_X||P_X) < \mathsf{E}_{\mathsf{L}}} \left\{ D\left(Q_X||P_X\right) + \mathcal{E}_d^+(\mathcal{S}, \mathsf{D}_{\mathsf{E}}, Q_X) \right\} + 2\epsilon + \delta \tag{311}$$

$$\overset{(d)}{\leq} \inf_{Q_X \in \mathcal{P}(\mathcal{X}): D(Q_X||P_X) < \mathsf{E}_{\mathsf{L}}} \left\{ D\left(Q_X||P_X\right) + \rho(Q_X) \right\} + 2\epsilon + \delta \tag{312}$$

$$\overset{(e)}{\leq} \mathsf{R} + 2\epsilon + 4\delta, \tag{313}$$

where (a) is because, by assumption, if $\mathcal{T}_n(Q_X)$ is empty then $\mathbb{P}\left[d_{\mathsf{E}}(\mathbf{X},\mathbf{Z}) \leq \mathsf{D}_{\mathsf{E}} | \mathbf{X} \in \mathcal{T}_n(Q_X)\right] = 0$, (b) is from (299) and (300), and (c) is from the third property of $\mathcal{S}$ promised by Lemma 13. The passage (d) follows from (294), and so it remains to prove (e). To this end, recall that $\mathbb{E}[r_n(\mathbf{X})] \leq \mathsf{R}$ for all n was assumed. Define, for $0 < \epsilon < \mathsf{E}_{\mathsf{L}}$, the *typical set*

$$\tilde{\mathcal{T}}(P_X, \epsilon) \triangleq \left\{ Q_X \in \mathcal{P}(\mathcal{X}) : D(Q_X||P_X) \leq \epsilon \right\}, \tag{314}$$

and with a slight abuse of notation, define $\tilde{\mathcal{T}}_n(P_X, \epsilon) \triangleq \tilde{\mathcal{T}}(P_X, \epsilon) \cap \mathcal{P}_n(\mathcal{X})$. Then, by the law of large numbers

$$\lim_{n\to\infty} \sum_{Q_X \in \tilde{\mathcal{T}}_n(P_X,\epsilon)} \mathbb{P}\left[\mathbf{X} \in \mathcal{T}_n(Q_X)\right] = 1. \tag{315}$$

Now, assume by contradiction, that for all $Q_X \in \tilde{\mathcal{T}}(P_X, \epsilon)$ we have $\rho(Q_X) \geq \mathsf{R} + 3\delta$. Since by construction $\rho(Q_X) \leq \overline{R}(\mathcal{S}, Q_X) + \delta$, the uniform convergence of $\mathbb{E}[r_n(\mathbf{X})|\mathbf{X} \in \mathcal{T}_n(Q_X)]$ to $\overline{R}(\mathcal{S}, Q_X)$ (see (11) and the discussion that follows) implies that there exists n_0 such that for all $n > n_0$

$$\begin{aligned} \mathbb{E}[r_n(\mathbf{X})|\mathbf{X} \in \mathcal{T}_n(Q_X)] &\geq \overline{R}(\mathcal{S}, Q_X) - \delta \\ &\geq \rho(Q_X) - 2\delta \\ &\geq \mathsf{R} + \delta, \end{aligned} \tag{316}$$

for all $Q_X \in \tilde{\mathcal{T}}_n(P_X, \epsilon)$. So, from (315), there exists n_1, such that for all $n > n_1$ we have that $\mathbb{P}\left[\mathbf{X} \in \tilde{\mathcal{T}}_n(P_X, \epsilon)\right] \geq \frac{1}{1+\delta/2\cdot\log|\mathcal{X}|}$, and then for all $n > \max\{n_0, n_1\}$

$$\mathbb{E}\left[r_n(\mathbf{X})\right] = \sum_{Q_X \in \mathcal{P}_n(\mathcal{X})} \mathbb{P}\left[\mathbf{X} \in \mathcal{T}_n(Q_X)\right] \cdot \mathbb{E}[r_n(\mathbf{X})|\mathbf{X} \in \mathcal{T}_n(Q_X)] \tag{317}$$

$$\geq \sum_{Q_X \in \tilde{\mathcal{T}}_n(P_X,\epsilon)} \mathbb{P}\left[\mathbf{X} \in \mathcal{T}_n(Q_X)\right] \cdot \mathbb{E}[r_n(\mathbf{X})|\mathbf{X} \in \mathcal{T}_n(Q_X)] \tag{318}$$

$$\geq \left(\min_{Q_X \in \tilde{\mathcal{T}}_n(P_X,\epsilon)} \mathbb{E}[r_n(\mathbf{X})|\mathbf{X} \in \mathcal{T}_n(Q_X)] \right) \cdot \sum_{Q_X \in \tilde{\mathcal{T}}_n(P_X,\epsilon)} \mathbb{P}\left[\mathbf{X} \in \mathcal{T}_n(Q_X)\right] \tag{319}$$

$$\overset{(a)}{\geq} (\mathsf{R}+\delta)\frac{1}{1+\delta/2\cdot\log|\mathcal{X}|} \tag{320}$$

$$> (\mathsf{R}+\delta)\frac{1}{1+\delta/\mathsf{R}} \tag{321}$$

$$= \mathsf{R}, \tag{322}$$

where (a) follows from (316). However, this is a contradiction to the fact that $\mathcal{S}_n$ satisfies $\mathbb{E}\left[r_n(\mathbf{X})\right] \leq \mathsf{R}$ for all n. Thus, there must exist $Q_X \in \tilde{\mathcal{T}}(P_X,\epsilon) \subset \tilde{\mathcal{T}}(P_X,\mathsf{E}_{\text{L}})$ such that $\rho(Q_X) < \mathsf{R}+3\delta$, which directly leads to (e) in (313). Since $\epsilon > 0$ and $\delta > 0$ are arbitrary, the first term in the upper bound of (18) is proved, i.e. $\mathcal{E}_d^+(\mathcal{S},\mathsf{D}_{\text{E}}) \leq \mathsf{R}$.

To prove the second term in the upper bound of (18), i.e. $\mathcal{E}_d^+(\mathcal{S},\mathsf{D}_{\text{E}}) \leq E_e^*(\mathsf{D}_{\text{E}})$, note that the eavesdropper can always ignore the cryptogram and *blindly* choose its estimate $\mathbf{z}$ (based only on the type Q_X). Thus, by similar arguments leading to (229), it can be shown that for all n sufficiently large

$$\mathcal{E}_d^+(\mathcal{S},\mathsf{D}_{\text{E}},Q_X) \leq R_{\text{E}}(Q_X,\mathsf{D}_{\text{E}}). \tag{323}$$

The method of types, as in (297) and the definition of $E_e^*(\mathsf{D}_{\text{E}})$ in (16), complete the proof. ■

REFERENCES

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[2] ——, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423,623–656, 1948.

[3] A. Wyner, "The wire-tap channel," *Bell System Technical Journal, The*, vol. 54, no. 8, pp. 1355–1387, October 1975.

[4] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *Information Theory, IEEE Transactions on*, vol. 24, no. 3, pp. 339–348, May 1978.

[5] D. Gunduz, E. Erkip, and H. Poor, "Secure lossless compression with side information," in *Information Theory Workshop, 2008. ITW '08. IEEE*, May 2008, pp. 169–173.

[6] ——, "Lossless compression with security constraints," in *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, July 2008, pp. 111–115.

[7] N. Merhav, "Shannon's secrecy system with informed receivers and its application to systematic coding for wiretapped channels," *Information Theory, IEEE Transactions on*, vol. 54, no. 6, pp. 2723–2734, June 2008.

[8] M. Hellman, "An extension of the Shannon theory approach to cryptography," *Information Theory, IEEE Transactions on*, vol. 23, no. 3, pp. 289–294, May 1977.

[9] H. Yamamoto, "Rate-distortion theory for the Shannon cipher system," *Information Theory, IEEE Transactions on*, vol. 43, no. 3, pp. 827–835, May 1997.

[10] S. C. Lu, "Random ciphering bounds on a class of secrecy systems and discrete message sources," *Information Theory, IEEE Transactions on*, vol. 25, no. 4, pp. 405–414, July 1979.

[11] C. Schieler and P. Cuff, "Secrecy is cheap if the adversary must reconstruct," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, July 2012, pp. 66–70.

[12] ——, "Rate-distortion theory for secrecy systems," *Information Theory, IEEE Transactions on*, vol. 60, no. 12, pp. 7584–7605, December 2014.

[13] P. Cuff, "Using a secret key to foil an eavesdropper," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, September 2010, pp. 1405–1411.

[14] C. Schieler and P. Cuff, "The henchman problem: Measuring secrecy by the minimum distortion in a list," in *Information Theory (ISIT), 2014 IEEE International Symposium on*, June 2014, pp. 596–600.

[15] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

[16] R. Ahlswede and G. Dueck, "Bad codes are good ciphers," *Problems of Control and Information Theory*, vol. 11, no. 5, 1982.

[17] N. Merhav, "A large-deviations notion of perfect secrecy," *Information Theory, IEEE Transactions on*, vol. 49, no. 2, pp. 506–508, February 2003.

[18] ——, "On the Shannon cipher system with a capacity-limited key-distribution channel," *Information Theory, IEEE Transactions on*, vol. 52, no. 3, pp. 1269–1273, March 2006.

[19] E. Haroutunian and A. Ghazaryan, "On the Shannon cipher system with a wiretapper guessing subject to distortion and reliability requirements," in *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, June-July 2002, pp. 324–.

[20] E. Arikan and N. Merhav, "Guessing subject to distortion," *Information Theory, IEEE Transactions on*, vol. 44, no. 3, pp. 1041–1056, May 1998.

[21] N. Merhav and E. Arikan, "The Shannon cipher system with a guessing wiretapper," *Information Theory, IEEE Transactions on*, vol. 45, no. 6, pp. 1860–1866, September 1999.

[22] E. Haroutunian, "On the Shannon cipher system with a wiretapper guessing subject to distortion and reliability requirements," August 2010, available online: http://arxiv.org/pdf/1008.0961.pdf.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[24] K. Marton, "Error exponent for source coding with a fidelity criterion," *Information Theory, IEEE Transactions on*, vol. 20, no. 2, pp. 197–199, March 1974.

[25] M. V. Burnashev, "Data transmission over a discrete channel with feedback: Random transmission time," *Problems of Information transmission*, pp. 250–265, 1976.

[26] B. Nakiboglu and L. Zheng, "Errors-and-erasures decoding for block codes with feedback," *Information Theory, IEEE Transactions on*, vol. 58, no. 1, pp. 24–49, January 2012.

[27] N. Weinberger and N. Merhav, "Optimum trade-offs between the error exponent and the excess-rate exponent of variable-rate Slepian-Wolf coding," *Information Theory, IEEE Transactions on*, vol. 61, no. 4, pp. 2165–2190, April 2015, extended version available online: http://arxiv.org/pdf/1401.0892v3.pdf.

[28] M. Loève, *Probability Theory I*. Springer, 1977.

[29] R. Ahlswede, "Coloring hypergraphs: A new approach to multi-user source coding, part II," *Journal of Combinatorics*, vol. 5, pp. 220–268, 1980.

[30] W. Rudin, *Principles of mathematical analysis*, 3rd ed. McGraw-Hill New York, 1976.

[31] N. Merhav, "Statistical physics and information theory," *Foundations and Trends in Communications and Information Theory*, vol. 6, no. 1-2, pp. 1–212, 2009.