Optimum Tradeoffs Between the Error Exponent and the Excess-Rate Exponent of Variable-Rate Slepian–Wolf Coding

Nir Weinberger, Student Member, IEEE, and Neri Merhav, Fellow, IEEE

Abstract—We analyze the optimal tradeoff between the error exponent and the excess-rate exponent for variable-rate Slepian–Wolf codes. In particular, we first derive upper (converse) bounds on the optimal error and excess-rate exponents, and then lower (achievable) bounds, via a simple class of variable-rate codes which assign the same rate to all source blocks of the same type class. Then, using the exponent bounds, we derive bounds on the optimal rate functions, namely, the minimal rate assigned to each type class, needed in order to achieve a given target error exponent. The resulting excess-rate exponent is then evaluated. Iterative algorithms are provided for the computation of both bounds on the optimal rate functions and their excess-rate exponents. The resulting Slepian–Wolf codes bridge between the two extremes of fixed-rate coding, which has minimal error exponent and maximal excess-rate exponent, and average-rate coding, which has maximal error exponent and minimal excess-rate exponent.

Index Terms—Slepian-Wolf coding, variable-rate coding, buffer overflow, excess-rate exponent, error exponent, reliability function, random-binning, alternating minimization.

I. INTRODUCTION

THE problem of distributed encoding of correlated sources has been studied extensively since the seminal paper of Slepian and Wolf [40]. That paper addresses the case, where a memoryless source \(\{(X_i, Y_i)\}\) needs to be compressed by two separate encoders, one for \(X_i\) and one for \(Y_i\). In a nutshell, the most significant result of [40] states that if \(Y_i\) is known at the decoder side, then \(X_i\) can be compressed at the rate of the conditional entropy of \(X_i\) given \(Y_i\). Since this is the minimal rate even for the case where \(Y_i\) is known also to the encoder, then no rate loss is incurred by the lack of knowledge of \(Y_i\) at the encoder. Early research has focused on asymptotic analysis of the decoding error probability for the ensemble of random binning codes. Gallager [21] has adapted his well known analysis techniques from random channel coding [19, Secs. 5.5-5.6] to the random binning ensemble of distributed source coding. Later, it was shown in [13] and [16] that the universal minimum conditional entropy decoder also achieves the same exponent. Expurgated error exponents were given in [14], assuming optimal decoding (non-universal). In [2, Appendix I], Ahlswede has shown the achievability of random binning and expurgated bounds via codebooks generated by permutations of good channel codes. The expurgated exponent analysis was then generalized to coded side information in [11] (with linear codes) and [36].

In all the above papers, fixed-rate coding was assumed, perhaps because, as is well known, Slepian-Wolf (SW) coding is, in some sense, analogous to channel coding (without feedback) [2], [5], [8], [14], for which variable-rate is usually of no use. More recently, it was recognized that variable-rate SW coding may have improved performance. For example, it was shown in [22], [27], and [38] that variable-rate SW codes might have lower redundancy (additional rate beyond the conditional entropy, for a given error probability). Other results on variable-rate coding can be found in [26], [28], and [33]. In another line of work, which is more relevant to this paper, it was observed that variable-rate coding under an average rate constraint [6], [7], [9] outperforms fixed-rate coding in terms of error exponents. The intuitive reason is that the empirical probability mass function (PMF) of the source tends to concentrate exponentially fast around the true PMF, and so in order to asymptotically satisfy an average rate constraint \(\bar{R}\), it is only required that the rates allocated to typical source blocks would have rate less than \(\bar{R}\) (see [6, Th. 1]). Other types, distant from the type of the source, can be assigned with arbitrary large rates, and thus effectively may be sent uncoded.

The expected value of the rate is, however, a rather soft requirement, and it provides a meaningful performance measure only in the case of many system uses, where the random rate concentrates around its expected value. Consider, for example, an on-line compression scheme, in which the codeword is buffered at the encoder before transmitted [23], [24]. If the instantaneous codeword length is larger than the buffer size, then the buffer overflows. If the decoder is aware of this event (using a dedicated feed-forward channel, e.g.) then this is an erasure event, and so, it is desirable to minimize this probability, while maintaining some given error probability. In a different case, the buffer length might be larger than the maximal codeword length, but the buffer is also used for other purposes (e.g., sending status data). If the data codewords have priority over all other uses, then it is desirable to minimize the occasions of blocking other usage of the buffer. This motivates...
us to take a somewhat different approach and address a more refined figure of merit for the rate. Specifically, we will be interested in the probability that the rate exceeds a certain threshold. While the aforementioned average-rate coding increases error exponents, its excess-rate probability is clearly inferior to that of fixed-rate coding.

It should be mentioned that in many other problems in information theory, instead of considering the average value of some cost (which for SW coding is the rate) more refined figures of merit are imposed, such as the excess probability or higher moments. Beyond lossless compression, which was mentioned above, other examples include excess distortion [12], [30], variable-rate channel coding with feedback [35], list size of a list decoder [41], and estimation [32] (see also [31] for intimately related problem of minimizing exponential moments of a cost function, and many references therein).

In this paper, we systematically analyze the trade-off between excess-rate exponent and the error exponent. Based on the analogy of SW and channel coding, we provide upper (converse) bounds on the error and excess-rate exponents of a general SW code. Then, we derive lower (achievability) bounds for a special class of SW codes, which assign the same coding rate to all source blocks of the same type class. The bounds on error exponents may be considered as a generalization of the error exponents of [6], [7], and [9] to the case where excess-rate performance is of importance. As will be seen, this requires a joint treatment of all possible types of the source at the same time, and not just the type of the source, as in average-rate coding. Both bounds will initially be expressed via fixed-composition reliability functions of channel codes (to be defined in the sequel), and only afterwards, specific known bounds (random coding, expurgated and sphere packing) on the reliability functions will be applied. This links the question whether assigning equal rates to source blocks of the same type class is asymptotically optimal, to the unsettled gap between the infimum and supremum reliability functions (to be also defined in the sequel) [15, Problem 10.7]. Whenever it can be verified that no gap exists, then assigning equal rates to types is optimal. However, similarly as in channel coding above the critical rate, where the reliability function is known exactly, the upper and lower bounds of SW exponents coincide for small error exponents, and then assigning equal rates to types class is optimal. Afterwards, for every type class, bounds on the minimal encoding rate, required to meet a prescribed value of error exponent, will be found, and corresponding bounds on the resulting excess-rate performance of the system will be derived. Since the computation of both the rate for a given type, and the excess-rate exponent, lead to optimization problems that lack closed-form solutions, we will provide explicit iterative algorithms that converge to the optimal solutions.

The outline of the remaining part of the paper is as follows. In Section II, we establish notation conventions and formulate SW codes. We also formulate channel codes and provide background of known results, which are useful for the analysis of SW codes. In Section III, we derive upper and lower bounds on the error exponent and excess-rate exponent of general SW codes, and discuss the trade-off between the two exponents. Then, in Section IV, we characterize the optimal rate allocation (in a sense that will be made precise), under an error exponent constraint, and in Section V, we analyze the resulting excess-rate exponent. In Section VI, we discuss computational aspects of the bounds on the optimal rate allocation, as well as the bounds on the optimal excess-rate exponent. Section VII demonstrates the results via a numerical example, and Section VIII summarizes the paper, along with directions for further research. Almost all proofs are deferred to Appendix A. If this is not the case, then the proof either appears right after the statement of the theorem or lemma, or a pointer is provided to an extended version of this paper [43]. The extended paper [43] also provides more detailed discussion on some topics. Specifically, [43, Appendix B] thoroughly discusses the fixed-composition reliability function of channel coding (see Subsection II-C) which is central to the analysis of SW codes. Finally, in Appendix B, we provide approximations of the optimal rate-function bounds, for the case of very weakly correlated sources.

II. PROBLEM FORMULATION

A. 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 their conditionings, if applicable. We will follow the standard notation conventions, e.g., \( P_X(x) \) will denote the \( X \)-marginal of \( P \), \( P_{Y|X}(y|x) \) will denote the conditional distribution of \( Y \) given \( X \), \( P_{XY}(x,y) \) will denote the joint distribution, and so on. The arguments will be omitted when we address the entire PMF, e.g., \( P_X \), \( P_{Y|X} \) and \( P_{XY} \).

Similarly, generic sources will be denoted by \( Q \), \( Q^* \), and in other forms, again subscripted by the relevant random variables/vectors/conditionings. The joint distribution induced by a PMF \( Q_X \) and conditional PMF \( Q_{Y|X} \) will be denoted by \( Q_X \times Q_{Y|X} \), and its \( Y \)-marginal will be denoted by \( (Q_X \times Q_{Y|X})_Y \), or simply by \( Q_Y \) when understood from context. An exceptional case will be the ‘hat’ notation. For this notation, \( \hat{Q}_X \) will denote the empirical distribution of a vector \( x \in \mathcal{X} \), i.e., the vector of relative frequencies \( \hat{Q}_X(x) \) of each symbol \( x \in \mathcal{X} \) in \( x \). The type class of \( x \in \mathcal{X} \), which will be denoted by \( T_n(\hat{Q}_X) \), is the set of all vectors \( x' \) with \( \hat{Q}_X = \hat{Q}_{x'} \).

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 $\mathcal{P}(X) \triangleq \bigcup_{n=1}^{\infty} \mathcal{P}_n(X)$. Similar notation for type classes will also be used for generic types $Q_X \in \mathcal{P}(X)$, i.e. $T_n(Q_X)$ will denote the set of all vectors $x$ with $Q_X = Q_x$.

The probability simplex for $X$ will be denoted by $\mathcal{Q}(X)$, and the simplex for the alphabet $X \times Y$ will be denoted by $\mathcal{Q}(X \times Y)$. The support of a PMF $Q_X$ will be denoted by $\text{supp}(Q_X) \triangleq \{x : Q_X(x) \neq 0\} \subseteq X$. For two PMFs $P_X, Q_X$ over the same finite alphabet $X$, we will denote the variation distance ($L_1$ norm) by

$$||P_X - Q_X|| \triangleq \sum_{x \in X} |P_X(x) - Q_X(x)|. \tag{1}$$

When optimizing a function of a distribution $Q_X$ over the entire probability simplex $\mathcal{Q}(X)$, 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 $\min_{Q \in \mathcal{Q}(X)} f(Q)$. The same will hold for optimization of a function of a distribution $Q_{XY}$ over the probability simplex $\mathcal{Q}(X \times Y)$.

The expectation operator with respect to (w.r.t.) a given distribution, e.g. $Q_{XY}$, will be denoted by $E_{Q_{XY}}[\cdot]$ where, again, the subscript will be omitted if the underlying probability distribution is clear from the context. The entropy of a given distribution, e.g. $Q_X$, will be denoted by $H(Q_X)$, and the binary entropy function will be denoted by $h_B(q)$ for $0 \leq q \leq 1$. The average conditional entropy of $Q_{Y|X}$ w.r.t. $Q_X$ will be denoted by $H(Q_{Y|X}|Q_X) \triangleq \sum_{x \in X} Q_X(x) H(Q_{Y|X}(\cdot|x))$, and the mutual information of a joint distribution $Q_{XY}$ will be denoted by $I(Q_{XY})$. The information divergence between two distributions, e.g. $P_{XY}$ and $Q_{XY}$, will be denoted by $D(P_{XY}||Q_{XY})$ and the average divergence between $Q_{Y|X}$ and $P_{Y|X}$ w.r.t. $Q_X$ will be denoted by $D(Q_{Y|X}||P_{Y|X}|Q_X) \triangleq \sum_{x \in X} Q_X(x) D(Q_{Y|X}(\cdot|x)||P_{Y|X}(\cdot|x))$. In all the information measures above, the PMF may also be an empirical PMF, for example, $H(\hat{Q}_X)$, and so on.

We will denote the Hamming distance between two vectors $x \in \mathcal{X}^n$ and $z \in \mathcal{X}^n$ by $d_H(x, z)$. The length of a string $b$ will be denoted by $|b|$, and the concatenation of the strings $b_1, b_2, \ldots$ will be denoted by $(b_1, b_2, \ldots)$. For a set $A$, we will denote its complement by $\bar{A}$, its closure by $\overline{A}$, its interior by $\text{int}(A)$, and its boundary by $\partial A$. If the set $A$ is finite, we will denote its cardinality by $|A|$. The probability of the event $A$ will be denoted by $\mathbb{P}(A)$, and $\mathbb{I}(A)$ will denote the indicator function of this event.

For two positive sequences, $\{a_n\}$ and $\{b_n\}$ the notation $a_n \asymp b_n$ will mean asymptotic equivalence in the exponential scale, that is, $\lim_{n \to \infty} 1/n \log \left( a_n / b_n \right) = 0$. Similarly, $a_n \preceq b_n$ will mean $\limsup_{n \to \infty} 1/n \log \left( b_n / a_n \right) = 0$, and so on. For a function $\lceil r \rceil$ will be defined as $\text{max} \{r, 0\}$, and $\lfloor r \rfloor$ will denote the ceiling function. For two integers, $a, b$, we denote by $a \mod b$ the modulo of $a$ w.r.t. $b$. Unless otherwise stated, logarithms and exponents will be understood to be taken to the natural base.

**B. Slepian-Wolf Coding**

Let $\{(X_i, Y_i)\}_{i=1}^{n}$ be $n$ independent copies of a pair of random variables $(X, Y)$. We assume that $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are finite alphabets, are distributed according to $P_{XY}(x, y) = \mathbb{P}(X = x, Y = y)$. It is assumed that $\text{supp}(P_X) = \mathcal{X}$ and that $\text{supp}(P_Y) = \mathcal{Y}$, otherwise, remove the irrelevant letters from their alphabet. We say that the conditional distribution $P_{Y|X}$ is not noiseless if there exists a pair of input letters $x, x' \in \mathcal{X}$ and an output letter $y \in \mathcal{Y}$ such that $P_{Y|X}(y|x) P_{Y|X}(y|x') > 0$, and assume this property for $P_{X|Y}$, and so, $H(P_{XY}|P_Y) > 0$.

A SW code $S_n$ for sequences of length $n$ is defined by a prefix code with encoder

$$s_n: \mathcal{X}^n \to \{0, 1\}^n \tag{2}$$

and a decoder

$$\sigma_n: \{0, 1\}^n \times \mathcal{Y}^n \to \mathcal{X}^n \tag{3}$$

where $\{0, 1\}^n$ is the set of all finite length binary strings. The encoder maps a source block $x$ into a binary string $s_n(x) \in \{0, 1\}^n$, where for $b \in \{0, 1\}^n$, the inverse image of $b$ is defined as

$$s_n^{-1}(b) \triangleq \{x \in \mathcal{X}^n : s_n(x) = b\} \tag{4}$$

and it is called a *bin*. The decoder $\sigma_n$, which observes $b = s_n(x)$ and the side information $y$, has to decide on the particular source block $x \in s_n^{-1}(b)$ to obtain a decoded source block $\hat{x} \triangleq \sigma_n(s_n(x), y)$. A sequence of SW codes $\{S_n\}_{n \geq 1}$, indexed by the block length $n$, will be denoted by $S$.

The error probability, for a given code $S_n = \{s_n(x), \sigma_n\}$, is denoted by $p_e(S_n) \triangleq \mathbb{P}(X \neq \hat{X})$. The *infimum error exponent* achieved for a sequence of codes $S$ is defined as

$$E_e^-(S) \triangleq \liminf_{n \to \infty} -\frac{1}{n} \log p_e(S_n) \tag{5}$$

and the *supremum error exponent* achieved is defined as

$$E_e^+(S) \triangleq \limsup_{n \to \infty} -\frac{1}{n} \log p_e(S_n). \tag{6}$$

While clearly, $E_e^-(S) \leq E_e^+(S)$, it is guaranteed that $p_e(S_n) \geq \exp \left[ -n E_e^+(S) \right]$ for all sufficiently large block lengths, while $p_e(S_n) = \exp \left[ -n E_e^-(S) \right]$ may hold only for some sub-sequence of block lengths. Thus, $E_e^+(S)$ is more robust to the choice of block length.

For a given $Q_X \in \mathcal{P}(X)$, we define, with a slight abuse of notation, the conditional infimum error exponent as

$$E_e^-(S, Q_X) \triangleq \liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{P}(X \neq \hat{X} | X \in T_n(Q_X)) \tag{7}$$

where we define $\mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) = 0$ if $T_n(Q_X)$ is empty. $E_e^+(S, Q_X)$ is defined analogously.

The coding rate of $x \in \mathcal{X}^n$ is defined as $r(x) \triangleq \log |x| / \log_2 e$. A SW code is termed a *fixed-rate* code of rate $R_0$ if $r(x) = R_0$ for all $x \in \mathcal{X}^n$. Otherwise it is called a *variable-rate* code, and has an *average rate* $\mathbb{E}[r(X)]$. We define the conditional rate of $Q_X \in \mathcal{P}(X)$ as

$$R(Q_X; S) \triangleq \limsup_{n \to \infty} \mathbb{E}[r(X) | X \in T_n(Q_X)], \tag{8}$$

where $\mathbb{E}[r(X) | X \in T_n(Q_X)] \triangleq 0$ if $T_n(Q_X)$ is empty. Since $r(x) = \log |x|$ allows the encoding of $x$ with zero error, it will be assumed that $R(Q_X; S)$ is finite. For a given
target rate $R$, the excess-rate probability, of a code $S_n$, is denoted by $p_e(S_n, R) \triangleq \mathbb{P}[r(X) \geq R]$, and the excess-rate exponent function, achieved for a sequence of codes $S$, is defined as

$$E_e(S, R) \triangleq \liminf_{n \to \infty} -\frac{1}{n} \log p_e(S_n, R).$$

(9)

For a given $Q_X \in \mathcal{P}(\mathcal{X})$, we define, with a slight abuse of notation, the conditional excess-rate exponent as

$$E_e(S, R, Q_X) \triangleq \liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[r(X) \geq R|X \in \mathcal{T}_n(Q_X)],$$

(10)

where $\mathbb{P}[r(X) \geq R|X \in \mathcal{T}_n(Q_X)] \triangleq 0$ if $\mathcal{T}_n(Q_X)$ is empty.

Throughout the paper, we will mainly be interested in the following sub-class of variable-rate SW codes.

**Definition 1:** A SW code $S_n$ is termed type-dependent, variable-rate code, if $r(x)$ depends on $x$ only via its type (empirical PMF). Namely, $\hat{Q}_X = \hat{Q}_X(r(x)$ implies $r(x) = r(\hat{x})$. Any finite function $\rho(\cdot) : Q_X \to \mathbb{R}^+$ is called a rate function.

A rate function is termed regular if there exists a constant $d > 0$ and a set $\mathcal{V} \triangleq \{Q_X \in \mathcal{Q}(\mathcal{X}) : D(Q_X||P_X) < d\}$, such that $\rho(\cdot)$ is continuous in $\mathcal{V}$, and equals some constant $R_0$ for $Q_X \in \mathcal{V}$.

The main objective of the paper is to derive the optimal trade-off between the error exponent and the excess-rate exponent, i.e., to find the maximal achievable excess-rate exponent, under a constraint on the error exponent. The subclass of type-dependent, variable-rate SW codes will be shown to achieve the optimal trade-off in a certain range of exponents, and the question of their optimality in other ranges will be discussed.

**C. Channel Coding**

In SW coding, the collection of source words that belong to the same bin, can be considered a channel code, and given the bin index, the SW decoder acts just as a channel decoder (with the exception that the prior probabilities of the source blocks in the bin may not necessarily be uniform). Thus, error exponents of SW codes are intimately related to error exponents of channel codes (see [7], [14]). Accordingly, we next define a few terms associated with channel codes, which will be needed in the sequel.

Consider a discrete memoryless channel $\{W(y|x), x \in \mathcal{X}, y \in \mathcal{Y}\}$ with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$, which are both finite. A channel code $C_n$ of block length $n$ is defined by an encoder

$$f_n : \{1, \ldots, [e^nR]\} \to \mathcal{X}^n$$

(11)

1In the definition of achievable excess-rate exponent, we use only limit inferior. It should be observed that for an operational meaning, the error exponent and excess-rate exponent should be jointly approached by a sub-sequence of block lengths. When the limit inferior is used for both the definition of the error exponent and the definition of the excess-rate exponent, any sufficiently large block length will approach the asymptotic limit of both exponents. When one of the exponents is defined as limit inferior, and the other is defined as limit superior, then there exists a sub-sequence of block lengths which achieve each of the exponents might be completely disjoint.

and a decoder

$$\phi_n : Y^n \to \{1, \ldots, [e^nR]\},$$

(12)

where $R$ is the rate of the code. We say that the channel code is a fixed composition code, if all codewords $\{f_n(m)\}$, $1 \leq m \leq [e^nR]$, belong to a single type class $T_n(Q_X)$. A sequence of channel codes will be denoted by $C = \{c_n\}_{n \geq 1}$. The error probability for a given channel code $C_n = \{f_n, \phi_n\}$ is denoted by $p_e(C_n) \triangleq \mathbb{P}(\phi_n(f_n(M)) \neq M)$, where $M$ is a uniform random variable over the set $\{1, \ldots, [e^nR]\}$. The infimum error exponent, achieved for a given sequence of channel codes $C$ is defined as

$$E^{-}_e(C) \triangleq \liminf_{n \to \infty} -\frac{1}{n} \log p_e(C_n)$$

(13)

and the supremum error exponent is defined as

$$E^{+}_e(C) \triangleq \limsup_{n \to \infty} -\frac{1}{n} \log p_e(C_n).$$

(14)

A number $E_e > 0$ is an achievable infimum (supremum) error exponent for the type $Q_X \in \mathcal{Q}(\mathcal{X})$ and the channel $W$ at rate $R$, if for any $\delta > 0$ there exists a sequence of types $Q_X^{(n)} \in \mathcal{P}_n(\mathcal{X})$ such that $Q_X^{(n)} \to Q_X$, and there exists a sequence of fixed composition channel codes $C_n \subseteq \mathcal{T}_n(Q_X^{(n)})$ with

$$\liminf_{n \to \infty} \frac{\log |C_n|}{n} \geq R - \delta$$

(15)

and $E^-_e(C) \geq E_e - \delta$ (respectively, $E^+_e(C) \geq E_e - \delta$). For a given rate $R$, a type $Q_X$, and a channel $W$, we let $E_e^+(R, Q_X, W)$ ($E_e^-(R, Q_X, W)$) be the largest achievable infimum (respectively, supremum) error exponent over all possible sequences of codes $C$ for the type $Q_X$. The functions $E_e^+(R, Q_X, W)$ and $E_e^-(R, Q_X, W)$ may be interpreted as infimum/supremum fixed-composition reliability functions of the channel $W$, when the type of the codewords must tend to $Q_X$.

We define by $C_0^-(Q_X, W)$ (respectively, $C_0^+(Q_X, W)$) the maximum of all rates such that $E_e^+(R, Q_X, W)$ (respectively, $E_e^-(R, Q_X, W)$) is infinite, which can be regarded as the zero-error capacity of the channel $W$ of fixed composition codes with codebook types which tends to $Q_X$. Fekete’s Lemma [15, Lemma 11.2] implies that $C_0^-(Q_X, W) = C_0^+(Q_X, W)$ and thus we will denote henceforth both quantities by $C_0(Q_X, W)$. Notice that when $W$ is not noiseless and $Q_X \in \mathcal{Q}(\mathcal{X})$, we have $C_0(Q_X, W) = 0$, namely, $C_0(Q_X, W)$ may be strictly positive only for types which belong to $\partial \mathcal{Q}(\mathcal{X})$. For any $Q_X \in \mathcal{Q}(\mathcal{X})$, we define

$$E_0^-(Q_X, W) \triangleq \lim_{R \to C_0(Q_X, W)} E_e^-(R, Q_X, W)$$

(16)

and $E_0^+(Q_X, W)$ is defined analogously.

Unfortunately, it is a long-standing open problem to find the exact values of $E_e^+(R, Q_X, W)$ and $E_e^-(R, Q_X, W)$ for an arbitrary rate $R \in [0, I(Q_X \times W)]$, and it is not even known if $E_e^+(R, Q_X, W) = E_e^-(R, Q_X, W)$ [15, Problem 10.7]. However, the following bounds on the fixed composition reliability function are well known when $Q_X \in \mathcal{P}_n(\mathcal{X})$. The random coding bound [15, Th. 10.2] is a lower bound on the infimum fixed-composition reliability
function, given by
\[
\mathcal{E}_s^*(\mathcal{R}, \mathcal{Q}_X, W) \geq E_{nc}(\mathcal{R}, \mathcal{Q}_X, W) \\
\triangleq \min_{Q_{Y|X}} \left\{ D(Q_{Y|X}||W|Q_X) + [I(Q_X \times Q_{Y|X}) - R]_+ \right\}
\]  
(17)

Similarly, the expurgated lower bound [15, Problem 10.18] is given by
\[
\mathcal{E}_s^*(\mathcal{R}, \mathcal{Q}_X, W) \geq E_{ac}(\mathcal{R}, \mathcal{Q}_X, W) \\
\triangleq \min_{Q_{X}: Q_X = \mathcal{Q}_X} \left\{ B(Q_{X|\tilde{X}}, W) + I(Q_{X|\tilde{X}}) - R \right\},
\]  
(18)

where
\[
B(Q_{X|\tilde{X}}, W) \triangleq \mathbb{E}_{Q_{X|\tilde{X}}}[d_W(X, \tilde{X})]
\]  
(19)

The sphere packing exponent [15, Th. 10.3] is an upper bound on the supremum fixed-composition reliability function and given by
\[
\mathcal{E}_s^*(\mathcal{R}, \mathcal{Q}_X, W) \leq E_{sp}(\mathcal{R}, \mathcal{Q}_X, W) \\
\triangleq \min_{Q_{Y|X}: I(Q_X \times Q_{Y|X}) \leq \mathcal{R}} D(Q_{Y|X}||W|Q_X)
\]  
(21)

which is valid for rates except \( R_{\infty}(\mathcal{Q}_X, W) \), defined as the infimum of all rates such that \( E_{sp}(\mathcal{R}, \mathcal{Q}_X, W) < \infty \).

An improved upper bound on the supremum fixed-composition reliability for low rates, is the straight line bound [15, Problem 10.30], [42, Sec. 3.8]. This bound is obtained by connecting the expurgated bound at \( \mathcal{R} = -\infty \), which is known to be tight [15, Problem 10.21], with the sphere packing bound. Since specifying our results on the optimal rate function (Section IV) and excess-rate exponents (Section V) of SW codes is fairly simple and does not contribute to intuition, we will not discuss this bound henceforth. On the same note, since \( E_{ac}(\mathcal{R}, \mathcal{Q}_X, W) \) and \( E_{sp}(\mathcal{R}, \mathcal{Q}_X, W) \) are not concave in \( \mathcal{Q}_X \), in general, then the error performance for a given fixed composition of type \( \mathcal{Q}_X \) can be improved by a certain time-sharing structure in the random coding mechanism. According to this structure, for each randomly selected codeword, the block length is optimally subdivided into codeword segments that are randomly drawn from optimally chosen types, whose weighted average (with weights proportional to the segment lengths) conforms with the given \( \mathcal{Q}_X \).

At zero-rate, the resulting expurgated error exponent is given by the upper concave envelope (UCE) of \( E_{ac}(0, \mathcal{Q}_X, W) \) [15, Problem 10.22]. Nonetheless, in many cases (see discussion in [25] and [34, Sec. 2]), \( E_{ac}(0, \mathcal{Q}_X, W) \) is already concave, and no improved bound can be obtained by taking the UCE (e.g., when \( |X| = 2 \), \( E_{ac}(0, \mathcal{Q}_X, W) \) is concave). In ordinary channel coding (without input constraints) this improvement is usually not discussed, because the time-sharing structure does increase the maximum of \( E_{nc}(\mathcal{R}, \mathcal{Q}_X, W) \) and \( E_{ac}(\mathcal{R}, \mathcal{Q}_X, W) \) over \( \mathcal{Q}_X \). However, for the utilization of channel codes as components of a SW code the value of \( E_{nc}(\mathcal{R}, \mathcal{Q}_X, W) \) and \( E_{ac}(\mathcal{R}, \mathcal{Q}_X, W) \) at any given \( \mathcal{Q}_X \) is of interest. Nonetheless, for the sake of simplicity of the exposition, throughout the sequel, we will not include this time-sharing mechanism in our discussions and derivations, although their inclusion is conceptually not difficult.

In [6, Proposition 4] these bounds were shown to hold for any \( \mathcal{Q}_X \in \mathcal{Q}(X) \) from continuity arguments. We will use the convention that all the above bounds are formally infinite for negative rates. It can be deduced from the above bounds [15, Corollary 10.4], that there exists a critical rate \( R_{cr}(\mathcal{Q}_X, W) \) such that for \( \mathcal{R} \in [R_{cr}(\mathcal{Q}_X, W), I(\mathcal{Q}_X \times W)] \), \( E_{nc}(\mathcal{R}, \mathcal{Q}_X, W) = E_{sp}(\mathcal{R}, \mathcal{Q}_X, W) \), and consequently \( \mathcal{E}_s^*(\mathcal{R}, \mathcal{Q}_X, W) = \mathcal{E}_e^*(\mathcal{R}, \mathcal{Q}_X, W) \).

III. ERROR AND EXCESS-RATE EXPONENTS

For a SW code, a trade-off exists between the error exponent \( E_e \), the target rate \( \mathcal{R} \), and the excess-rate exponent \( E_e \).

In Subsection III-A, we discuss informally some known results regarding error exponents of fixed-rate SW codes and variable-rate SW codes, under an average rate constraint. We also discuss the excess-rate exponent function that they achieve.

Then, in Subsection III-B, upper bounds (converse results) will be found on the supremum error and excess-rate exponents, and lower bounds (achievability results) on the infimum error and excess-rate exponents will be derived for type-dependent, variable-rate SW codes. It will be apparent that the gap between the lower and upper bounds is only due to the gap which exists, in general, between, the infimum and supremum channel reliability functions. Thus, whenever the channel reliability functions are equal, type-dependent, variable-rate SW codes are optimal. For this reason, as well as their intuitive plausibility, we will later analyze optimal (in a sense that will be made precise) type-dependent, variable-rate SW codes.

A. Previous Work

For a sequence of fixed-rate SW codes \( \mathcal{S} \) at rate \( \mathcal{R}_0 \), the excess-rate exponent function is trivially given by
\[
E_e(\mathcal{S}, \mathcal{R}) = \begin{cases} 
0, & \mathcal{R} \leq \mathcal{R}_0 \\
\infty, & \text{otherwise}
\end{cases}
\]  
(22)

Evidently, this function bears a strong dichotomy between rates below and above \( \mathcal{R}_0 \). Bounds on the error exponents for fixed-rate SW coding were derived in [13, Ths. 2 and 3], [2, Th. 1], and [14, Th. 2]. The analysis is essentially based on considering each type class of the source separately. Loosely speaking, for any given \( \mathcal{Q}_X \in \mathcal{P}_n(\mathcal{X}) \), there exists a partition of the type class \( \mathcal{T}_n(\mathcal{Q}_X) \) into bins, such that every bin corresponds to a channel code of rate \( H(\mathcal{Q}_X) - \mathcal{R}_0 \), which achieves an error exponent function \( E_e^*(H(\mathcal{Q}_X) - \mathcal{R}_0, \mathcal{Q}_X, P_{Y|X}) \). Since \( \mathcal{P}(\mathcal{T}_n(\mathcal{Q}_X)) = \exp \{ -nD(\mathcal{Q}_X||P_X) \} \), and the number of types increases only polynomially, the error exponent is given by
\[
E_e^{-}(\mathcal{S}) \geq \min_{\mathcal{Q}_X} \left\{ D(\mathcal{Q}_X||P_X) + E_e^*(H(\mathcal{Q}_X) - \mathcal{R}_0, \mathcal{Q}_X, P_{Y|X}) \right\},
\]  
(23)

\[2\] We will prove (23) rigorously in Theorem 5.
It was observed in [6] and [7] that sequences of variable-rate SW codes may have better error exponents than those of fixed-rate SW codes, when an average rate constraint is imposed, i.e. \( E[r(X)] \leq R_0 \). Intuitively, since asymptotically, the average rate is only determined by the rate of types \( \{Q_X\} \) that are ‘close’ (in a sense that was made precise in [6, Th. 1] and [7, Th. 1 and 2]) to the source \( P_X \), one can allocate large rates to a-typical source blocks, transmit them uncoded using \( n \log_2 |x| \) bits, and the decoder will have zero-error for source blocks from these type classes. The result is that for such variable-rate SW codes, the supremum error exponent equals the upper bound and a lower bound for \( E^+_{\mathcal{R}}(S) \) can be improved, e.g., by coding each of the source blocks in a type close to \( P_X \), and obtaining the same error exponent (24), and the excess-rate function
\[
E^+_{\mathcal{R}}(S) = E^+_{\mathcal{R}}(H(P_X) - R_0, P_X, P_Y|X). \tag{24}
\]
This can be thought of as a generalization of [2, Th. 1] to variable-rate codes under an average-rate constraint. However, since the probability that \( Q_X \) would be away from \( P_X \) decays with an arbitrary small exponent, the resulting excess-rate exponent function is given by \( E^+_{\mathcal{R}}(S, R) = 0 \) for \( R \leq \log |x| \), which is inferior to the infinite excess-rate exponent of fixed-rate coding for \( R \in (R_0, \log |x|) \). This excess-rate function can be improved, e.g., by coding each of the source blocks in the ‘uncoded type classes’ with \( \log_x |x| \approx nH(Q_X) \) bits, and obtaining the same error exponent (24), and the excess-rate exponent for \( R \in (R_0, \log |x|) \) will be
\[
E^+_{\mathcal{R}}(S, R) = \begin{cases} 0, & R \leq R_0 \\ \min_{H(Q_X) \geq R} D(Q_X || P_X), & R_0 < R \leq \log |x|. \end{cases} \tag{25}
\]
While this excess-rate exponent may be positive for \( R \in (R_0, \log |x|) \), it is nonetheless finite, in contrast to fixed-rate coding (22). In this paper, we will analyze systematically the trade-off between the error and excess-rate exponents for variable-rate codes, where the two above cases, i.e., fixed-rate and variable-rate with average rate constraint, may be considered as two extremes of this trade-off.

Since in [6] and [7] the focus was on coding the source type \( P_X \); the essence of [6, Th. 1] and [7, Ths. 1 and 2] is an upper bound and a lower bound for \( E^+_{\mathcal{R}}(S, P_X) \). Nonetheless, the proofs of these bounds are similar for any given type \( Q_X \in \mathcal{P}(X) \). We will use the following variation of [6, Th. 1].

**Theorem 2 (Variation of [6, Th. 1]):** Let \( S \) be an arbitrary sequence of SW codes. Then, for every \( Q_X \in \mathcal{P}(X) \)
\[
E^+_{\mathcal{R}}(S, Q_X) \leq E^+_{\mathcal{R}}(H(Q_X) - R(Q_X; S), Q_X, P_Y|X). \tag{26}
\]
Also, for any \( Q_X \in \mathcal{P}(X) \) and \( \inf \{Q(X) \} \) there exists a sequence of type-dependent SW codes \( S^* \) with rates \( r^*(x) \), such that for any \( \delta > 0 \) and sufficiently large \( n \), we have
\[
r^*(x) \leq R(Q_X; S) + \delta \quad \text{for all } x \in T_\delta(Q_X) \quad \text{and} \quad E^+(S^*, Q_X) \geq E^+(H(Q_X) - R(Q_X; S), Q_X, P_Y|X) - \delta. \tag{27}
\]

In [6] and [7], the achievability result actually obtained was
\[
E^+_{\mathcal{R}}(S, Q_X) \geq E^+_{\mathcal{R}}(H(Q_X) - R(Q_X; S), Q_X, P_Y|X) - \delta, \tag{28}
\]
and (24) was proved. Before we proceed, we make two comments on the difference in the proof technique from [6, Th. 1] and why the stronger statement (27) is required.

The proof of (27) is based on the additional properties of optimal channel codes stated in Lemma 20 in Appendix A (see [43, Appendix B] for more details). Notice that in channel coding, the type of the fixed composition code is under the control of the code designer. Indeed, according to the definition of the infimum fixed-composition reliability function, to achieve \( E^+(R, Q_X, W) \), one needs to find a sequence of types \( Q_X^{(n)} \to Q_X \) with required error probability, but it is not required that \( Q_X^{(n)} = Q_X \) for all \( n \). In the proof of Theorem 2, we permute channel codes into SW codes, and conditioning on the event that the source block belongs to the type \( Q_X \), we would like to achieve conditional error exponent of \( E^+(R, Q_X, W) \) for the SW code. Thus, for SW coding, the type is determined by the source, and there is no flexibility to choose a ‘nearby’ type, since the type of the source block is not under the control of the code designer. Lemma 20 states that the infimum reliability function can be achieved with codes of type \( Q_X \) exactly, when the block length is sufficiently large.

The reason that (27) is required, is because in the proof of [6, Th. 1] and Theorem 2 a single channel code is constructed and utilized for SW coding of a single type \( Q_X \). By contrast, when considering the more refined notion of excess-rate, all types \( \mathcal{P}(\mathcal{X}) \) of the source should be considered at the same time, and as a result, many channel codes should be constructed (see the proof of Theorem 5 heretofore). Now, consider the simplified case of SW coding for just two different types. In this case, two channel codes are required for a ‘good’ SW coding of the two types. However, if the codes are designed to achieve the supremum reliability function, there is no guarantee that the block lengths of the codes will match, because the limit superior might not be achieved by the same sub-sequence of block lengths for both types. Specifically, for any given block length such that one of the codes has ‘good’ error probability (i.e., close to the probability guaranteed by the supremum reliability function), the other code might have ‘poor’ error probability, and vice versa. Since in order to construct a good SW code, we need to find a sequence of block lengths such that both channel codes have good error probability, this can only be guaranteed for the lower error exponent of the infimum reliability function. Indeed, for the infimum reliability function, good error probability is assured for all sufficiently large block lengths, and so, when the block length is sufficiently large, both channel codes, if properly designed, will have error probability close to the one guaranteed by the infimum reliability function.

**B. Bounds on Exponents for General SW Codes**

In this subsection, we derive upper and lower bounds on the error exponent and excess-rate exponent, which hold for any sequence of variable-rate SW codes. Unlike the case of...
Theorem 2, the exponent bounds, in this subsection should consider all possible types in \( \mathcal{P}(X) \).

**Theorem 3:** Let \( S \) be an arbitrary sequence of SW codes. Then,

\[
\mathcal{E}^+_e(S) \leq \inf_{Q_X \in \mathcal{P}(X)} \left\{ D(Q_X \| P_X) + \mathcal{E}^+(H(Q_X) - \rho(Q_X), Q_X, P_{Y|X}) \right\} - \mathcal{R}(Q_X; S, Q_X, P_{Y|X}). \tag{29}
\]

**Theorem 4:** Let \( S \) be any arbitrary sequence of SW codes. Then,

\[
\mathcal{E}_e(S, R) \leq \inf_{Q_X \in \mathcal{P}(X)} \left\{ D(Q_X \| P_X) + \mathcal{E}_e(S, R, Q_X) \right\}. \tag{30}
\]

Next, we derive an achievable exponent and excess-rate exponent for type-dependent, variable-rate SW codes. The proof is based on the achievability result of Theorem 2, but when considering the notion of excess-rate exponent, attention need to be given to all types of the source.

**Theorem 5:** For any given rate function \( \rho(Q_X) \), there exists a sequence of type-dependent, variable-rate SW codes \( S \) such that

\[
\mathcal{E}^{-}_e(S) \geq \inf_{Q_X \in \mathcal{P}(X)} \left\{ D(Q_X \| P_X) + \mathcal{E}^-(H(Q_X) - \rho(Q_X), Q_X, P_{Y|X}) \right\} \tag{31}
\]

and

\[
\mathcal{E}_e(S, R) \geq \inf_{Q_X \in \mathcal{P}(X): \rho(Q_X) \geq R} D(Q_X \| P_X). \tag{32}
\]

The proof is deferred to Appendix A, but here we provide an intuitive outline of the SW code constructed. From Theorem 2, it is possible to construct a SW code \( S^*_n(Q_X) \) for any given \( Q_X \in \text{int} \{ \mathcal{P}(X) \} \), with conditional error probability converging to about \( \exp[-n\mathcal{E}_e^+(H(Q_X) - \rho(Q_X), Q_X, P_{Y|X})] \). However, to obtain a SW code which satisfies (31) for a sufficiently long block length, the conditional error probability should be about \( \exp[-n\mathcal{E}_e^+(H(Q_X) - \rho(Q_X), Q_X, P_{Y|X})] \) uniformly over all types. Indeed, if uniform convergence is not satisfied then, for any given finite block length, there might be types \( Q_X \) such that the error probability of \( S^*_n(Q_X) \) is still far from its limit, and the error probability of this code may be a dominant factor in the total error probability. Thus, we have to prove uniform convergence of the error probability. Our strategy is as follows. We choose a large block length \( n \) such that the types of \( \mathcal{P}_{n0}(X) \) are good approximations for all types in \( \mathcal{P}(X) \), and construct good SW codes \( S^*_n(Q_X) \) for all \( Q_X \in \mathcal{P}_{n0}(X) \). Since \( |\mathcal{P}_{n0}(X)| \) is finite, uniform convergence of the error probability of \( S^*_n(Q_X) \) holds. For any given \( n \), upon observing a block from the source, we will modify it (namely, by truncating it and altering some of its components), so that the modified source block would have a type within \( \mathcal{P}_{n0}(X) \), and can then be encoded by one of the ‘good’ SW codes \( S^*_n(Q_X) \). The encoded modified vector will be sent to the decoder, along with the modification data. Then, at the decoder, the side information vector will be modified accordingly, so it appears as resulting from the memoryless source \( P_{Y|X} \), but conditioning on the modified source block. Thus, the decoder of \( S^*_n(Q_X) \) can be used to decode the modified vector, and the modification data can be used to recover the actual source block.

**Remark 6:** According to Theorem 5 and the proof of the achievability part of Theorem 2, it is implicit that the random binning exponent, defined as

\[
\min_{Q_{XY}} \left\{ D(Q_{XY} \| P_{XY}) + \left[ R - H(Q_{X|Y}|Q_Y) \right]_+ \right\}. \tag{33}
\]

may be achieved by using permutations of a channel code which achieve the random coding exponent. However, as is well known, for fixed-rate SW coding [21], one can achieve the error exponent by simple random binning, i.e. assigning source blocks to bins independently, with a uniform probability distributions over the bins. The random binning procedure can be generalized to type-dependent, variable-rate SW codes. In a nutshell, for any given rate function \( \rho(Q_X) \), we generate \( n\rho(Q_X) \) bins for each of the types \( Q_X \in \mathcal{P}_n(X) \). Then, we assign a bin to each \( x \in T_n(Q_X) \) by independent random selection, with a uniform distribution over all the bins of type class \( T_n(Q_X) \). The encoded codeword is a concatenation of a binary string representing the type of the source block (which adds negligible rate), and a binary string of length \( n \times \rho(Q_X) \) nats, representing the bin. The decoder, which is aware of both the type and the bin of the current source block, may be either maximum likelihood (ML) decoder, or even a universal minimum conditional entropy decoder (which does not depend on \( P_{XY} \)). The exact exponent of this ensemble for both decoders is

\[
\min_{Q_{XY}} \left\{ D(Q_{XY} \| P_{XY}) + \left[ \rho(Q_X) - H(Q_{X|Y}|Q_Y) \right]_+ \right\}. \tag{34}
\]

a result analogous to [20] for random channel coding.

### C. Trade-Off Between Exponents

As common in variable-rate SW coding, a trade-off exists between the error exponent and excess-rate exponent. In the remaining part of the paper, we explore this trade-off by requiring the achievability of a certain target error exponent \( \mathcal{E}_E \) with maximal excess-rate exponent. Theorem 5 shows that in order to achieve a target error exponent \( \mathcal{E}_e \), a type-dependent, variable-rate SW code may be employed. Then, Theorems 3 and 4 provide upper bounds which quantify the gap from optimal performance. Comparing Theorem 5 with Theorems 3 and 4, it is evident that there might be two origins for a gap between the bounds. The first one lies in the error exponent expression, and the second is in the excess-rate exponent. We now discuss these differences.

First, in general, it is yet to be known whether the inequality \( \mathcal{E}_s^+(R, Q_X, W) \leq \mathcal{E}_E(R, Q_X, W) \) may be strict. Thus, if for the minimizers in (32) and (31) a strict inequality occurs, then a gap exists between the upper and lower bounds for the SW code. Nonetheless, it is also well known that for \( R \geq R_{e}(Q_X, W) \), \( \mathcal{E}_s^+(R, Q_X, W) = \mathcal{E}_E(R, Q_X, W) \) is guaranteed, and so there are cases in which the upper and lower bounds coincide, especially at low target error exponents \( \mathcal{E}_E \). Second, on substituting a rate function \( \rho(Q_X) \)

---

4For more details, see [43, Appendix C].

5In (32) and (31), a minimum might not be achieved. In this case, the last statement should be valid for all sequence of distributions which achieves the infimum.
of a type-dependent, variable-rate codes in (30), the resulting upper bound is different from the lower bound of (32), only if the function

$$\inf_{Q_X \in \mathcal{P}(X)} D(Q_X || P_X)$$

(35)

is not left-continuous in $R$. As will turn out, for the class of rate functions of interest, left-continuity is satisfied, and the upper and lower bounds coincide. Thus, from the above discussion, we conclude that type-dependent, variable-rate SW codes are optimal for sufficiently low target error exponents $E_e$.

Since from Theorem 5, any target error exponent can be achieved with type-dependent, variable-rate SW codes, and because they are provably optimal in some domain, we henceforth consider only such SW codes. We will define optimal rate functions as follows.

**Definition 7:** A rate function $\rho^*(Q_X, E_e)$ is said to be inf-optimal, if for any $\delta > 0$, there exists a sequence of type-dependent, variable-rate SW codes $S$ with $R(Q_X; S) \leq \rho^*(Q_X, E_e) + \delta$ and $E_e^* (S) \geq E_e$, and for every other rate function $\rho(Q_X)$ with the above property, we have $\rho^*(Q_X, E_e) \leq \rho(Q_X)$, for all $Q_X \in \mathcal{P}(X)$. The sup-optimal rate function $\overline{\rho}^*(Q_X, E_e)$ is defined analogously.

Notice that by definition, we have

$$\overline{\rho}^*(Q_X, E_e) \leq \rho^*(Q_X, E_e).$$

(36)

In Section IV, we will obtain bounds on the optimal rate functions for any given $E_e$, and in Section V, we will obtain bounds on the excess-rate performance for these optimal rate functions.

**IV. OPTIMAL RATE FUNCTIONS**

In this section, we explore the optimal rate functions, for any given $E_e$. Before discussing specific bounds, we characterize them using the inverse of the fixed composition reliability function.

Theorem 5 implies that for $\rho^*(Q_X, E_e)$ to be inf-optimal, we must have

$$E_e \leq D(Q_X || P_X) + E^*_e (H(Q_X) - \rho^*(Q_X, E_e), Q_X, P_{Y|X})$$

(37)

for any given $Q_X \in \mathcal{P}(X)$.

We have the following proposition regarding general properties of $E^*_e(R, Q_X, W)$ and $E^*_e(R, Q_X, W)$ (for a proof, see [43, Appendix B, Proposition 27]). The following corollary is immediate.

**Proposition 8:** For a given channel $W$, the functions $E^*_e(R, Q_X, W)$ and $E^*_e(R, Q_X, W)$ are:

1. Strictly decreasing functions of $R$ in the interval $(C_0(Q_X, W), I(Q_X \times W))$.
2. Continuous functions of $(R, Q_X)$ for $R \in (C_0(Q_X, W), I(Q_X \times W))$.

**Corollary 9:** As a function of $R$, the function $E^*_e(R, Q_X, W)$ has a continuous inverse $E^*_e(R, Q_X, W)$ across the interval $E_e \in [0, E_0(Q_X, W))$. An analogous result holds for $E^*_e(R, Q_X, W)$.

Define $E_{e,0} \triangleq D(Q_X || P_X)$, where for brevity, here and in definitions throughout the section, the dependency in $Q_X$ of the defined quantities is omitted. Define the intervals

$$J_1 \triangleq [0, D(Q_X || P_X)],$$

$$J^-_2 \triangleq (D(Q_X || P_X), D(Q_X || P_X) + E_0^-(Q_X, W)),$$

where $E_0^-(Q_X, W)$ is as defined in (16), and

$$J^-_3 \triangleq [D(Q_X || P_X) + E_0^-(Q_X, W), \infty).$$

(39)

The intervals $J^+_2$ and $J^-_3$ are defined analogously. Now, Corollary 9 immediately implies the following:

$$\rho^*(Q_X, E_e) \begin{cases} 0, & E_e \in J_1 \\ H(Q_X) - R^*(E_e - D(Q_X || P_X), Q_X, P_{Y|X}), & E_e \in J^-_2 \\ H(Q_X) - C_0(Q_X, P_{Y|X}), & E_e \in J^-_3, \end{cases}$$

(41)

where $R^*(E_e - D(Q_X || P_X), Q_X, P_{Y|X})$ is as defined in Corollary 9.

Similarly, Theorem 3 implies that $\overline{\rho}^*(Q_X, E_e)$ cannot be sup-optimal unless

$$E_e \leq D(Q_X || P_X) + \overline{E}^*_e (H(Q_X) - \overline{\rho}^*(Q_X, E_e), Q_X, P_{Y|X})$$

(42)

for any given $Q_X \in \mathcal{P}(X)$. Now Corollary 9 implies:

$$\overline{\rho}^*(Q_X, E_e) \begin{cases} 0, & E_e \in J_1 \\ H(Q_X) - \overline{R}^*(E_e - D(Q_X || P_X), Q_X, P_{Y|X}), & E_e \in J^+_2 \\ H(Q_X) - C_0(Q_X, P_{Y|X}), & E_e \in J^+_3. \end{cases}$$

(43)

In Definition 7, $\rho^*(Q_X, E_e)$ is only defined for $Q_X \in \mathcal{P}(X)$. This is because the value of $\rho^*(Q_X, E_e)$ for $Q_X \in \mathcal{Q}(X) \setminus \{\mathcal{P}(X)\}$ (any irrational PMF) has no operational meaning, and does not affect exponents (see Theorems 2, 4, and 5). Thus, for $Q_X \in \mathcal{Q}(X) \setminus \{\mathcal{P}(X)\}$, we may arbitrarily define it as the lower semi-continuous extension of $\rho^*(Q_X, E_e)$. Specifically, for any given $Q_X \in \mathcal{Q}(X) \setminus \{\mathcal{P}(X)\}$ we henceforth define

$$\rho^*(Q_X, E_e) \triangleq \lim_{\epsilon \downarrow 0} \inf_{Q_X \in \mathcal{Q}(X) \setminus \{\mathcal{P}(X)\}, \|Q_X - Q_X^\epsilon\|_1 \leq \epsilon} \rho^*(Q_X, E_e),$$

(44)

and the same convention will be used for $\overline{\rho}^*(Q_X, E_e)$.

**Lemma 10:** The rate function $\rho^*(Q_X, E_e)$ is regular and strictly increasing in the range

$$E_e \in (D(Q_X || P_X), E_0^-(Q_X, W)).$$

(45)

The same properties hold for $\overline{\rho}^*(Q_X, E_e)$.

**Proof:** These properties follow directly from Corollary 9.

Next, we provide specific bounds on the optimal rate functions. Generally, any bound on the reliability function may be used, but we will focus on the random binning exponent and
expurgated exponent as lower bounds to the largest achievable exponent, and the sphere packing exponent as an upper bound. In essence, these bounds are generalizations of the random binning bound [13, Th. 2], [14, Th. 2], the expurgated bound, which follows from [14, Th. 2], and the sphere packing bound [13, Th. 3] for type-dependent, variable-rate SW coding. For the sake of simplicity, we assume that \( C_0(Q_X, P_Y|X) = 0 \) for all \( Q_X \), and so the expurgated and sphere packing exponents are finite for every positive rate. The results are easily generalized to the case of \( C_0(Q_X, P_Y|X) > 0 \).

We first need a few more definitions. Let

\[
Q_{Y|X} \triangleq \arg \min_{Q_{Y|X}} \left\{ I(Q_X \times Q_{Y|X}) + D(Q_{Y|X}||P_{Y||X}|Q_X) \right\},
\]

\[
Q_{X|Y} \triangleq \arg \min_{Q_{X|Y}} \left\{ B(Q_X \times Q_X, P_Y|X) \right\},
\]

where \( B(Q_X, P_Y|X) \) is defined in (19). Next, define\(^6\)

\[
E_{e,a-rb} \triangleq D(Q_X||P_X) + D(Q_{Y|X}||P_{Y||X}|Q_X),
\]

\[
E_{e,max-rb} \triangleq D(Q_X||P_X) + I(Q_X \times Q_{Y|X}) + D(Q_{Y|X}||P_{Y||X}|Q_X),
\]

\[
E_{e,a-ex} \triangleq D(Q_X||P_X) + B(Q_X \times Q_{X|Y}, P_X|X),
\]

\[
E_{e,max-ex} \triangleq D(Q_X||P_X) + B(Q_X \times Q_Y, P_Y|X),
\]

and

\[
E_{e,max-sp} \triangleq D(Q_X||P_X) + D(Q_X \times (Q_X \times P_Y|X) || P_{X}Y).
\]

Also, define the sets

\[
A_{rb} \triangleq \left\{ Q_{Y|X} : D(Q_X \times Q_{Y|X}||P_{XY}) \leq \varepsilon \right\},
\]

\[
A_{ex} \triangleq \left\{ Q_{X|Y} : Q_{X} = Q_X, E_e = D(Q_X||P_X) + B(Q_X \times Q_X, P_Y|X) \right\},
\]

and let \( A_{sp} \triangleq A_{rb} \). The random binning rate function is defined as

\[
\rho_{rb}(Q_X, E_e) \triangleq \begin{cases} 0, & E_e \in \mathcal{J}_1 \\ E_e + H(Q_X) - D(Q_X||P_X) - \min_{Q_{Y|X} \in A_{rb}} \left\{ I(Q_X \times Q_{Y|X}) + D(Q_{Y|X}||P_{Y||X}|Q_X) \right\}, & E_e \in \mathcal{J}_{rb,2} \\ E_e - E_{e,a-rb} + H(Q_X) - I(Q_X \times Q_{Y|X}), & E_e \in \mathcal{J}_{rb,3} \\ H(Q_X), & E_e \in \mathcal{J}_{rb,4}, \end{cases}
\]

where we have defined the intervals

\[
\mathcal{J}_{rb,2} \triangleq \{ E_{e,0}, E_{e,a-rb} \},
\]

\[
\mathcal{J}_{rb,3} \triangleq \{ E_{e,a-rb}, E_{e,max-rb} \},
\]

\[
\mathcal{J}_{rb,4} \triangleq \{ E_{e,max-rb}, \infty \}.
\]

The expurgated rate function is defined as

\[
\rho_{ex}(Q_X, E_e) \triangleq \begin{cases} 0, & E_e \in \mathcal{J}_1 \\ E_e - E_{e,a-ex} + H(Q_X) - I(Q_X \times Q_{X|Y}), & E_e \in \mathcal{J}_{ex,2} \\ H(Q_X) - \min_{Q_{Y|X} \in A_{ex}} I(Q_X \times Q_{Y|X}), & E_e \in \mathcal{J}_{ex,3} \\ H(Q_X), & E_e \in \mathcal{J}_{ex,4}, \end{cases}
\]

where we have defined the intervals

\[
\mathcal{J}_{ex,2} \triangleq \{ E_{e,0}, E_{e,a-ex} \},
\]

\[
\mathcal{J}_{ex,3} \triangleq \{ E_{e,a-ex}, E_{e,max-ex} \},
\]

and

\[
\mathcal{J}_{ex,4} \triangleq \{ E_{e,max-ex}, \infty \}.
\]

Finally, the sphere packing rate function is defined as

\[
\rho_{sp}(Q_X, E_e) \triangleq \begin{cases} 0, & E_e \in \mathcal{J}_1 \\ H(Q_X) - \min_{Q_{Y|X} \in A_{sp}} I(Q_X \times Q_{Y|X}), & E_e \in \mathcal{J}_{sp,2} \\ H(Q_X), & E_e \in \mathcal{J}_{sp,3}, \end{cases}
\]

where we have defined the intervals

\[
\mathcal{J}_{sp,2} \triangleq \{ E_{e,0}, E_{e,max-sp} \},
\]

and

\[
\mathcal{J}_{sp,3} \triangleq \{ E_{e,max-sp}, \infty \}.
\]

**Theorem 11**: For any given \( E_e \) and \( Q_X \in \mathcal{P}(X) \)

\[
\rho_{sp}(Q_X, E_e) \leq \rho^*(Q_X, E_e) \leq \rho_{rb}(Q_X, E_e) \leq \rho_{ex}(Q_X, E_e),
\]

Due to the similarity between the random binning bound and sphere packing bound, we obtain the known property from channel coding: For any \( Q_X \) there exists \( E_{e,cr}(Q_X) \) such that if \( E_e \leq E_{e,cr}(Q_X) \) we get \( \rho_{rb}(Q_X, E_e) = \rho_{sp}(Q_X, E_e) \). Thus, for any required \( E_e \), if \( E_e \leq E_{e,cr}(Q_X) \) then the optimal rate function is exactly known. Specifically, the right limit of the optimal rate function \( \rho_{rb}(Q_X, E_e) \) at its discontinuity point \( E_{e,0} \) can be easily evaluated from (55) to be

\[
\lim_{E_e \downarrow E_{e,0}} \rho_{rb}(Q_X, E_e) = H(Q_X) - D(P_Y||Q_Y) \]

where \( Q_Y = \sum_{x \in X} Q_X(x) P_{Y|X}(y|x) \). Namely, the resulting rate is the conditional entropy \( H(Q_X|Y) \) of the distribution \( Q_{XY} = Q_X \times P_Y|X \). Especially, for \( Q_X = P_X \) we have that \( \rho_{rb}(Q_X, \varepsilon) \geq H(P_{X|Y}||P_Y) \), for all \( \varepsilon > 0 \), as expected. The following lemma provides several simple properties of the rate functions \( \rho_{rb}(Q_X, E_e) \), \( \rho_{ex}(Q_X, E_e) \) and \( \rho_{sp}(Q_X, E_e) \).
Lemma 12: The rate functions $\rho_{\text{b}}(Q_X, E_r)$, $\rho_{\text{ex}}(Q_X, E_r)$ and $\rho_{\text{sp}}(Q_X, E_r)$ have the following properties:

- Strictly positive for $E_r > E_{r,0}$.
- Strictly increasing as a function of $E_r \geq E_{r,0}$ and $E_r \leq E_{r,\max\text{-rb}}$ (for random binning) or $E_r \leq E_{r,\max\text{-ex}}$ (for expurgated) and $E_r \geq E_{r,\max\text{-sp}}$ (for sphere packing).
- Concave in $E_r \in (E_{r,0}, \infty)$.
- Regular rate functions.

As for capacity and error exponents in channel coding, the computation of the bounds on the optimal rate function requires the solution of a non-trivial optimization problem. We defer the discussion on this matter to Section VI, where we discuss iterative algorithms for the computation of the bounds on the optimal rate functions, as well as their excess-rate performance. Nonetheless, in Appendix B, we provide analytic approximations for $\rho_{\text{b}}(Q_X, E_r)$ and $\rho_{\text{sp}}(Q_X, E_r)$ in the case of weakly correlated sources.\(^3\)

V. Excess-Rate Performance

In this section, we evaluate the excess-rate exponent of the optimal rate functions bounds, as defined in Section IV. This results in lower and upper bounds on the maximal achievable excess-rate exponent for a given error exponent, and thus the characterization of the optimal trade-off between error exponent and excess-rate exponent.

Notice that for a general rate function $\rho(\cdot)$, and target rates $R \in \mathbb{E}[r(X)], \max QX, \rho(QX))$, the excess-rate exponent is strictly positive and finite. The next lemma shows that the upper bound of (30) and the lower bound of (32) coincide for regular rate functions. Since in Lemma 12 it was shown that inf/sup optimal rate functions as well as the random binning, expurgated and sphere packing rate functions are all regular rate functions, this means that we have the exact expression for their excess-rate performance.

Lemma 13: For a regular rate function $\rho(Q_X)$

$$\inf_{Q_X: \rho(Q_X) = R} D(Q_X || P_X) = \min_{Q_X: \rho(Q_X) = R} D(Q_X || P_X). \quad \text{(68)}$$

We now mention a few general properties of excess-rate exponents functions.

Lemma 14: Let $\rho(Q_X)$ be a rate function, and $R_{\max} = \sup QX, \rho(QX))$. If $\rho(Q_X)$ is regular, let $R'_{\max} = \sup QX \in \mathbb{E} \rho(QX)$. The excess-rate exponent $E_r(R)$ for the rate function $\rho(Q_X)$ has the following properties:

- $E_r(R) = 0$ for $R \in [0, \rho(P_X)]$.
- $E_r(R) = \infty$ for $R \in (R_{\max}, \infty)$.
- $E_r(R)$ is increasing in $[\rho(P_X), R_{\max}]$. If $\rho(Q_X)$ is regular, then $E_r(R)$ is strictly increasing in $[\rho(P_X), R'_{\max}]$.
- $E_r(R)$ is continuous in $[\rho(P_X), R_{\max}]$ except for a countable number of points. If $\rho(Q_X)$ is regular, then $E_r(R)$ is left-continuous in $[\rho(P_X), R'_{\max}]$.

In the rest of the section, we assume that a target error exponent $E_r$ is given and fixed. Thus, for brevity, we omit the notation of the dependence of various quantities on it. We define the excess-rate exponent of the inf-optimal rate function as

$$\bar{E}_r^*(R) \triangleq \min_{Q_X: \rho(Q_X) \geq R} D(Q_X || P_X), \quad \text{(69)}$$

and analogously, define $\bar{E}_r^*(R)$. Similarly, we define the random-binning excess-rate exponent as

$$E_{r,\text{rb}}(R) \triangleq \min_{Q_X: \rho_{\text{rb}}(Q_X, E_r) \geq R} D(Q_X || P_X), \quad \text{(70)}$$

and analogously define $E_{r,\text{ex}}(R)$ and $E_{r,\text{sp}}(R)$. For a given $R$, we evidently have

$$\max\{E_{r,\text{rb}}(R), E_{r,\text{ex}}(R), E_{r,\text{sp}}(R)\} \leq \bar{E}_r^*(R) \leq E_r^*(R) \leq E_{r,\text{sp}}(R). \quad \text{(71)}$$

For some $R$, let the minimizer in (70) be $Q_{X, r}^*$. Then, if $\rho_{\text{b}}(Q_{X, r}^*, E_r) = \rho_{\text{sp}}(Q_{X, r}^*, E_r)$, it is easily verified that the bounds in (71) are tight, and $E_r^*(R) = \bar{E}_r^*(R) = E_{r,\text{rb}}(R)$. In other cases, one can use the upper bound at the rate function $\rho_{\text{sp}} = \min\{\rho_{\text{b}}(Q_X, E_r), \rho_{\text{ex}}(Q_X, E_r)\}$ to obtain an excess-rate exponent $E_{r,\text{sp}}(R)$, defined similarly to (70). Then, an improvement over the random-binning and expurgated excess-rate exponents is guaranteed, since

$$\max\{E_{r,\text{rb}}(R), E_{r,\text{ex}}(R)\} \leq E_{r,\text{sp}}(R). \quad \text{(72)}$$

Next, we evaluate the bounds on the optimal excess-rate exponent, e.g., as in (70). However, as we have seen, $\rho_{\text{b}}(Q_X, E_r)$, as well as the other rate functions, are not given analytically, and performing the maximization in (70) directly may be prohibitively complex, especially when $|X|$ is large. Thus, we describe an indirect method to evaluate the excess-rate bounds. For a given $R$, any curve $E_r = E_r(R)$ may be characterized by a condition that verifies whether the rate and excess-rate pair $(R, E_r)$ is either below or above the curve. The proof is based on the following lemma, that introduces a rate function designed to achieve pointwise $(R, E_r)$, but not necessarily $E_r$.

Lemma 15: Let

$$\hat{\rho}(Q_X; R, E_r) \triangleq \begin{cases} R, & D(Q_X || P_X) < E_r, \\ R_0, & \text{otherwise}. \end{cases} \quad \text{(73)}$$

Then, if there exists $R_0$ such that $\hat{\rho}(Q_X; R, E_r)$ achieves infimum error exponent $E_r$, then $\rho^*(Q_X, E_r)$ achieves infimum error exponent $E_r$, with rate $R$ and excess-rate exponent $E_r$. If $\hat{\rho}(Q_X; R, E_r)$ does not achieve supremum error exponent $E_r$ then $\bar{E}_r^*(Q_X, E_r)$ does not achieve supremum error exponent $E_r$ with rate $R$ and excess-rate exponent $E_r$.\(^2\)

Proof:

(\Rightarrow) Assume that $\hat{\rho}(Q_X; R, E_r)$ achieves $(R, E_r)$ with an infimum error exponent $E_r$. Clearly the definition of an optimal rate function imply that $\rho^*(Q_X, E_r)$ also achieves $(R, E_r)$.

(\Rightarrow) Assume that $\bar{E}_r^*(Q_X, E_r)$ achieves $(R, E_r)$. If $Q_X$ satisfies $D(Q_X || P_X) \geq E_r$ then for

$$R_0 \geq \max_{Q_X: D(Q_X || P_X) \geq E_r} \bar{E}_r^*(Q_X, E_r), \quad \text{(74)}$$

we get $\hat{\rho}(Q_X; R, E_r) \geq \bar{E}_r^*(Q_X, E_r)$. Else, if $\hat{\rho}(Q_X; R, E_r) > R$ for some $Q_X$ that satisfies $D(Q_X || P_X) < E_r$, then $\bar{E}_r^*(Q_X, E_r)$ does not achieve $(R, E_r)$.

\(^2\)The expurgated bound is not very useful in this regime [42, Sec. 3.4].
using Lemma 13. Thus, we must have \( \hat{\rho}(Q_X; R, E_e) \geq \overline{\rho}(Q_X, E_e) \) for all \( Q_X \) and this implies that \( \hat{\rho}(Q_X; R, E_e) \) also achieves supremum error exponent \( E_e \). It is easy to see that \( \hat{\rho}(Q_X; R, E_e) \) has excess-rate exponent \( E_e \) at rate \( R \) directly from its construction and Lemma 13.

Notice that the rate function \( \hat{\rho}(Q_X; R, E_e) \), introduced in the previous lemma, has only pointwise optimal excess-rate exponent, in the sense that for the given \((R, E_e)\) it achieves the optimal trade-off between the error exponent and excess-rate exponent. By contrast, the optimal rate functions \( \rho^*(Q_X, E_e) \) and \( \overline{\rho}(Q_X, E_e) \) achieve the optimal excess-rate exponent, at any given rate.

Define for a given \((R, E_e)\)
\[
\Gamma_{ib}(t, Q_X, Q_{Y|X}) \triangleq D(Q_X||P_X) + D(Q_{Y|X}||P_{Y|X}|Q_X)
+ t \cdot \left[ R - H(Q_{X|Y}|Q_Y) \right]
\]
\[e_{ib}(t) \triangleq \min_{Q_X: D(Q_X||P_X) \leq E_e} \min_{Q_{Y|X}} \Gamma_{ib}(t, Q_X, Q_{Y|X}), \quad \Gamma_{ib}(t, Q_X, Q_{Y|X}) \triangleq \Gamma_{ib}(Q_X, Q_{Y|X})
\]
(75)
and
\[
\Gamma_{ex}(t, Q_X, Q_{X|X}) \triangleq D(Q_X||P_X) + B(Q_{X|X}, P_{X|X})
+ t \cdot \left[ R - H(Q_{X|X}) \right]
\]
\[e_{ex}(t) \triangleq \min_{Q_X: D(Q_X||P_X) \leq E_e} \min_{Q_{X|X}} \Gamma_{ex}(t, Q_X, Q_{X|X}). \quad \Gamma_{ex}(t, Q_X, Q_{X|X}) \triangleq \Gamma_{ex}(Q_X, Q_{X|X})
\]
(76)
Also, define
\[
\Gamma_{sp}(t, Q_X, Q_{Y|X}) \triangleq \Gamma_{ib}(t, Q_X, Q_{Y|X})
\]
\[e_{sp}(t) \triangleq e_{ib}(t).
\]
(79)
(80)

**Theorem 16:** If
\[
\max_{0 \leq t \leq 1} \left\{ \max_{0 \leq t \leq I} e_{ib}(t), \max_{I \geq 1} e_{ex}(t) \right\} \geq E_e
\]
(81)
then there exists a sequence of SW codes with infimum error exponent \( E_e \), and excess-rate exponent \( E_e \) at rate \( R \).

Conversely, if
\[
\max_{I \geq 0} e_{sp}(t) < E_e
\]
(82)
then there is no sequence of SW codes with supremum error exponent \( E_e \), and excess-rate exponent \( E_e \) at rate \( R \).

The functions \( e_{ib}(t), e_{ex}(t) \) and \( e_{sp}(t) \) are concave functions of \( t \) (as pointwise minimum of linear functions in \( t \)), and thus the maximization over \( t \) is relatively simple to perform. In addition, \( \max_{0 \leq t \leq 1} e_{ib}(t), \max_{I \geq 1} e_{ex}(t) \) and \( \max_{I \geq 0} e_{sp}(t) \) are non-increasing functions of \( E_e \) and so for any given constraint on \( E_e \) and target rate \( R \), a simple line search algorithm will find \( E_e(R) = \min \{ E_e : (R, E_e) \text{ is achievable for } E_e \} \). Thus, the computational problem is to compute \( e_{ib}(t), e_{ex}(t) \) and \( e_{sp}(t) \), for any given \( t \). We address this matter in Section VI.

We conclude this section with a short discussion. As evident from its definition \( 7 \), for a given constraint \( E_e \) on the error exponent, the rate function \( \overline{\rho}(Q_X, E_e) \) must be employed in order to satisfy the constraint \( E_e \). Any other feasible rate function \( \rho(Q_X) \) will satisfy \( \rho(Q_X) \geq \overline{\rho}(Q_X, E_e) \) for some \( Q_X \in \mathcal{P}_n(A) \), and will consequently have inferior excess-rate exponent, no matter what is the target rate. In other words, the excess-rate exponent is maximized by \( \overline{\rho}(Q_X, E_e) \) simultaneously for every \( R \). Specifically, the average rate and the maximal rate (which is required to avoid buffer overflow) are determined. Indeed, for a given \( E_e \), the average rate is determined to be
\[
\mathbb{E}[r(X)] \rightarrow \overline{\rho}^*(Q_X, E_e) \triangleq R_0
\]
(83)
as \( n \to \infty \), and the maximal rate is determined to be
\[
\max r(x) \rightarrow \overline{\rho}^*(Q_X, E_e) = \max_{Q_X} \overline{\rho}^*(Q_X, E_e) \triangleq R_1
\]
(84)
as \( n \to \infty \). If, for example, a buffer of size \([nR_1 \cdot \log_2 e]\) bits is used, then the buffer never overflows. Further reducing the buffer size to \([nR_1 \cdot \log_2 e]\) bits with \( R < R_1 \), will result in a buffer overflow probability of \( \approx e^{-nE_e(R)} \). In a completely equivalent way, we can invert this relation. Consider the practical case in which we are given a buffer of length \([nR_1 \cdot \log_2 e]\), and would like to achieve an average rate \( R_0 \) without buffer overflow events, we can use the rate function
\[
\rho(Q_X) = \begin{cases} R_0, & ||P_X - Q_X|| \leq \epsilon \\ \min \{ H(Q_X, R_1), 1 \}, & ||P_X - Q_X|| > \epsilon \end{cases}
\]
(85)
and by taking \( \epsilon \downarrow \), the maximal error exponent that can be achieved this way is given by
\[
\max \left\{ E_e : \overline{\rho}^*(P_X, E_e) \leq R_0, \right. \\
\left. \max \overline{\rho}^*(Q_X, E_e) \leq \min \{ H(Q_X, R_1), 1 \} \right\}
\]
(86)
In addition, we can find the required rate for fixed-rate coding and coding under average rate constraint. In the case of fixed-rate coding, to ensure an infimum error exponent of \( E_e \) one must use \( \rho(Q_X) = R_1 = \max_{Q_X} \overline{\rho}^*(Q_X, E_e) \) for all \( Q_X \in \mathcal{P}_n(A) \). For coding under average rate constraint, to ensure a supremum error exponent of \( E_e \), one can choose, as discussed in Subsection III-A, for any given \( \epsilon > 0 \), the rate function
\[
\rho(Q_X) = \begin{cases} \overline{\rho}^*(P_X, E_e), & ||P_X - Q_X|| \leq \epsilon \\ H(Q_X), & ||P_X - Q_X|| > \epsilon \end{cases}
\]
(87)
and by taking \( \epsilon \downarrow \), the average rate is asymptotically given by \( \overline{\rho}^*(P_X, E_e) \). It is also evident that if \( \max_{Q_X, E_e} \overline{\rho}(Q_X, E_e) = \overline{\rho}(P_X, E_e) \) then fixed-rate coding is optimal and the excess-rate exponent cannot be improved beyond that of fixed-rate.

Finally, consider the scenario in which an excess-rate event for rate \( R \) and a decoder error event are both considered a failure of the system. Then, the total probability of a system

\[8\] An analogous discussion holds for \( \overline{\rho}(Q_X, E_e) \).
failure satisfies
\[ \mathbb{P}\{\text{system failure}\} \triangleq \mathbb{P}(r(X) \geq R \cup \hat{X} \neq X) \leq e^{-nE_e} + e^{-nE^*_e(R)} \]  
(88)
and the achievable exponent of the system failure probability, is given by
\[ \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}\{\text{system failure}\} = \max_{E_e} \{E_e, E^*_e(R)\}. \]  
(89)
Since \( E^*_e(R) \) is a non-increasing function of \( E_e \), then typically the maximum will be the solution for the fixed point equation
\[ E_e = E^*_e(R). \]  
(90)
Therefore, the resulting system failure exponent can only be better compared to the case of fixed rate (\( E^*_e(R) = \infty \)).

VI. COMPUTATIONAL ALGORITHMS

As we have seen in Sections IV and V, in order to compute the bounds on the optimal rate functions and the resulting excess-rate performance, optimization problems need to be solved. In essence, since the bounds on the optimal rate functions stem from the bounds on channel coding error exponents, any computational algorithm for channel coding error exponents may be used [3], [29]. However, these classical algorithms are given for Gallager-style bounds [19], not the form of Csiszár and Körner [15], used in this paper. Moreover, they form the basis for the computational algorithm of the excess-rate performance for the random binning and sphere packing bounds.

For the random binning and sphere packing rate functions (55), (63), it is required to compute \( \nu_{eh}(P_{XY}, Q_X, E_e, \eta) \)
\[ \begin{align*}
\nu_{eh}(P_{XY}, Q_X, E_e, \eta) & \triangleq \min_{Q_{Y|X} \in \mathcal{Q}_{Y|X}} D(Q_{Y|X} \| P_{Y|X}) \leq E_e \\
& \left\{ D(Q_{Y|X} \| Q_{Y|X}^r P_{Y|X}) + \eta \cdot D(Q_{Y|X} \| P_{Y|X}^r Q_X) \right\},
\end{align*} \]  
(91)
where \( \eta = 1 \) for (55) and \( 0 \) for (63). For expurgated rate function (59) it is required to compute \( \nu_{ex}(P_{XY}, Q_X, E_e) \)
\[ \begin{align*}
\nu_{ex}(P_{XY}, Q_X, E_e) & \triangleq \min_{Q_{X|X^\alpha} \in \mathcal{Q}_{X|X^\alpha}^\alpha} D(Q_{X|X^\alpha} \| Q^\alpha_{X|X^\alpha} Q_X). 
\end{align*} \]  
(92)
Moreover, to compute the bounds on the excess-rate performance in (76) and (78), the values of \( \nu_{eh}(P_{XY}, R, E_e, t) \) and \( \nu_{ex}(P_{XY}, R, E_e, t) \) need to be computed. In this section, we provide explicit iterative algorithms to compute \( \nu_{eh}(P_{XY}, Q_X, E_e, \eta) \), \( \nu_{ex}(P_{XY}, Q_X, E_e) \) and \( \nu_{eh}(P_{XY}, R, E_e, t) \), and prove their correctness. The merit of these algorithms is that they require at most a one-dimensional optimization, regardless of the alphabet sizes \( |\mathcal{X}| \) and \( |\mathcal{Y}| \). The optimization problem of \( e_{ex}(P_{XY}, R, E_e, t) \) is briefly discussed, and shown to be convex, rendering it feasible to compute using generic algorithms.

Throughout, we will utilize an auxiliary PMF \( \tilde{Q}_Y \). For \( 0 \leq \alpha \leq 1 \), define the geometric combination mapping \( \tilde{M}_\alpha(P_{XY}, \tilde{Q}_Y, \alpha) \) whose output \( \tilde{Q}_{Y|X} \) satisfies
\[ \tilde{Q}_{Y|X}(y|x) \triangleq \psi_x p_{Y|X}^{\alpha}(y|x) \tilde{Q}_Y^{1-\alpha}(y), \]  
(93)
for all \( x \in \mathcal{X}, y \in \mathcal{Y} \), where \( \psi_x \) is a normalization factor, chosen such that \( \sum_{y \in \mathcal{Y}} \tilde{Q}_{Y|X}(y|x) = 1 \) for all \( x \in \mathcal{X} \). Algorithm 1 provides a method to compute \( \nu_{eh}(P_{XY}, Q_X, E_e, \eta) \).

Lemma 17: Algorithm 1 outputs \( \nu_{eh}(P_{XY}, Q_X, E_e, \eta) \).

Algorithm 1 is presented for a specific \( E_e \), but it is also useful if one is interested in the full curve \( \rho_{eh}(Q_X, E_e) \). To compute the second term in the random binning rate function (55) one needs to compute
\[ \begin{align*}
\min_{Q_{Y|X} \in \mathcal{Q}_{Y|X}} I(Q_X \| Q_{Y|X}) + D(Q_{Y|X} \| P_{Y|X} Q_X) 
\left( \begin{array}{l}
= \min_{Q_{Y|X}} \left\{ D(Q_{Y|X} \| Q_{Y|X}^r P_{Y|X}) + \lambda (D(Q_{Y|X} \| P_{Y|X}) + D(Q_{Y|X} \| P_{Y|X} Q_X) - E_e) \right\} \right.
\left. + \max_{\tilde{Q}_{Y|X}} \left\{ \lambda (D(Q_{Y|X} \| P_{Y|X}) + D(Q_{Y|X} \| P_{Y|X} Q_X) - E_e) \\
+ (1 + \lambda) D(Q_{Y|X} \| P_{Y|X} Q_X) \right\} \right.
\\end{array} \right) \]  
(95)
where \( \lambda \) is because the minimization problem is convex. The KKT optimality conditions [4, Sec. 5.5.3] imply that for any given \( \lambda \in [0, \infty) \) the inner minimizer \( Q_{Y|X}^*(\lambda) \) of last line in (95) is also the optimal solution for (95), whenever the
Algorithm 2: Alternating Minimization Algorithm for the Computation of $e_{eb}(P_{XY}, R, E_r, t)$

**Input:** A source $P_{XY}$, a target rate $R$, a target excess-rate $E_r$ and $t \geq 0$.  

**Output:** The value of $e_{eb}(P_{XY}, R, E_r, t)$.

1. Initialize $\hat{Q}_Y$ randomly such that $\text{supp}(\hat{Q}_Y) = \mathcal{Y}$, and set $\hat{Q}_{Y|X} = M_{\hat{Q}}(P_{X|Y}, \hat{Q}_Y, \frac{t}{1+t})$ and compute $h_1$ and $h_2$.
2. Iterate over the following steps until convergence:
   a) Set $\hat{Q}_X = M_{\hat{Q}}(P_X, h_1, h_2, 0, t)$. If $D(\hat{Q}_X || P_X) < E_r$ then set $\lambda = 0$. Else, find $\lambda^* > 0$ that satisfies
      \[ D(M_{\hat{Q}}(P_X, h_1, h_2, \lambda, t) || P_X) = E_r \]
      and set $\hat{Q}_X = M_{\hat{Q}}(P_X, h_1, h_2, \lambda^*, t)$.  
   b) Set $\hat{Q}_Y(y) = \sum_{x \in \mathcal{X}} \hat{Q}_X(x) \hat{Q}_{Y|X}(y|x)$ for all $y \in \mathcal{Y}$, set $\hat{Q}_{Y|X} = M_{\hat{Q}}(P_Y, \hat{Q}_Y, \frac{1}{1+t})$ and compute $h_1$ and $h_2$.  
3. Let the converged variables be $\lambda^*$ and $\hat{Q}_Y$. Set $\hat{Q}_X$, $\hat{Q}_{Y|X}$ in (76). Return.

error exponent constraint in $A_{eb}$ is given by

$$E_r(\lambda) = D(Q_X || P_X) + D(Q_{Y|X}(\lambda) || P_{Y|X} Q_X). \quad (96)$$

Clearly, Algorithm 1 is suitable for the inner minimization in (95), simply by setting $\eta = \lambda + 1$ and $E_r = \infty$. Equivalently, this means that in step 2a of the algorithm, we always set $\alpha^* = \frac{\eta}{\eta - 1} = \frac{1}{\lambda + 1}$. Otherwise stated, when $\alpha^*$ varies from 1 to $\frac{1}{\lambda}$, the curved part of $\rho_{eb}(Q_X, E_r)$ is exhausted.

Next, Algorithm 2 provides a method to compute $e_{eb}(P_{XY}, R, E_r, t)$. The technique is somewhat similar to Algorithm 1, but here an additional optimization is carried out over $Q_X$. For this algorithm, we define

$$h_{1,t}(x) = \frac{D(\hat{Q}_{Y|X}(\cdot|x)) || P_{Y|X} Q_X)}{t+1}, \quad (97)$$

$$h_{2,t}(x) = \frac{D(\hat{Q}_{Y|X}(\cdot|x)) || \hat{Q}_Y)}{t+1}, \quad (98)$$

where $\hat{Q}_{Y|X} = M_{\hat{Q}}(P_{Y|X}, \hat{Q}_Y, \frac{1}{1+t})$, as well as the mapping $M_{\hat{Q}}(P_X, h_1, h_2, \lambda, t)$ whose output $\hat{Q}_X$ satisfies

$$\hat{Q}_X(x) = \psi \cdot [P_X(x)]^\frac{1}{1+t+\lambda} \cdot h_{1,t}(x) - \frac{t}{1+t+\lambda} \cdot h_{2,t}(x), \quad (99)$$

for all $x \in \mathcal{X}$, where $\psi$ is a normalization factor, such that $\sum_{x \in \mathcal{X}} \hat{Q}_X(x) = 1$.

**Lemma 18:** Algorithm 2 outputs $e_{eb}(P_{XY}, R, E_r, t)$.

Next, Algorithm 3 provides a method to compute $v_{ex}(P_{XY}, Q_X, E_r)$. We define the Bhattacharyya mapping $M_b(Q_{\tilde{X}|X}, P_{X|\tilde{X}}, \lambda)$ whose output $Q_{\tilde{X}|X}$ satisfies

$$Q_{\tilde{X}|X}(\tilde{x}|x) = \psi_x Q_{\tilde{X}|X}(\tilde{x}|x) \exp \left[ -\lambda \cdot d_{P_{X|\tilde{X}}}(x, \tilde{x}) \right] \quad (100)$$

for all $x, \tilde{x} \in \mathcal{X}$, and $\psi_x$ is a normalization constant, such that $\sum_{\tilde{x} \in \mathcal{X}} Q_{\tilde{X}|X}(\tilde{x}|x) = 1$, for any $x \in \text{supp}(Q_X)$. Similarly, define the lumping mapping $M_b(Q_{\tilde{X}|X})$ whose output $Q_{\tilde{X}|X}$ satisfies

$$Q_{\tilde{X}|X}(\tilde{x}|x) = \frac{Q_{\tilde{X}|X}(\tilde{x}|x)}{Q_{\tilde{X}|X}(x)} \exp \left[ -\lambda \cdot d_{P_{X|\tilde{X}}}(x, \tilde{x}) \right] \quad (101)$$

for all $x, \tilde{x} \in \mathcal{X}$ and $\psi_x$ is a normalization constant, such that $\sum_{\tilde{x} \in \mathcal{X}} Q_{\tilde{X}|X}(\tilde{x}|x) = 1$, for any $x \in \text{supp}(Q_X)$. Similarly, define the lumping mapping $M_b(Q_{\tilde{X}|X})$ whose output $Q_{\tilde{X}|X}$ satisfies

$$Q_{\tilde{X}|X}(\tilde{x}|x) = \frac{Q_{\tilde{X}|X}(\tilde{x}|x)}{Q_{\tilde{X}|X}(x)} \exp \left[ -\lambda \cdot d_{P_{X|\tilde{X}}}(x, \tilde{x}) \right] \quad (102)$$

It can be easily seen that the resulting optimization problem is convex in the variables $(Q_{\tilde{X}}, Q_{\tilde{X}|X})$, and can be solved by any general solver. Unfortunately, we have not been able to prove that alternating minimization algorithm converges (and even in this event, there is no explicit solution for the optimal $Q_{\tilde{X}|X}$ given some $Q_{\tilde{X}}$).
VII. A NUMERICAL EXAMPLE

In this section, we provide a simple numerical example to illustrate the bounds obtained in previous sections, utilizing the computational algorithms of Section VI. Let the alphabets be $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y} = \{0, \chi, 1\}$. $P_X$ be given by $P_X(0) = 1 - P_X(1) = 0.2$, and $P_{Y|X}$ be given by the following transition probability matrix

$$P_{Y|X} = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.05 & 0.15 & 0.8 \end{bmatrix}.$$  

(103)

Figure 1 shows the bounds on the optimal rate functions (in nats) for $Q_X$ given by $Q_X(0) = 1 - Q_X(1) = 0.25$ as a function of $E_r$. The points at which $\rho_{\mathrm{rb}}(Q_X, E_r)$ becomes (respectively, ceases to be) affine with a unity slope are indicated by a vertical lines. For small $E_r$, the random binning and sphere packing bounds coincide, and so, $\overline{\rho}^s(Q_X, E_r) = \rho_{\mathrm{rb}}(Q_X, E_r) = \rho_\star(Q_X, E_r) = \rho^\star(Q_X, E_r)$.

Figure 2 shows the bounds on the optimal rate functions (in nats), for all possible types (indexed by $Q_X(0)$) for $E_r = 0.05$ and $E_r = 0.2$. It can be seen that indeed this optimal function is in the form of a regular rate function, and that for $E_r = 0.05$ the optimal rate function is exactly known, for all types of the source. For comparison, the entropies $H(Q_X)$ and $H(Q_X|Q_Y)$ where $Q_{XY} = Q_X \times P_{Y|X}$ are also plotted, and the rates for $P_X$ are marked. The bounds on the optimal excess-rate exponent are computed and plotted in Figure 3 for $E_r = 0.05$ and in Figure 4 for $E_r = 0.2$. As before, for the smaller $E_r$ the optimal excess-rate exponent is obtained exactly, while a gap exists for the larger $E_r$. It can be verified that Figures 2 and 3 are consistent.

For example, for $E_r = 0.05$ it can be seen in Figure 2 that for the type $Q_X = P_X$, the rate is $\rho^s(P_X, 0.05) = \overline{\rho}^s(P_X, 0.05) \approx 0.377$ nats so the excess-rate exponent is $\overline{E}^s_r(0.377) = \overline{E}^\star_r(0.377) = 0$. Following the discussion at the end of Section V, this is the minimal average rate required to achieve an error exponent of $E_r = 0.05$. Then, as $Q_X(0)$ increases, the rate also increases, up to its maximal value of $\rho^s(Q_X, 0.05) = \overline{\rho}^s(Q_X, 0.05) \approx 0.4$, for $Q_X(0) \approx 0.2574$. This maximal rate is the minimal rate required to totally eliminate excess-rate events. The excess-rate exponent of this rate is determined by the divergence of this type from the true source $P_X$, and given by $\overline{E}^s_r(0.4) = \overline{E}^\star_r(0.4) \approx D(Q_X^\star \| P_X) \approx 10^{-2}$. This is the maximal value of $\overline{E}^s_r(R)$ shown in Figure 3, and for larger rates, clearly $\overline{E}^s_r(R) = \infty$.

For comparison, we also consider fixed-rate coding. From Figure 3, for $E_r = 0.05$ we have $\overline{E}^s_r(0.3921) = 2 \cdot 10^{-3}$. It can be found that if one uses fixed-rate coding, at rate $R_0 = 0.3921$, for all $Q_X$ then the error exponent achieved is...
only $E_e \approx 0.045$. Therefore, if the finite excess-rate exponent of variable-rate coding is tolerated, then this provides an improvement in the error exponent over fixed-rate coding.

VIII. SUMMARY

In this paper, we have considered the trade-off between error and excess-rate exponents for variable-rate SW coding. The cases of fixed-rate coding and variable-rate coding under average constraints may be considered as two extreme points in this trade-off. In fixed-rate coding the same rate is assigned to all possible types, and so, the maximal excess-rate exponent is achieved, but at the price of minimal error exponent. In average-rate coding, the main concern is the coding of the true type of the source, and all other types are sent uncoded. The resulting error exponent is maximal, but at the price of minimal excess-rate exponent. Thus, for a coding system with more stringent instantaneous rate demands, it is necessary to lose some of the gains in error exponent of variable-rate coding, and improve the excess-rate exponent. In this work, we have derived bounds on rate functions which achieve the optimal trade-off, and analyzed their excess-rate performance, for a given requirement on the error exponent.

Before we conclude, we briefly outline two possible extensions. In many practical cases, there is some uncertainty regarding the source $P_{XY} = P_X \times P_{Y|X}$. Clearly, if independence between $X$ and $Y$ is a possible scenario, then in this worst case, the side information $Y$ is useless (when no feedback link exists). In other cases, it may be known that $P_{XY} \in \mathcal{F} \subset Q(\mathcal{X} \times \mathcal{Y})$ for some family of distributions $\mathcal{F}$. In this case, a possible requirement is that the rate function $\rho(Q_X)$ will be chosen to achieve error exponent $E_e$ uniformly for all sources in $\mathcal{F}$. With a slight change and abuse of notation, we define, e.g., the infimum optimal rate function for the source $P_{XY}$ as $\rho^*(Q_X, E_e; P_{XY})$ and the optimal rate function for the family $\mathcal{F}$ as

$$\rho^*(Q_X, E_e; \mathcal{F}) \triangleq \max_{P_{XY} \in \mathcal{F}} \rho^*(Q_X, E_e; P_{XY}). \quad (104)$$

This maximization is (relatively) easy to perform if, e.g., the conditional probability $P_{Y|X}$ is known exactly, and in addition, a nominal $P_X$ is known such that the actual $P_X$ satisfies $D(P_X||P_X) \leq U$, for some given uncertainty level $U > 0$ (recall Pinsker’s inequality [10, Lemma 11.6.1] and see also the discussion in [37]). A direction for future research is to derive bounds on optimal rate functions and their excess-rate performance which are robust for source uncertainty of various kinds.

In this paper, we have focused on the SW scenario in which the side information vector $Y$ is known exactly to the source. Similar techniques can also be applied to the more general case of SW coding, where the side information is also encoded. In this case, there are two encoders, $s_n$ for encoding $X$ and $s'_n$ for encoding $Y$, while the central decoder $\sigma_n$ now uses both codewords $s_n(x)$ and $s'_n(y)$. For type-dependent, variable-rate codes, two rate functions $p_X(Q_X)$ and $p_Y(Q_Y)$ may be defined accordingly. While bounds on the resulting error exponent may be derived, the trade-off in this case is more complicated.

First, there are two excess-rate exponents, one for each of the encoders. Second, a trickle of coordination might be required between the two encoders in order to ensure a required error exponent. Specifically, at least one of the encoders needs to know the current rate (or equivalently, the type class of the current source block) of the other encoder.

APPENDIX A

Proof of Theorem 2:

Upper Bound (26): Follows exactly as in the proof of [6, Th. 1].

Lower Bound (27): We shall use the following Lemma, whose proof can be found in [43, Appendix B, Lemma 26].

Lemma 20: Let $Q_X \in \mathcal{P}(\mathcal{X}) \cap \overline{\mathcal{Q}}(\mathcal{X})$, and let $n_0 \in \mathbb{N}$ be the minimal block length such that $Q_X \in \mathcal{P}_{n_0}(\mathcal{X})$. Let $\mathcal{C}$ be a sequence of fixed composition codes $C_n \subseteq \mathcal{C}_{n_0}$ such that $Q_{X}^{(n)} \to Q_X$, and $\liminf_{n \to \infty} \frac{1}{n} \log |C_n| \geq R$. Then, for any $\epsilon > 0$, there exists a sequence of fixed composition channel codes $C_{n_m} \subseteq \mathcal{T}_{n_m}(Q_X)$ such that $\liminf_{n \to \infty} \frac{1}{n} \log |C_n| \geq R - \epsilon$, and $\liminf_{n \to \infty} \frac{1}{n} \log p_e(C_{n_m}) \geq E_e^*(C) - \epsilon$.

For brevity, we will omit the notation of the dependence of $\mathcal{R}(Q_X; S)$ in $S$ and denote it by $\mathcal{R}(Q_X)$. Assume that $Q_X \in \text{int} \mathcal{Q}(\mathcal{X})$, and $Q_X \in \mathcal{P}_{n_0}(\mathcal{X})$ for some minimal $n_0 \in \mathbb{N}$. Since the statement in (27) is only about the conditional error exponent of the type $Q_X$, it is clear that the SW codes constructed, may only encode $x \in \mathcal{T}_{n}(Q_{X})$, so any block lengths $n \mod n_0 \neq 0$ should be considered, as otherwise $\mathcal{T}_{n}(Q_X)$ is empty, and the conditional error probability is 0, by definition.

Let $\delta > 0$ be given, and let $\mathcal{C}$ be a sequence of constant composition channel codes of type $Q_X^{(n)} \to Q_X$, asymptotic rate $\liminf_{n \to \infty} \frac{1}{n} \log |C_n| \geq H(Q_X) - \mathcal{R}(Q_X) - \frac{\delta}{2}$, which also achieves the infimum reliability function for the channel $P_{Y|X}$, i.e.,

$$\liminf_{n \to \infty} - \frac{1}{n} \log p_e(C_{n}) \geq E_e^*(H(Q_X) - \mathcal{R}(Q_X), Q_X, P_{Y|X}) - \delta. \quad (A.1)$$

As is well known, simple expurgation arguments can assure that (A.1) holds for the maximal error probability (over all codewords) and not just the average error probability, and so this will be assumed henceforth. From this point onward, the proof follows the steps of the proof of [6, Th. 1], and only for the sake of completeness, we provide it here. From Lemma 20, it can be assumed without loss of generality that for $n$ sufficiently large, whenever, $n \mod n_0 = 0$, the codebook satisfies $C_n \in \mathcal{T}_{n}(Q_{X})$. Now, assume that $n$ is sufficiently large and that $n \mod n_0 \neq 0$. From the covering lemma [1, Sec. 6, Covering Lemma 2], one can find

$$T_n = \exp \left[ n \left( \mathcal{R}(Q_X) + \delta \right) \right] \quad (A.2)$$

permutations $\{\pi_{n, t}\}_{t=1}^{T_n}$, such that $T_n(Q_X) = \bigcup_{t=1}^{T_n} \pi_{n, t}(C_{n})$, where $\pi_{n, t}(C_n)$ means that the same permutation $\pi_{n, t}()$
operates on codewords in the codebook. Since the channel $P_{Y|X}$ is memoryless then clearly $p_x(\pi(C_n)) = p_x(C_n)$ for any permutation $\pi$, since the decoder can always apply the inverse permutation on $Y$ and decode as if the codebook is $C_n$. Let us define the following sequence of SW codes $S^n = \{S^n_1, S^n_2, \ldots\}$ from the channel codes $C = \{C_n = (f_n, \phi_n)\}$.

- **Codebook Construction**: Generate the codebook $C_n$ and enumerate the permutations $\{\pi_{n,t}\}_{t=1}^{T_n}$ such that $T_n(Q_X) = \bigcup_{t=1}^{T_n} \pi_{n,t}(C_n)$. The above information is revealed to both the encoder and the decoder off-line.

- **Encoding**: Upon observing $x$, determine its empirical distribution $Q_X$. If $\hat{Q}_X \neq Q_X$ the codeword is $s^n_0(x) = 0$. Else, find $t^*(x) = \min_{t'}: x \in \pi_{n,t'}(C_n(Q_X))$. The codeword is $s^n_0(x) = (1, t^*(x))$ where $t(\pi)$ is the binary representation of $t$ in $\lceil \log_2 T_n \rceil$ bits.

- **Decoding**: If $s^n_0(x) = 0$ then declare an error. Else, recover from $s^n_0(x)$ the permutation $\pi = \pi_{n,t'}(x)$. Find $\hat{y}(\pi) \neq t^*(\pi(\phi_n(\pi^{-1}(\pi(\hat{y}))))$, and if $\hat{y}(\pi) = t^*(\pi)$ then decode $\hat{x} = \pi(\phi_n(\pi^{-1}(\pi(\hat{y}))))$, and otherwise declare an error.

The conditional average rate of $S^n$ satisfies $\limsup_{n \to \infty} \mathbb{E}[r^*(X) | X \in T_n(Q_X)] = R(Q_X) + \delta$. Since all source blocks in $T_n(Q_X)$ are equiprobable, the conditional error probability satisfies

$$
\liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{P}(\sigma^n_0(s^n_0(X), Y) \neq \hat{X} | X \in T_n(Q_X)) = \liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{P}(\phi_n(Y) \neq \hat{X} | X \in T_n(Q_X))
$$

$$
= E^+(H(Q_X)) - R(Q_X), P_{X|Y}, \min_{Q_X} D(Q_X || P_X) - \delta,
$$

where it should be emphasized that whenever $n \text{ mod } n_0 = 0$ then $\mathbb{P}(\sigma^n_0(s^n_0(X), Y) \neq \hat{X} | X \in T_n(Q_X)) = 0$ by convention. The result follows since $\delta > 0$ was arbitrary. Before concluding the proof, we make the following remark.

**Remark 21**: In the proof, the actual choice of the decoder was implicit since the SW codes are constructed from channel codes. However, as is well known, the optimal decoder in terms of minimum error probability is to decode $\hat{x} \in T_n(Q_X) \cap s^{-1}_n(s_0(x))$ that maximizes $P_{X,Y}(\hat{x} | y)$. Since all $\hat{x} \in s^{-1}_n(s_0(x))$ are in the class $Q_X$, they have the same probability $P_X(\hat{x})$, so this decoding rule is equivalent to maximizing $P_{X,Y}(\hat{x} | y)$, which is the ML decoding rule. Nonetheless, there are cases in which other decoders, such as the minimum conditional entropy decoder, also achieve the same error exponent (see [13] or [43, Appendix C] for a precise definition). This decoder has the merit of not depending on $P_{XY}$ and is therefore a universal decoder.

**Proof of Theorem 3**: Since $|P_n(X)| \leq (n + 1)^{|X|}$, the error probability satisfies

$$
p_e(S_n) = \sum_{Q_X \in P_n(X)} \mathbb{P}(T_n(Q_X)) \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \leq \max_{Q_X \in P_n(X)} e^{-nD(Q_X || P_X)} \cdot \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \leq \exp\left(-n \cdot \min_{Q_X \in P_n(X)} \left\{ D(Q_X || P_X) - \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\} \right). \quad (A.4)
$$

Now, for every $\epsilon > 0$, let $Q_X^* \in P(X)$ be such that

$$
D(Q_X^* || P_X) + \limsup_{n \to \infty} \left\{ \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X^*)) \right\} \leq \inf_{Q_X \in P(X)} \left\{ D(Q_X || P_X) + \limsup_{n \to \infty} \left\{ \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\} + \epsilon \right. \quad (A.5)
$$

and let $m_0$ be sufficiently large so that

$$
\sup_{n > m_0} \left\{ \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X^*)) \right\} \leq \limsup_{n \to \infty} \left\{ \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\} + \epsilon. \quad (A.6)
$$

Then,

$$
\mathcal{E}_e^+(S) = \limsup_{n \to \infty} \min_{Q_X \in P_n(X)} \left\{ D(Q_X || P_X) - \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\} \leq \limsup_{n \to \infty} \inf_{n \geq m_0} \min_{Q_X \in P_n(X)} \left\{ D(Q_X || P_X) - \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\}
$$

$$
= \inf_{Q_X \in P(X)} \left\{ D(Q_X || P_X) + \sup_{n \geq m_0} \left\{ \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\} \right\} \leq D(Q_X^* || P_X) + \sup_{n \geq m_0} \left\{ \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\}
$$

$$
\leq D(Q_X^* || P_X) + \limsup_{n \to \infty} \left\{ \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) \right\} + \epsilon.
$$

(A.7)
where (a) is because, by assumption, if \(T_n(Q_X)\) is empty then \(\mathbb{P}(\hat{X} \neq X | X \in T_n(Q_X)) = 0\), and (b) is from (A.5) and (A.6). The inequality (c) is due to the upper bound of Theorem 2.

\[ \text{Proof of Theorem 4:} \text{ The proof is similar to the proof of Theorem 3 and thus omitted. It can be found in [43, Appendix A, proof of Theorem 4].} \]

\[ \text{Proof of Theorem 5:} \text{ We will use the following two lemmas which can be proved by routine arguments. Their proofs can be found in [43, Appendix A, proofs of Lemma 20 and Lemma 21].} \]

Lemma 22: Let \(Q_X, Q'_X \in P_n(\mathcal{X})\) and assume that
\[ \|Q_X - Q'_X\| = \frac{d^*}{n} \] where \(d^* > 0\). If \(x \in T_n(Q_X)\) then
\[ \min_{x \in T_n(Q'_X)} d_{HI}(x, \hat{x}) \leq d^*. \] (A.8)

Lemma 23: Let \(Q_X \in P_n(\mathcal{X})\) and \(x \in T_n(Q_X)\). For \(1 \leq k < n\) we have
\[ \|\hat{Q}_X - Q_{x(1:n-k)}\| \leq \|X||\frac{k}{n-k}. \] (A.9)

We can now prove Theorem 5. Let \(\epsilon > 0\) be given, and find \(n_0\) sufficiently large such that for any \(Q_X \in P_n(\mathcal{X})\) there exists \(Q_X \in P_{n_0}n(\mathcal{X}) \cap \text{int} Q(\mathcal{X})\) such that \(\|Q_X - Q'_X\| \leq \frac{\epsilon}{2}\). For a given pair of vectors \(x, x' \in \mathcal{X}^n\), define the binary vector \(\Delta_{xx'} = \{i : x(i) \neq x'(i)\}\). Also let \(n_1 = n_0 + 2n_0|\mathcal{X}|\). We construct the following SW codes \(S\) for all \(n > n_1\), 1:

- **Codebook construction:***
  - Compute \(k^*(n) \triangleq n/n_0\).
  - Assign a binary string \(r_1(Q_X)\) for each type in \(T_{n_0}(\mathcal{X})\).
  - Assign a binary string \(r_2(x)\) for each letter \(x \in \mathcal{X}\).
    - For any vector \(x \in \mathcal{X}^m\), define
      \[ r_2(x) \triangleq (r_2(x(1)), \ldots, r_2(x(m))). \] (A.11)
  - Assign a binary string \(r_3(b)\) for each binary vector \(b \in \{0, 1\}^n\) such that \(d_H(0, b) \leq \lfloor \frac{\epsilon}{4} \rfloor\), where \(0\) is the all-zero vector of length \(n\).
  - Construct the SW codes \(S_{k^*(n)-n_0}(Q_X) = (s_{k^*(n)-n_0, Q_X}^*, s_{k^*(n)-n_0, Q_X}^+\) of rate \(\rho(Q_X)\) as in Theorem 2, for all \(Q_X \in P_{n_0}(\mathcal{X}) \cap \text{int} Q(\mathcal{X})\).
  - For any given \(Q_X \in P_n(\mathcal{X})\) find
    \[ \Phi_\epsilon(Q_X) \triangleq \arg \min_{Q_X \in P_{n}(\mathcal{X}) \cap \text{int} Q(\mathcal{X})} \|Q_X - Q'_X\|. \] (A.12)

The above data is revealed to both the encoder and the decoder off-line.

- **Encoding:** Upon observing \(x\), determine its empirical distribution \(Q_x\) and find
  \[ \hat{W} = \arg \min_{W \in T_{n}(n_0)} d_{HI}(x(1 : k^*(n) \cdot n_0), W). \] (A.13)

Let \(x' = x(1 : k^*(n) \cdot n_0)\), and encode the source block \(x\) as:
\[ s_n(x) = \left( \tau_1(\hat{Q}_X), \tau_2(\hat{W}(\mathcal{H}_{x'})), \tau_3(\hat{Q}_X), \tau_2(x(\mathcal{H}_{x'})), \tau_3(\Delta_X), \tau_2(x(k^*(n) \cdot n_0 + 1 : n)), s_{k^*(n)-n_0}(Q_X) (w) \right). \] (A.14)

- **Decoding:** Upon observing \(y\) and \(s_n(x)\):
  - From \(s_n(x)\), recover \(\hat{Q}_X\) and determine \(\Phi_\epsilon(\hat{Q}_X)\). Recover \(\Delta_{x'}, \mathcal{H}_{x'}, x(\mathcal{H}_{x'}),\) and \(x(k^*(n) \cdot n_0 + 1 : n)\).
  - Generate a vector \(\hat{y}' \in \mathcal{X}^{k^*(n)-n_0}\) as follows: For any index \(1 \leq i \leq k^*(n) \cdot n_0\). If \(\Delta_{x'} = 0\) then set \(y'(i) = y(i)\). Otherwise, draw \(y'(i)\) according to the conditional distribution \(P_{Y|X}(y'|x(i))\).
  - Decode
    \[ \hat{w} = s_{k^*(n)-n_0}(\Phi_\epsilon(\hat{Q}_X)) \left( s_{k^*(n)-n_0}(\Phi_\epsilon(\hat{Q}_X))(w), y' \right). \] (A.15)
  - The decoded source block is
    \[ \hat{x}(i) = \begin{cases} \hat{w}(i), & 1 \leq i \leq k^*(n) \cdot n_0, \quad \Delta_{x'}(i) = 0 \\ x(i), & 1 \leq i \leq k^*(n) \cdot n_0, \quad \Delta_{x'}(i) = 1 \\ x(i), & k^*(n) \cdot n_0 + 1 \leq i \leq n. \end{cases} \] (A.16)

To prove that such coding is possible, notice that from Lemma 23 and the fact that \(n > n_1\), we have
\[ \|\hat{Q}_X - Q_x\| \leq \frac{\epsilon}{2} \] (A.17)
and by the triangle inequality
\[ \|\hat{Q}_X - \hat{Q}_W\| \leq \|\hat{Q}_X - Q_x\| + \|Q_x - Q_w\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \] (A.18)

Thus, Lemma 22 implies that
\[ d_{HI}(\Delta_{x'}, \Phi_\epsilon(\hat{Q}_X)) = d_{HI}(\hat{y}', \hat{w}) \leq \left[ \frac{ne}{2} \right]. \] (A.19)

Let us now analyze the resulting asymptotic error probability of \(S\). For any given \(\delta > 0\)
\[ \mathbb{E}_\epsilon(S) \triangleq \lim_{n \to \infty} \min_{Q_X \in P_n(\mathcal{X})} \min_{W \in T_{n}(n_0)} \left\{ D(Q_X || P_X) - \frac{1}{n} \log \mathbb{P}(\hat{X} \neq X | X \in T_{n}(Q_X)) \right\} \]
\[ \triangleq \lim_{n \to \infty} \min_{Q_X \in P_n(\mathcal{X})} \frac{1}{n} \log \mathbb{P}(\hat{W} \neq W | W \in T_{k^*(n)-n_0}(\Phi_\epsilon(Q_X))) \]
\[ = \lim_{n \to \infty} \min_{Q_X \in P_n(\mathcal{X})} \min_{W \in T_{n}(n_0)} \left\{ D(Q_X || P_X) \right\} \] (A.20)
where the passages are explained as follows:

- **Equality (a)** is as in (A.4). Notice that the error event \([\hat{X} \neq X]\) in this equation for the code \(S_n\).

- **Equality (b)** is because an error \(\hat{x} \neq x\) occurs only when the decoder \(\sigma_{k^*(n)n_0}^{*}(Q_X)\) makes an error, since the vector \(v\) is generated memorlessly according to \(P_{Y|X}\), conditioned on \(w\). Notice that the error event \([\hat{W} \neq W]\) in this equation and the following is for the code \(S_{k^*(n)n_0}^{*}(\Phi_{e}(Q_X))\).

- **Inequality (c)** is because there exists \(n_2\) sufficiently large, such that for all \(n > n_2\) the error probability of the decoder \(\sigma_{k^*(n)n_0}^{*}(Q_X)\) satisfies

\[
\frac{k^*(n) \cdot n_0}{n} \log \mathbb{P}(\hat{W} \neq W | W \in T_{k^*(n)n_0}(\Phi_{e}(Q_X))) \\
\geq E^*_k(H(\Phi_{e}(Q_X)) - \rho(\Phi_{e}(Q_X), \Phi_{e}(Q_X), P_{Y|X}) - \delta)
\]

uniformly for all \(Q_X \in \mathcal{P}_{n_0}(Q_X)\) (notice also that \(k^*(n)n_0 \to 1\) as \(n \to \infty\)).

- **Inequality (d)** is because \(D(Q_X||P_X)\) is a continuous function of \(Q_X\) in \(Q(\mathcal{X})\) (as \(supp(P_X) = \mathcal{X}\)), and thus uniformly continuous, and where \(\delta_1 > 0\) and \(\delta_1 \downarrow 0\) as \(\epsilon \downarrow 0\).

Regarding the rate, observe that the resulting codes of \(S\) are type-dependent, variable-rate SW codes, since \(S_{k^*(n)n_0}^{*}(Q_X)\) are such. Let us analyze the total rate required to encode \(x \in Q_X\):

- **Since \(|\mathcal{P}_n(\mathcal{X})| \leq (n+1)^{k^*(n)}\) then for \(n\) sufficiently large \(\frac{1}{n} |\mathcal{P}_n(\mathcal{X})| \leq \frac{|\mathcal{X}|}{n} \cdot \log(n+1) \leq \delta\).** (A.23)

- **Encoding of all possible binary vectors \(b \in \{0, 1\}^{n-k^*(n)n_0}\) such that \(d_B(0, b) \leq \lceil \frac{\varepsilon n}{2} \rceil\) requires a rate of [10, Ch. 13.2]**

\[
\frac{1}{n} |r_2(\mathcal{W}_n^{(k^*(n)n_0)}) + r_2(\mathcal{X}_n^{(k^*(n)n_0)})| \leq \delta
\]

for \(n\) sufficiently large.

- **Encoding the components of \(x(\mathcal{H}_n^{(k^*(n)n_0)})\) and \(w(\mathcal{H}_n^{(k^*(n)n_0)})\) letter-wise with zero error, requires a rate of**

\[
\frac{1}{n} |r_2(\mathcal{X}_n^{(k^*(n)n_0)}) + r_2(\mathcal{W}_n^{(k^*(n)n_0)})| \leq \frac{\epsilon}{2} \log |\mathcal{X}| + \frac{\delta}{n} \leq \log |\mathcal{X}| + \delta
\]

for \(n\) sufficiently large.

- **By construction, \(S_{k^*(n)n_0}^{*}(\Phi_{e}(Q_X))\) is a type-dependent, variable-rate SW code of rate \(\rho(\Phi_{e}(Q_X))\) and thus for sufficiently large \(n\)**

\[
\frac{1}{n} |s_{k^*(n)n_0}^{*}(\Phi_{e}(Q_X))(w)| \leq \frac{k^*(n) \cdot n_0}{n} \rho(\Phi_{e}(Q_X)) + \delta
\]

(A.27)

uniformly over \(Q_X\).
codes which assign rate \( \rho(\Phi_{x}(Q_{X})) + q \) to the type \( Q_{X} \), and (c) is again by the uniform continuity of \( D(Q_{X} \| P_{X}) \) in \( Q(\mathcal{X}) \). We obtain the desired result by taking \( \delta \downarrow 0 \) and then \( \epsilon \downarrow 0 \).

Before completing the proof, we make the following two remarks.

Remark 24: The vector actually coded is \( w \) (A.13), not the original source block \( x \). Thus, after modifying \( x \) to \( w \), the distribution of \( w \) may not be uniform within its type class (even when conditioned on the event that \( x \) belongs to some type class), which might affect (A.20). There are two possibilities to circumvent this.\(^{14}\) The first is to use common randomness at the encoder and decoder, and to generate a uniformly random permutation. Prior to encoding, the source block \( x \) is permuted, and the decoder simply applies the inverse permutation after decoding. In this case, the uniform distribution of \( w \) is assured. The second possibility is to construct the SW codes from channel codes (as was done in Theorem 2) which have maximal error probability according to the reliability function (see (A.1)), and not just the average error probability. As is well known, such a channel code can be generated from a good average error probability codebook, by simply expurgating the worst half of the codebook. The rate loss is negligible, and here too, good error probability is assured uniformly over \( w \) in the type class.

Remark 25: In the proof above, the actual decoders \( \sigma_{n}^{*}Q_{X} \) of \( S_{n}^{*}(Q_{X}) \) were not specified, and any decoder which achieves the error exponent for the underlying channel code can be used. Thus, in the proof of Theorem 5, a randomized decoder was required, in order to mimic the channel operation \( P_{Y|X} \) for the vector \( w \). However, this might not be required if \( \sigma_{n}^{*}Q_{X} \) is more specific. For example, if the decoder \( \sigma_{w}^{*}Q_{X} \) is the ML decoder, then instead of drawing \( y(i) \) according to the conditional distribution \( P_{Y|X}(\cdot|w(i)) \), it can be simply set to the letter with maximal likelihood, i.e., \( y(i) = \arg\max_{y \in \mathcal{Y}} P_{Y|X}(y|w(i)) \). This only improves the error probability, and thus the results of Theorem 5 remain valid.

Proof of Theorem 11: We will only prove the random binning bound, as the derivations for the expurgated and sphere packing bounds follow the same lines. (See [43, Appendix A, proof of Theorem 10] for details).

From Theorem 5, we may clearly assume that \( \rho_{b}(Q_{X}, E_{e}) \leq H(Q_{X}) \), as otherwise the random coding bound in (17) is infinite, and \( E_{e} \) is trivially achieved. Now, from the random coding bound in (17), the condition in (37) will be satisfied for a rate function \( \rho_{b}(Q_{X}, E_{e}) \) which satisfies

\[
E_{e} \leq D(Q_{X} \| P_{X}) + \min_{Q_{Y|X}} \left[ D(Q_{Y|X} \| P_{Y|X} | Q_{X}) + \rho_{b}(Q_{X}, E_{e}) - H(Q_{X} | Q_{Y}) \right] \geq 0.
\]

(A.30)

Clearly, if \( E_{e} \leq D(Q_{X} \| P_{X}) \), no actual constraint is imposed on the rate, and (A.30) is satisfied even for \( \rho_{b}(Q_{X}, E_{e}) = 0 \).

\(^{14}\)This matter was not addressed in the body of the proof in order not to over-complicate it.

Otherwise, (A.30) is equivalent to

\[
E_{e} \leq D(Q_{X} \| P_{X}) + \min_{Q_{Y|X}} \left[ D(Q_{Y|X} \| P_{Y|X} | Q_{X}) + \rho_{b}(Q_{X}, E_{e}) - H(Q_{X} | Q_{Y}) \right] \geq 0.
\]

(A.31)

or

\[
\rho_{b}(Q_{X}, E_{e}) \geq \min_{Q_{Y|X}} \left[ \frac{E_{e} - D(Q_{Y|X} \| P_{Y|X})}{\lambda} + H(Q_{X} | Q_{Y}) \right] = \max_{Q_{Y|X} \in A} \left\{ E_{e} - D(Q_{Y|X} \| P_{X}) - D(Q_{Y|X} \| P_{Y|X} | Q_{X}) + H(Q_{X} | Q_{Y}) \right\}
\]

(A.32)

which directly leads to the second term in (55). For the third term in (55), let us notice that for \( E_{e} \geq D(Q_{X} \| P_{X}) + D(Q_{Y|X} \| P_{Y|X} | Q_{X}) \) we have that \( \rho_{b}(Q_{X}, E_{e}) \) is affine with slope 1. Indeed, using (55) we get for \( E_{e} > D(Q_{X} \| P_{X}) \)

\[
\rho_{b}(Q_{X}, E_{e}) \geq E_{e} + H(Q_{X}) - D(Q_{X} \| P_{X}) - I(Q_{X} \times Q_{Y} | X) + D(Q_{Y|X} \| P_{Y|X} | Q_{X})
\]

(A.33)

and for \( E_{e} \geq D(Q_{X} \| P_{X}) + D(Q_{Y|X} \| P_{Y|X} | Q_{X}) \) equality is achieved since \( Q_{Y|X} \in A_{n} \). For the fourth term in (55), notice that the minimal \( E_{e} \) such that \( \rho_{b}(Q_{X}, E_{e}) = H(Q_{X}) \) is given by

\[
E_{e} = D(Q_{X} \| P_{X}) + I(Q_{X} \times Q_{Y} | X) + D(Q_{Y|X} \| P_{Y|X} | Q_{X}),
\]

(A.34)

Proof of Lemma 12: These properties are simple to derive. For details, see [43, Appendix A, Proof of Lemma 11].

Proof of Lemma 13: This can be proved if we show that the infimum of \( \inf_{Q_{X} \in \mathcal{V}} D(Q_{X} \| P_{X}) \) is attained, and that the function \( \min_{Q_{X} \in \mathcal{V}} \rho_{b}(Q_{X}, E_{e}) \geq R \) \( D(Q_{X} \| P_{X}) \) is left-continuous in \( R \). We begin by showing that the infimum of \( \inf_{Q_{X} \in \mathcal{V}} D(Q_{X} \| P_{X}) \) is attained. Recall that \( \rho(Q_{X}) \) is regular, and so there exists a \( d > 0 \) such that \( \rho(Q_{X}) \) is continuous in \( V = \{ Q_{X} \in \mathcal{V} | D(Q_{X} \| P_{X}) < d \} \), and equals a constant \( \rho(Q_{X}) = R_{0} \), for \( Q_{X} \in \mathcal{V}^{c} \). Thus,

\[
\inf_{Q_{X} \in \mathcal{V}^{c}} D(Q_{X} \| P_{X}) = \min_{Q_{X} \in \mathcal{V}^{c}} \inf_{\rho(Q_{X}) \geq R} D(Q_{X} \| P_{X}) = \min_{Q_{X} \in \mathcal{V}^{c}} \inf_{\rho(Q_{X}) \geq R} D(Q_{X} \| P_{X})
\]

(A.35)

and so, if \( \inf_{Q_{X} \in \mathcal{V}^{c}} D(Q_{X} \| P_{X}) \) is not attained, then the infimum of \( \inf_{Q_{X} \in \mathcal{V}} \rho_{b}(Q_{X}, E_{e}) \geq R \) \( D(Q_{X} \| P_{X}) \) is not attained for some \( Q_{X} \in V \), and so

\[
\inf_{Q_{X} \in V} D(Q_{X} \| P_{X}) = d.
\]

(A.36)

However, in this case, there also must exist a sequence \( \hat{Q}^{(n)}_{X} \in V \) such that \( \rho(\hat{Q}^{(n)}_{X}) \to R \) and \( \rho(\hat{Q}^{(n)}_{X}) \geq R \). But
since $D(\mathcal{Q}^{(n)}_X || P_X) < d$ this is a contradiction side of (68) then

$$\min_{Q_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X) \geq D(\mathcal{Q}^{(n)}_X || P_X) - \delta$$

Now, to show left continuity of $\min_{Q_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X)$ as a function of $R$, let $\delta > 0$ be given. For any $\varepsilon > 0$ we clearly have

$$\min_{Q_X: \rho(\mathcal{Q}_X) \geq R-\varepsilon} D(\mathcal{Q}_X || P_X) \leq \min_{Q_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X) = D(\mathcal{Q}^{(n)}_X || P_X). \quad (A.37)$$

To obtain the reversed inequality, we divide the proof into two cases, depending on whether $\mathcal{Q}^{(n)}_X \in \mathcal{V}$ or not.

**Case 1:** $Q_X^* \in \mathcal{V}$. Recall that $\rho(\mathcal{Q}_X) = \rho(\mathcal{Q}_X; R)$ is continuous and finite inside the interior of $\mathcal{V}$, and $D(\mathcal{Q}_X || P_X)$ is a continuous function of $Q_X$. Now, we may define for any $Q_X \in \mathcal{V}$ such that $\rho(\mathcal{Q}_X) \geq R$, the closed neighborhood

$$D(\mathcal{Q}_X, R, \delta) \triangleq \{ \hat{Q}_X : D(\hat{Q}_X || P_X) \geq D(\mathcal{Q}_X || P_X) - \delta \} \cap \mathcal{V}. \quad (A.38)$$

Also, we may define the set

$$\mathcal{V}'(R) \triangleq \left\{ Q_X \in \partial \mathcal{V} : \lim_{\hat{Q}_X \to Q_X} \rho(\hat{Q}_X) = R \right\}, \quad (A.39)$$

where $\partial \mathcal{V} = \mathcal{I} \mathcal{V}$ is the boundary of $\mathcal{V}$, and for any $Q_X \in \mathcal{V}'(R)$

$$D'(\mathcal{Q}_X, R, \delta) \triangleq \{ Q_X \} \cup D(\mathcal{Q}_X, R, \delta). \quad (A.40)$$

Now, consider the set

$$\mathcal{U} \triangleq \mathcal{V} \setminus \bigcup_{\{Q_X \in \mathcal{V}'(R)\}} D'(\mathcal{Q}_X, R, \delta)$$

and let $R' \triangleq \sup_{Q_X \in \mathcal{U}} \rho(\mathcal{Q}_X)$. Then we must have $R' < R$. To see this, assume conversely, that $R' = R$ and let $\mathcal{Q}_X \in \mathcal{U}$ achieve the maximum, namely, $\rho(\mathcal{Q}_X) = R$. Now, either $\mathcal{Q}_X \in \mathcal{V}$ or the supremum is not attained, but both cases lead to contradiction. Indeed, if the supremum is attained at some $\mathcal{Q}_X \in \mathcal{V}$ then $D(\mathcal{Q}_X, R, \delta) \in \mathcal{U}$ and so $\mathcal{Q}_X \notin \mathcal{U}$ which is a contradiction. Otherwise, there exists a sequence $\mathcal{Q}^{(n)}_X \in \mathcal{U}$ such that $\rho(\mathcal{Q}^{(n)}_X) \to R$. Assume that an arbitrary convergent sub-sequence of $\mathcal{Q}^{(n)}_X$ converges to $\mathcal{Q}_X \in \mathcal{V}$. But, the definition of $D'(\mathcal{Q}_X, R, \delta)$ and the continuity of $D(\mathcal{Q}_X || P_X)$ in $\mathcal{V}$ imply that for any sufficiently large $n$ we must have $\mathcal{Q}^{(n)}_X \notin \mathcal{U}$, which is a contradiction. Now, consider two sub-cases:

1) $R > R_0$. If we choose $\varepsilon < \min\{R - R', R - R_0\}$ we have

$$\min_{Q_X: \rho(\mathcal{Q}_X) \geq R-\varepsilon} D(\mathcal{Q}_X || P_X) \geq \min_{\mathcal{Q}_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X) \geq D(\mathcal{Q}^{(n)}_X || P_X) - \delta \quad (A.42)$$

since the left most minimization is over a smaller set.

2) $R \leq R_0$. Since $\mathcal{Q}^*_X$ is the minimizer for the right hand side of (68) then

$$\min_{\mathcal{Q}_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X) \geq D(\mathcal{Q}^*_X || P_X) \quad (A.43)$$

and if we choose $\varepsilon \leq R - R'$ we also have

$$\min_{\mathcal{Q}_X: \rho(\mathcal{Q}_X) \geq R-\varepsilon} D(\mathcal{Q}_X || P_X) \geq \min_{\mathcal{Q}_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X) \geq D(\mathcal{Q}^*_X || P_X) - \delta. \quad (A.44)$$

**Case 2:** $Q_X^* \notin \mathcal{V}$. In this case we clearly have $R_0 \in \mathcal{V}$ and

$$\inf_{Q_X: \rho(\mathcal{Q}_X) \geq R} \min_{\mathcal{Q}_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X) = d. \quad (A.45)$$

Now, if we let $R \triangleq \sup_{Q_X \in \mathcal{V}} \rho(\mathcal{Q}_X)$, then either this supremum is not attained or $R \in \mathcal{V}$. To see this, assume conversely, that the supremum is attained by some $Q_X \in \mathcal{V}$ and also $R \notin \mathcal{V}$. Then this implies

$$\inf_{Q_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X) \leq D(\mathcal{Q}^*_X || P_X) \leq d \quad (A.46)$$

which is a contradiction. Now, we have two sub-cases:

1) If $R \in \mathcal{V}$, we can choose $\varepsilon = \mathcal{R} - R' > 0$ and obtain

$$\min_{Q_X: \rho(\mathcal{Q}_X) \geq R-\varepsilon} D(\mathcal{Q}_X || P_X) \geq D(\mathcal{Q}^*_X || P_X) \quad (A.47)$$

2) Otherwise, suppose that $\sup_{Q_X \in \mathcal{V}} \rho(\mathcal{Q}_X)$ is not attained and $R \notin \mathcal{V}$. If $R \in \mathcal{V}$ then there exists a sequence $\mathcal{Q}^{(n)}_X \in \mathcal{V}$ such that $\rho(\mathcal{Q}^{(n)}_X) \to R$, and so there exists $n_0$ such that $\rho(\mathcal{Q}^{(n)}_X) > R$ which contradicts the optimality of $\mathcal{Q}^*_X$, and so we must have $R \notin \mathcal{V}$. In this case, $\rho(\mathcal{Q}^*_X) < R$ for all $Q_X \in \mathcal{V}$, so define

$$\mathcal{W} \triangleq \{ Q_X \in \mathcal{Q}(X) : D(\mathcal{Q}_X || P_X) \leq d - \delta \} \quad (A.48)$$

and let $R' \triangleq \max_{Q_X \in \mathcal{W}} \rho(\mathcal{Q}_X)$, where clearly $R' < R$. Then, for $\varepsilon = R - R' > 0$

$$\min_{Q_X: \rho(\mathcal{Q}_X) \geq R-\varepsilon} D(\mathcal{Q}_X || P_X) \geq d - \delta = D(\mathcal{Q}^*_X || P_X) - \delta. \quad (A.49)$$

To conclude, in both cases, for any given $\delta > 0$ we can find $\varepsilon > 0$ such that

$$\min_{Q_X: \rho(\mathcal{Q}_X) \geq R-\varepsilon} D(\mathcal{Q}_X || P_X) \geq D(\mathcal{Q}^*_X || P_X) - \delta. \quad (A.50)$$

This means that $\min_{Q_X: \rho(\mathcal{Q}_X) \geq R} D(\mathcal{Q}_X || P_X)$ is left-continuous as a function of $R$, and the desired result is obtained.

**Proof of Lemma 14:** These properties are simple to derive. For details, see [43, Appendix A, Proof of Lemma 13].

**Proof of Theorem 16:** We will only prove the random binning bound, as the derivations for the expurgated and sphere packing bounds follow the same lines (See [43, Appendix A, Proof of Th. 15] for details).

For any given $(R, E_r)$ we may use the condition of Lemma 15. Notice that $\hat{\rho}(\mathcal{Q}_X; R, E_r)$ is a regular rate function, and so the excess-rate exponent in Lemma 13 is applicable. From Theorem 5 and the random coding bound in (17), the rate function $\rho(\mathcal{Q}_X; R, E_r)$ will achieve infimum
error exponent $E_e$ if
\[
E_e \leq D(Q_X||P_X) + \min_{Q_{Y|X}} \left\{ D(Q_{Y|X}||P_{Y|X}|Q_X) + \left[ R - H(Q_{Y|X}|Q_Y) \right] \right\}
\]
\[
= \min_{Q_X} \left\{ D(Q_X||P_X) + \min_{Q_{Y|X}} \left[ D(Q_{Y|X}||P_{Y|X}|Q_X) + \left[ R - H(Q_{Y|X}|Q_Y) \right] \right] \right\}
\]
\[
\leq \min_{Q_X} \left\{ D(Q_X||P_X) + \min_{0 \leq t \leq 1} \left[ D(Q_X||P_X) + D(Q_{Y|X}||P_{Y|X}|Q_X) + t \left[ R - H(Q_{Y|X}|Q_Y) \right] \right] \right\}
\]
\[
\equiv \max_{0 \leq t \leq 1} \min_{Q_X} \left\{ \min_{Q_{Y|X}} \left[ D(Q_X||P_X) + D(Q_{Y|X}||P_{Y|X}|Q_X) + t \left[ R - H(Q_{Y|X}|Q_Y) \right] \right] \right\}
\]
where (a) is because the minimization problem in (A.52) is convex in $Q_X$ (over the convex set $\{Q_X \in Q(X) : D(Q_X||P_X) \leq E_e\}$) and $\{Q_{Y|X}\}$, and the maximization problem is linear in $t$ (over the convex set $[0,1]$), and thus also concave. Therefore, we can interchange the maximization and minimization [39] order, and obtain the condition
\[
\max_{0 \leq t \leq 1} E_e(t) \geq E_e.
\]

Proof of Lemma 17: Introducing an auxiliary PMF $\tilde{Q}_Y$ and using [10, Lemma 10.8.1] we get that
\[
\nu_{ab}(P_{XY}, Q_X, E_e, \eta) = \min_{Q_{Y|X}: D(Q_{Y|X}) \leq E_e} \min_{Q_Y} \left\{ D(Q_{Y|X}||\tilde{Q}_Y|Q_X) + \eta \cdot D(Q_{Y|X}||P_{Y|X}|Q_X) \right\}
\]
\[
= \min_{Q_{Y|X}} \left\{ D(Q_{Y|X}||\tilde{Q}_Y|Q_X) + \eta \cdot D(Q_{Y|X}||P_{Y|X}|Q_X) \right\}
\]
\[
= \min_{Q_{Y|X}} \left\{ D(Q_{Y|X}||\tilde{Q}_Y|Q_X) + \eta \cdot D(Q_{Y|X}||P_{Y|X}|Q_X) \right\}
\]
Notice that (A.54) is an optimization problem over $(Q_{Y|X}, \tilde{Q}_Y)$ and consider utilizing an alternating minimization algorithm, where for a given $\tilde{Q}_Y$, the minimizer $Q_{Y|X}$ is found, and vice versa. We divide the rest of the proof into two main parts. In the first part, we prove that the alternating minimization algorithm indeed converges to the optimal solution, and in the second part, we solve the two individual optimization problems (resulting from keeping one of the optimization variables fixed).

Part 1: In [17, Sec. 5.2] and [18] sufficient conditions were derived for the convergence of an alternating minimization algorithm. Specifically, these conditions are met for a minimization problem of the form
\[
\inf_{Q_1 \in Q_1} \inf_{Q_2 \in Q_2} D(Q_1||Q_2)
\]
where $Q_1$ and $Q_2$ are two positive measures (which may not necessarily sum to 1) over a finite alphabet $Z$, and $Q_1$, $Q_2$ are two convex sets. To prove that alternating minimization algorithm converges for the optimization problem (A.54), we now show that it can be written in the form of (A.55). The objective function of (A.54) is given by
\[
D(Q_{Y|X}||\tilde{Q}_Y|Q_X) + \eta \cdot D(Q_{Y|X}||P_{Y|X}|Q_X)
\]
\[
= \sum_{x,y} Q_X(x) Q_{Y|X}(y|x) \log \frac{Q_{Y|X}(y|x)}{Q_Y(y) P_{Y|X}(y|x)}
\]
\[
= (1 + \eta) \sum_{x,y} Q_{XY}(x, y)
\]
\[
\times \log \frac{Q_{XY}(x, y)}{\tilde{Q}_Y(y) \tilde{Q}_Y(y) P_{Y|X}(y|x)}.
\]
Thus, if we let $Z = X \times Y$ and consider the measures $Q_{XY}$ and $\tilde{Q}_Y$, then the objective function is of the form of (A.55). Now, the feasible set for $Q_{XY}$ is
\[
\left\{ Q_{XY} : \sum_{y \in Y} Q_{XY}(x, y) = Q_X(x), D(Q_{XY}||P_{XY}) \leq E_e \right\}
\]
which is a convex set. Now, it is easy to show that [10, Lemma 10.8.1] hold even if the feasible region of $\tilde{Q}_Y$ is extended from the simplex $Q(Y)$ to the set
\[
\tilde{Q}(Y) \triangleq \left\{ \tilde{Q}_Y : \sum_{y \in Y} \tilde{Q}_Y(y) \leq 1, \tilde{Q}_Y(y) \geq 0 \text{ for all } y \in Y \right\}
\]
which is also a convex set. Now, define the feasible set for the variables $\tilde{Q}_{XY}$ as
\[
\tilde{Q} \triangleq \left\{ \tilde{Q}_{XY} : \exists \tilde{Q}_Y \in \tilde{Q}(Y) \text{ so that } \forall (x, y) \in X \times Y, \tilde{Q}_{XY}(x, y) = \left[ \tilde{Q}_Y(y) \right]^{1+\eta} P_{Y|X}(y|x) \right\}
\]
We show that $\tilde{Q}$ is also a convex set. Let
\[
\tilde{Q}_{i,XY}(x, y) = \left[ \tilde{Q}_{i,Y}(y) \right]^{1+\eta} P_{Y|X}(y|x)\right\}^{1+\eta} Q_X(x)
\]
for $\tilde{Q}_{i,Y} \in \tilde{Q}(Y)$, $i = 0, 1$, and $0 \leq \alpha \leq 1$. Then,
\[
\tilde{Q}_{a,XY} \triangleq (1 - a) \tilde{Q}_{0,XY} + a \tilde{Q}_{1,XY}
\]
\[
= P_{Y|X} Q_X \cdot \left( \left( 1 - a \right) \tilde{Q}_{0,Y} + a \tilde{Q}_{1,Y} \right).
\]
Thus, to show that $\tilde{Q}_{a,XY} \in \tilde{Q}$ all is needed to prove is that $\tilde{Q}_{a,Y} \triangleq \left( (1 - a) \tilde{Q}_{0,Y} + a \tilde{Q}_{1,Y} \right)^{1+\eta} \in \tilde{Q}(Y)$.
As positivity of $\hat{Q}_{x,y}$ is clear, it remains to verify that $\sum_{y \in \mathcal{Y}} \hat{Q}_{x,y}(y) \leq 1$. Indeed, we have
\[
\sum_{y \in \mathcal{Y}} \hat{Q}_{x,y}(y) = \sum_{y \in \mathcal{Y}} \left( (1 - \alpha) \hat{Q}_{0,y}^{1+\eta} + \alpha \hat{Q}_{1,y}^{1+\eta} \right) \leq (1 - \alpha) \left( \sum_{y \in \mathcal{Y}} \hat{Q}_{0,y}(y) \right)^{1+\eta} + \alpha \left( \sum_{y \in \mathcal{Y}} \hat{Q}_{1,y}(y) \right)^{1+\eta} \leq [(1 - \alpha) + \alpha]^{1+\eta} = 1
\] (A.62)
where \((a)\) follows from a variant of Minkowski’s inequality [42, Sec. 3A.1, inequality \((k)\)], and \((b)\) is from the fact that both $t^{1+\eta}$ and $t^{1+\eta}$ are increasing functions of $t \in \mathbb{R}_+$ when $\eta \geq 0$, and $\hat{Q}_t \in \hat{Q}(\mathcal{Y})$. Thus the optimization problem \((A.54)\) is of the form \((A.55)\) and an alternating minimization algorithm converges to the optimal, unique, solution, which we denote by $(\hat{Q}_{X,Y}^*, \hat{Q}_Y^*)$.

Part 2: First, suppose that $\hat{Q}_Y$ is given. The minimizer $\hat{Q}_{x|Y}$ can be found by using the Karush-Kuhn-Tucker (KKT) conditions for convex problems [4, Sec. 5.5.3]. The resulting solution satisfies that for any $x \in \mathcal{X}$ such that $Q_x(x) \neq 0$
\[
\hat{Q}_{x|Y}(y|x) = P_{Y|x}(y|x)^{\alpha} \left[ \hat{Q}_Y(y) \right]^{1-\alpha} \tag{A.63}
\]
for some $\frac{\eta}{1+\eta} \leq \alpha \leq 1$, where $P_{Y|x}$ is a normalization constant.

Clearly, from \((93)\) we have $Q_{x|Y}^* = \mathbb{M}_{\hat{Q}_Y} P_{Y|x} \hat{Q}_Y \alpha^*$. The value of $Q_{x|Y}^*$ for $x \in \mathcal{X}$ with $Q_x(x) = 0$ is immaterial as it does not affect the optimal value of the objective function. Also, it is evident that the solution $Q_{x|Y}^*$ is indeed positive.

To find the optimal $\hat{Q}_{Y|X}$, we need to choose $\alpha$ in order to satisfy the constraint $D(\hat{Q}_{y|X}^* || P_{Y|x}) \leq \mathcal{E}_r$. For this, the complementary slackness condition [4, Sec. 5.5.2] implies that $\alpha$ should be chosen either to satisfy
\[
D(\hat{Q}_{y|X}^* || P_{Y|x}) = \mathcal{E}_r - D(Q_x || P_x) \tag{A.64}
\]
and then $\frac{\eta}{1+\eta} < \alpha \leq 1$, or $\alpha = \frac{\eta}{1+\eta}$ and then
\[
D(\hat{Q}_{y|X}^* || P_{Y|x}) < \mathcal{E}_r - D(Q_x || P_x). \tag{A.65}
\]
To find $\alpha^*$ that satisfies the complementary slackness condition, note that $D(\hat{Q}_{y|X}^* || P_{Y|x})$ is a monotonically decreasing function of $\alpha$. Indeed, it is easy to see that if $\hat{Q}_Y$ is initialized such that
\[
\text{supp}(\hat{Q}_Y) = \text{supp} \left( \sum_{x \in \mathcal{X}} Q_x(x) P_{Y|x}(y|x) \right) \tag{A.66}
\]
then this remains true for all iterations. Then, it follows from [15, Problem 2.14(c)] that for any given $x \in \mathcal{X}$ such that $Q_x(x) \neq 0$, we have that $D(\hat{Q}_{y|X}^* || P_{Y|x}(x|x))$ is a decreasing function of $\alpha$, and thus their average $D(\hat{Q}_{y|X}^* || P_{Y|x} || Q_x)$ is also a decreasing function of $\alpha$. Thus, if for $\alpha = \frac{\eta}{1+\eta}$ we have $D(\hat{Q}_{y|X}^* || P_{Y|x} || Q_x) < \mathcal{E}_r - D(Q_x || P_x)$ then $\alpha^* = \frac{\eta}{1+\eta}$. Otherwise, we have $D(\hat{Q}_{y|X}^* || P_{Y|x} || Q_x) > \mathcal{E}_r - D(Q_x || P_x)$ and $D(\hat{Q}_{y|X}^* || P_{Y|x} || Q_x) = 0 < \mathcal{E}_r - D(Q_x || P_x)$. Thus, in the later case, a simple search finds $\alpha^*$.

Second, assume that $Q_{Y|X}$ is given. The minimizer $\hat{Q}_Y$ can be found using [10, Lemma 10.8.1] to be
\[
\hat{Q}_Y(y) = \sum_{x \in \mathcal{X}} Q_x(x) \hat{Q}_Y(x|y). \tag{A.67}
\]
It is easily seen that Algorithm 1 indeed implements the procedure described in this proof. \hfill $$

**Proof of Lemma 18:** Introducing an auxiliary PMF $\hat{Q}_Y$ and using [10, Lemma 10.8.1], we get that
\[
e_{\text{th}}(P_{XY}, \mathcal{R}, \mathcal{E}_r, t) \triangleq \min_{\hat{Q}_Y} \min_{Q_x : D(Q_x || P_x) \leq \mathcal{E}_r} \min_{Q_{Y|x}} \left[ D(Q_x || P_x) + D(Q_{Y|x} || P_{Y|x} || Q_x) + t \cdot \left( \mathcal{R} - H(Q_x) + D(Q_{Y|x} || \hat{Q}_Y || Q_x) \right) \right]. \tag{A.68}
\]
Now, Algorithm 2 is an alternating minimization algorithm, that keeps all parameters but one fixed, and optimizes over the non-fixed parameter. Now, for a given $t \geq 0$, the objective function in \((76)\) is given by
\[
(1+t) \sum_{x,y} Q_{XY}(x,y) \log \frac{Q_{XY}(x,y)}{P_{XY}(x,y)} \left[ \hat{Q}_Y(y) \right]^{1+\eta} + t \mathcal{R}. \tag{A.69}
\]
The same technique that was used in the proof of Lemma 17, shows that this optimization problem is of the form \((A.55)\) (with additional constant $\mathcal{R}$). Thus, an alternating minimization algorithm converges to the optimal solution.

We now turn to the minimization of individual variables, assuming that all other variables are fixed, for a given $t \geq 0$. First, consider the minimization over $Q_{XY}$, which itself can be separated to an unconstrained minimization over $Q_{Y|X}$ and a constrained minimization over $Q_X$. The minimizer $\hat{Q}_{Y|X}$ can be found using standard Lagrange optimization methods. The result is $\hat{Q}_{Y|X} = \mathbb{M}_{\hat{Q}_Y} P_{Y|x} \hat{Q}_Y \alpha^*$, for all $x \in \mathcal{X}$ such that $Q_X(x) \neq 0$, and arbitrary otherwise, since the value of $Q_{Y|X}^*(y|x)$ for $x \in \mathcal{X}$ such that $Q_X(x) = 0$ does not affect the value of the optimization problem. For this optimal choice, using the definitions of $h_1, \ldots, h_N$ we obtain
\[
\min_{Q_x} \min_{Q_{Y|x} : D(Q_{Y|x} || P_{Y|x}) \leq \mathcal{E}_r} \left[ D(Q_x || P_x) + \sum_{x \in \mathcal{X}} Q_X(x) h_1(x) \right]
\]
\[
+ t \cdot \left( \mathcal{R} - H(Q_x) + \sum_{x \in \mathcal{X}} Q_X(x) h_2(x) \right). \tag{A.70}
\]
Next, we optimize over $Q_X$ using the KKT conditions, and the result is
\[
Q_{X}^*(x) = \psi \cdot \left[ P_{X}(x) \right]^{1+\eta} \times \exp \left( \frac{1}{1+\eta} \cdot h_1(x) - \frac{t}{1+\eta} \cdot h_2(x) \right) \tag{A.71}
\]
where $\psi$ is a normalization constant, and $\lambda \geq 0$. From the definition (99), it is evident that $Q^\star_{X|Y} = \text{Min}(P_X, h_1, h_2, \lambda, t)$. Using the complementary slackness condition [4, Sec. 5.5.2], $\lambda$ should be found such that either $D(Q^\star_{X|Y}||P_X) = E_\epsilon$ or $\lambda = 0$. It can be verified from [15, Problem 2.14(c)] that $D(Q^\star_{X|Y}||P_X)$ is a monotonic decreasing function of $\lambda$ and thus the above search is relatively simple. To see this, observe that initializing $\tilde{Q}_{X|Y}$ with support $Y$ implies that in the first iteration $\text{supp}(Q^\star_{X|Y}) = \text{supp}(P_{Y|X})$ which assures that $h_1, h_2, \lambda, t$ are finite. As $\text{supp}(Q^\star_{X|Y}) = \text{supp}(P_X) = \mathcal{X}$ for all $\lambda > 0$ and $t \geq 0$ then $\text{supp}(\tilde{Q}_{Y}) = \mathcal{Y}$ for all iterations (cf. (A.76)). Thus, for any $t \not= 0$ we may express $Q^\star_{X|Y}$ as

$$Q^\star_{X|Y} = \psi \cdot \left[ P_X \right]^{\frac{1-h_1(\lambda)}{t}} \cdot \tilde{P}_X(x),$$

where

$$\tilde{P}_X(x) = \psi \cdot \exp \left( -\frac{h_1(\lambda)}{t} - h_2, (x) \right).$$

and $\psi$ is a normalization factor. Setting $\alpha = \frac{1-h_1(\lambda)}{t}$ we get that $D(Q^\star_{X|Y}||P_X)$ is a decreasing function of $\alpha$. Since $\alpha$ is a monotonically increasing function of $\lambda$ this implies that $D(Q^\star_{X|Y}||P_X)$ is also a decreasing function of $\lambda$. For $t = 0$ we may write again

$$Q^\star_{X|Y} = \psi \cdot P_X^{\frac{1-h_1(\lambda)}{t}} \cdot \tilde{P}_X(x),$$

where now

$$\tilde{P}_X(x) = \psi \cdot \exp (-h_1, (x) + \log P_X(x)).$$

Similar arguments show that $D(Q^\star_{X|Y}||P_X)$ is a decreasing function of $\lambda$.

The optimal $\tilde{Q}^\star_{Y}$ for a given $t$ and $Q_X, Q_{Y|X}$ is simply

$$\tilde{Q}^\star_{Y}(y) = \sum_{x \in \mathcal{X}} Q_X(x) Q_{Y|X}(y|x),$$

using [10, Lemma 10.8.1].

**Proof of Lemma 19:** Clearly (92) can be written as

$$\nu_{ex}(P_{XY}, Q_X, E_{\epsilon}) = \min_{\hat{Q}_{X|Y} \in G} D(\hat{Q}_{X|Y}||Q_X \times Q_X),$$

where $G \equiv G_1 \cap G_2 \cap G_3$ and

$$G_1 = \{ Q_{X|Y} : B(\hat{Q}_{X|Y}) = E_{\epsilon} + D(Q_X||P_X) \},$$

$$G_2 = \{ \hat{Q} = Q_X \},$$

$$G_3 = \{ \hat{Q} = Q_X \}.$$

It may be easily seen that $\{G_i\}_{i=1}^3$ are linear families. In [17, Th. 5.1], the convergence of an iterative algorithm for a minimization problem of the form

$$\min_{\hat{Q} \in G} D(\hat{Q}||Q),$$

where $G$ is the intersection of a finite number of linear families, was proved. The minimizer of (A.81) is called the I-projection of $Q$ onto $G$ and denoted $\hat{Q}^*$. The algorithm is called iterative scaling and works as follows: First $\hat{Q}^{(0)} = Q$ is initialized. Then, $\hat{Q}^{(1)}$ is the I-projection of $\hat{Q}^{(0)}$ onto $G_1$, $\hat{Q}^{(2)}$ is the I-projection of $\hat{Q}^{(1)}$ onto $G_2$, and so on, where for $n > L$, $\hat{Q}^{(n)}$ is the I-projection of $\hat{Q}^{(n-1)}$ onto $G_{n \mod L}$. Such a procedure converges to $\hat{Q}^*$.

Thus, to use the iterative scaling algorithm for the case at hand, we initialize $\hat{Q}^{(0)}(x, \bar{x}) = Q_X(x) Q_{X|Y}^{\bar{x}}(\bar{x})$ for all $(x, \bar{x}) \in \mathcal{X} \times \mathcal{X}$, and then we need to find for any given PMF $\hat{Q}_{X|Y}$ the I-projections

$$\min_{\hat{Q}_{X|Y} \in G} D(\hat{Q}_{X|Y}||Q_X \times Q_{X|Y}).$$

for $i = 1, 2, 3$. In what follows we will perform the I-projection on $G_1 \cap G_2$ jointly, and then on $G_3$. In this case, $\hat{Q}_{X|Y}$ is of the form $Q_X \times \hat{Q}_{X|Y}$ for all iterations.

First, for $G_1 \cap G_2$, we need to solve

$$\min_{\hat{Q}_{X|Y} : B(\hat{Q}_{X|Y}) + D(Q_X||P_X) = E_\epsilon} D(Q_X \times \hat{Q}_{X|Y}||Q_X \times \hat{Q}_{X|Y}).$$

Standard Lagrangian optimization implies that for any $x \in \mathcal{X}$ such that $Q_X(x) \not= 0$ we must have

$$\hat{Q}^{\star}_{X|Y}(x|x) = \psi_x \hat{Q}_{X|Y}^{*}(x|x) \exp [-\lambda \cdot d_{P_{Y|X}}(x, \bar{x})],$$

where $\psi_x$ is a normalization constant, such that

$$\sum_{x \in \mathcal{X}} \hat{Q}^{\star}_{X|Y}(x|x) = 1,$$

for any $x \in \text{supp}(Q_X)$, namely

$$\hat{Q}^{\star}_{X|Y}(x|x) = \text{Min}(\hat{Q}_{X|Y}^{*}(x|x), P_{Y|X}, \lambda).$$

The value of $\hat{Q}^{\star}_{X|Y}(x|x)$ for $x \in \mathcal{X}$ with $Q_X(x) = 0$ is immaterial as it does not affect the optimal value of the objective function. Also, it is evident that the solution $\hat{Q}^{\star}_{X|Y}$ is indeed positive. Finally, we need to find $\hat{\lambda} \in \mathbb{R}$ such that the constraint $B(Q_X \times \hat{Q}_{X|Y}^{*}) + D(Q_X||P_X) = E_\epsilon$ is satisfied, namely

$$E_\epsilon + D(Q_X||P_X) = \sum_{x \in \mathcal{X} \setminus \mathcal{X}_0} Q_X(x) \hat{Q}^{\star}_{X|Y}(x|x),$$

$$\times \exp (-\lambda \cdot d_{P_{Y|X}}(x, \bar{x})).$$

Second, the linear family $G_3$ induces a simple constraint on the $\bar{x}$-marginal of $\hat{Q}_{X|Y}$. In this case, the lumping property [17, Lemma 4.1 and Sec. 5.1] implies that the I-projection onto $G_2$ is given by $\hat{Q}^{\star}_{X|Y} = \hat{M}(\hat{Q}_{X|Y})$, which evidently satisfies $\hat{Q}^{\star}_{X|Y} \in G_2$.

It is easily seen that Algorithm 3 indeed implements the procedure described in this proof.

**APPENDIX B**

Consider the case of very weakly correlated sources, namely

$$P_y(x|y) = P_Y(y) \cdot (1 + \epsilon_{xy}),$$

[17] In channel coding, this is referred to as “very noisy channel” [42, Sec. 3.4].
where for all $x \in \mathcal{X}$ we have $\sum_{y \in \mathcal{Y}} \epsilon_{xy} = 0$ and $|\epsilon_{xy}| \ll 1$ for all $(x, y) \in (\mathcal{X}, \mathcal{Y})$. Consider again the minimization problem in (55)

$$
\min_{Q_{Y|X}: D(Q_{Y|X}||E_{xy}) \leq E_x, \forall x \in \mathcal{X}} \left\{ I(Q_X \times Q_{Y|X}) + D(Q_{Y|X}||P_{Y|X}|Q_X) \right\},
$$

(B.2)

which from [10, Lemma 10.8.1] is equivalent to

$$
\min_{Q_Y, \tilde{Q}_{Y|X}} \min_{Q_{Y|X}: D(Q_{Y|X}||P_{Y|X}) \leq E_y} \left\{ D(Q_{Y|X}||\tilde{Q}_Y|Q_X) + D \left( (Q_{Y|X}||P_{Y|X}|Q_X) \right) \right\}. \quad (B.3)
$$

Now, notice that if $X$ and $Y$ are independent, then the optimal solution is $\tilde{Q}_Y^* = Q_{Y|X} = P_Y$ for all $x \in \mathcal{X}$ and both divergences vanish. A continuity argument then implies that for the low dependence case, the two divergences at the optimal solution are close to 0. Therefore, we can use the following Euclidean approximation [17, Th. 4.1]: For two PMFs $P_X, Q_X$ such that $\text{supp}(P_X) = \mathcal{X}$ and $Q_X \approx P_X$ we have that

$$
D(Q_X||P_X) \approx \frac{1}{2} \sum_{x \in \mathcal{X}} (Q_X(x) - P_X(x))^2. \quad (B.4)
$$

Moreover, for another PMF $\tilde{P}_X$, if $\tilde{P}_X \approx P_X$ then

$$
D(Q_X||P_X) \approx \frac{1}{2} \sum_{x \in \mathcal{X}} \frac{(Q_X(x) - \tilde{P}_X(x))^2}{\tilde{P}_X(x)}, \quad (B.5)
$$

which also shows that $D(P_X||Q_X) \approx D(Q_X||P_X)$. Now, the objective function of the minimization problem can be approximated as

$$
D(Q_{Y|X}||\tilde{Q}_Y|Q_X) + D(Q_{Y|X}||P_{Y|X}|Q_X)
\approx \frac{1}{2} E_{\tilde{Q}_Y} \left[ \sum_{y \in \mathcal{Y}} \frac{(Q_{Y|X}(y|x) - \tilde{Q}_Y(y))^2}{\tilde{Q}_Y(y)} \right]
+ \sum_{y \in \mathcal{Y}} \frac{(Q_{Y|X}(y|x) - P_{Y|X}(y|x))^2}{P_{Y|X}(y|x)},
\approx \frac{1}{2} E_{\tilde{Q}_Y} \left[ \sum_{y \in \mathcal{Y}} \frac{(Q_{Y|X}(y|x) - \tilde{Q}_Y(y))^2}{\tilde{Q}_Y(y)} \right]
+ \sum_{y \in \mathcal{Y}} \frac{(Q_{Y|X}(y|x) - P_{Y|X}(y|x))^2}{P_{Y|X}(y|x)}, \quad (B.6)
$$

and similarly the constraint $Q_{Y|X} \in A_{hb}$ is approximated by

$$
\frac{1}{2} E_{Q_X} \left[ \sum_{y \in \mathcal{Y}} \frac{(Q_{Y|X}(y|x) - P_{Y|X}(y|x))^2}{P_{Y|X}(y|x)} \right] \leq E_x - D(Q_X||P_X). \quad (B.7)
$$

Standard Lagrangian optimization over $Q_{Y|X}$ for a given $\tilde{Q}_Y$ results

$$
Q_{Y|X}^* = a P_{Y|X} + (1-a) \tilde{Q}_Y \quad (B.8)
$$

for some $\frac{1}{2} \leq a \leq 1$, where $a$ is either chosen to satisfy the constraint or $a = \frac{1}{2}$. Now, for any given $a$ the resulting value of the optimization problem is

$$
\left[ \frac{a^2 + (a - 1)^2}{2} \right] \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \frac{P_{Y}(y|x) - \tilde{Q}_Y(y)^2}{P_{Y}(y)}, \quad (B.9)
$$

and the minimizer $\tilde{Q}_Y$ can be easily found to be

$$
\tilde{Q}_Y^*(y) = \sum_{x \in \mathcal{X}} Q_X(x) P_{Y|X}(y|x). \quad (B.10)
$$

Notice that the optimal solution $\tilde{Q}_Y^*$ does not depend on $a$. Thus, for a given $E_x \geq D(Q_X||P_X)$ the optimal value of $a$ is given by $a^* = \max(\bar{a}, \frac{1}{2})$ where $\bar{a}$ achieves equality in (B.7),

$$
\tilde{a} = 1 - \frac{E_x - D(Q_X||P_X)}{\frac{1}{2} \sum_{x \in \mathcal{X}} Q_X(x) \sum_{y \in \mathcal{Y}} (P_{Y}(y|x) - \tilde{Q}_Y^*(y))^2}, \quad (B.11)
$$

using (B.5). Then, in the case of very weakly correlated sources, the optimal rate function can be approximated by

$$
\rho_{hb}(Q_X, E_x) \approx E_x + H(Q_X) - D(Q_X||P_X) - \left[ a^{*2} + (a^{*} - 1)^2 \right] D(P_Y||\tilde{Q}_Y^*|Q_X), \quad (B.12)
$$

where $a^{*}$ is given analytically as a function of $E_x$. In addition, similar approximations for the unconstrained minimization problem (46) show that

$$
D(Q_{Y|X}||P_{Y|X}|Q_X) \approx \frac{1}{4} D(P_{Y|X}||\tilde{Q}_Y^*|Q_X). \quad (B.13)
$$

Thus, for $D(Q_X||P_X) \leq E_x \leq D(Q_X||P_X) + \frac{1}{4} D(P_{Y|X}||\tilde{Q}_Y^*|Q_X)$ we have $\tilde{a} \leq \frac{1}{2}$ and by substituting $\tilde{a}$ in (B.12) we obtain

$$
\rho_{hb}(Q_X, E_x) \approx H(Q_X) - E_x + D(Q_X||P_X) - D(P_{Y|X}||\tilde{Q}_Y^*|Q_X) + 2D(P_{Y|X}||\tilde{Q}_Y^*|Q_X)(E_x - D(Q_X||P_X)). \quad (B.14)
$$

For $\rho_{op}(Q_X, E_x)$ the analysis is similar, and in this case $a^{*} \in (0, 1)$, so we obtain the exact same expression as in (B.14), but this time it is valid for the entire range $D(Q_X||P_X) \leq E_x \leq D(Q_X \times P_{Y|X}|X,y)||P_{XY})$. Notice that $D(P_{Y|X}||\tilde{Q}_Y^*|Q_X)$ is the mutual information of the joint distribution $Q_X \times P_{Y|X}$ and thus is a measure of the independence between $X$ and $Y$. As $D(P_{Y|X}||\tilde{Q}_Y^*|Q_X) \to 0$ then $X$ and $Y$ become “more” independent, and then the rate function $\rho_{hb}(Q_X, E_x)$ is affine for almost the entire range $E_x \geq D(Q_X||P_X)$. Indeed, in this case, the main error event is associated with “bad binning”,

where $\mathcal{X}$ and $\mathcal{Y}$ are the sets of possible values for $X$ and $Y$, respectively.
i.e. at least two source blocks of the same type are mapped to the same bin by the random binning procedure.

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor, Jun Chen, for providing them the unpublished manuscript [6], and for his useful comments. Specifically, the proof of Theorem 5 follows from the proof of Theorem 2 which is a reproduction of [6, Th. 1]. Useful comments made by the anonymous referees are also acknowledged with thanks.

REFERENCES

[1] R. Ahlswede, “Coding hypergraphs: A new approach to multi-user source coding—Part II,” J. Combinat., vol. 5, no. 3, pp. 220–268, 1980.
[2] R. Ahlswede and G. Dueck, “Good codes can be produced by a few permutations,” IEEE Trans. Inf. Theory, vol. 28, no. 3, pp. 430–443, May 1982.
[3] S. Arimoto, “Computation of random code exponent functions,” IEEE Trans. Inf. Theory, vol. 22, no. 6, pp. 665–671, Nov. 1976.
[4] Cambridge University Press, 2004.
[5] J. Chen, D.-K. He, A. Jagmohan, and L. A. Lstras-Montaño, “On the linear codebook-level duality between Slepian–Wolf coding and channel coding,” IEEE Trans. Inf. Theory, vol. 55, no. 12, pp. 5575–5590, Dec. 2009.
[6] J. Chen, D.-K. He, A. Jagmohan, and L. A. Lstras-Montaño. On the Reliability Function of Variable-Rate Slepian–Wolf Coding. [Online]. Available: http://www.ece.uci.edu/~jchen/IT/variableSW.pdf.
[7] J. Chen, D.-K. He, A. Jagmohan, and L. A. Lastras-Montaño, “On the reliability function of variable-rate Slepian–Wolf coding,” in Proc. 45th Annu. Allerton Conf. Commun., Control, Comput., Sep. 2007, pp. 1–31.
[8] J. Chen, D.-K. He, and A. Jagmohan, “On the duality between Slepian–Wolf coding and channel coding under mismatched decoding,” IEEE Trans. Inf. Theory, vol. 55, no. 9, pp. 4006–4018, Sep. 2009.
[9] J. Chen, D.-K. He, A. Jagmohan, and L. A. Lastras-Montaño, “On universal variable-rate Slepian–Wolf coding,” in Proc. IEEE Int. Conf. Commun., May 2008, pp. 1426–1430.
[10] T. M. Cover and J. A. Thomas, Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing). New York, NY, USA: Wiley, 2006.
[11] I. Csiszár, “Linear codes for sources and source networks: Error exponents, universal coding,” IEEE Trans. Inf. Theory, vol. 28, no. 4, pp. 585–592, Jul. 1982.
[12] I. Csiszár, “On the error exponent of source-channel transmission with a distortion threshold,” IEEE Trans. Inf. Theory, vol. 28, no. 6, pp. 823–828, Nov. 1982.
[13] I. Csiszár and J. Körner, “Towards a general theory of source networks,” IEEE Trans. Inf. Theory, vol. 26, no. 2, pp. 155–165, Mar. 1980.
[14] I. Csiszár and J. Körner, “Graph decomposition: A new key to coding theorems,” IEEE Trans. Inf. Theory, vol. 27, no. 1, pp. 5–12, Jan. 1981.
[15] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[16] I. Csiszár, J. Körner, and K. Marton, “A new look at the error exponent of a discrete memoryless channel,” in Proc. Int. Symp. Inf. Theory, 1977, p. 107.
[17] I. Csiszár and P. C. Shields, Information Theory and Statistics: A Tutorial (Foundations and Trends in Communications and Information Theory). Norwell, MA, USA: Now Publishers Inc., 2004.
[18] I. Csiszár and G. Tusnády, “Information geometry and alternating minimization procedures,” Statist. Decisions, vol. 1, no. 1, pp. 205–237, 1984.
[19] R. G. Gallager, Information Theory and Reliable Communication. New York, NY, USA: Wiley, 1968.
[20] R. Gallager, “The random coding bound is tight for the average code (Corresp.),” IEEE Trans. Inf. Theory, vol. 19, no. 2, pp. 244–246, Mar. 1973.
[21] R. G. Gallager, (1976), Source Coding With Side Information and Universal Coding, LDS-P—P—937, M.I.T. [Online]. Available: http://web.mit.edu/gallager/www/papers/paper5.pdf
[22] D.-K. He, L. A. Lastras-Montaño, E.-H. Yang, A. Jagmohan, and J. Chen, “On the redundancy of Slepian–Wolf coding,” IEEE Trans. Inf. Theory, vol. 55, no. 12, pp. 5607–5627, Dec. 2009.
[23] P. A. Humblet, “Generalization of Huffman coding to minimize the probability of buffer overflow (Corresp.),” IEEE Trans. Inf. Theory, vol. 27, no. 2, pp. 230–232, Mar. 1981.
[24] F. Jelinek, “Buffer overflow in variable length coding of fixed rate sources,” IEEE Trans. Inf. Theory, vol. 14, no. 3, pp. 490–501, May 1968.
[25] F. Jelinek, “Evaluation of expurgated bound exponents,” IEEE Trans. Inf. Theory, vol. 14, no. 3, pp. 501–505, May 1968.
[26] B. G. Kelly and A. B. Wagner, “Improved source coding exponents via Witsenhausen’s rate,” IEEE Trans. Inf. Theory, vol. 57, no. 9, pp. 5615–5633, Sep. 2011.
[27] S. Kuzuoka, “On the redundancy of variable-rate Slepian–Wolf coding,” in Proc. Int. Symp. Inf. Theory Appl. (ISITA), Oct. 2012, pp. 155–159.
[28] S. Kuzuoka and S. Watanabe, “An information-spectrum approach to weak variable-length Slepian–Wolf coding,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Jun./Jul. 2014, pp. 2684–2688.
[29] J. R. Lesh, “Computational algorithms for coding bound exponents,” Ph.D. dissertation, Dept. Elect. Eng., Univ. California, Los Angeles, CA, USA, 1976.
[30] K. Marton, “Error exponent for source coding with a fidelity criterion,” IEEE Trans. Inf. Theory, vol. 20, no. 2, pp. 197–199, Mar. 1974.
[31] N. Merhav, “On optimum strategies for minimizing the exponential moments of a loss function,” Commun. Inf. Syst., vol. 11, no. 4, pp. 343–368, 2011.
[32] N. Merhav, “On optimum parameter modulation–estimation from a large deviations perspective,” IEEE Trans. Inf. Theory, vol. 58, no. 12, pp. 7215–7225, Dec. 2012.
[33] N. Merhav, “Erasure/list exponents for Slepian–Wolf decoding,” IEEE Trans. Inf. Theory, vol. 60, no. 8, pp. 4463–4471, Aug. 2014. [Online]. Available: http://arxiv.org/pdf/1305.5626.pdf
[34] N. Merhav, “On zero-rate error exponents of finite-state channels with input-dependent states,” IEEE Trans. Inf. Theory, vol. 61, no. 2, pp. 741–750, Feb. 2015. [Online]. Available: http://arxiv.org/pdf/1406.7092.pdf
[35] B. Nakiboglu and L. Zheng, “Errors- and-erasures decoding for block codes with feedback,” IEEE Trans. Inf. Theory, vol. 58, no. 1, pp. 24–49, Jan. 2012.
[36] Y. Oohama and T. Han, “Universal coding for the Slepian–Wolf data compression system and the strong converse theorem,” IEEE Trans. Inf. Theory, vol. 40, no. 6, pp. 1908–1919, Nov. 1994.
[37] F. Rezaei and C. D. Charalambous, “Robust coding for uncertain sources: A minimax approach,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Sep. 2005, pp. 1539–1543.
[38] S. Sarvotham, D. Baron, and R. G. Baraniuk, “Variable-rate universal Slepian–Wolf coding with feedback,” in Proc. 39th Asilomar Conf. Signals, Syst. Comput. Rec., Oct./Nov. 2005, pp. 8–12.
[39] M. Sion, “On general minimax theorems,” Pacific J. Math., vol. 8, no. 1, pp. 171–176, 1958.
[40] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” IEEE Trans. Inf. Theory, vol. 19, no. 4, pp. 471–480, Jul. 1973.
[41] I. E. Telatar, “Exponential bounds for list size moments and error probability,” in Proc. IEEE Inf. Theory Workshop (ITW), Jun. 1998, p. 60.
[42] A. J. Viterbi and J. K. Omura, Principles of Digital Communication and Coding. New York, NY, USA: Dover, 2009.
[43] N. Weinberger and N. Merhav. (Nov. 2014). “Optimum trade-offs between the error exponent and the excess-rate exponent of variable-rate Slepian–Wolf coding.” [Online]. Available: http://arxiv.org/pdf/1401.0892v3.pdf

Nir Weinberger (S’14) received the B.Sc. and M.Sc. degrees (both summa cum laude) from Tel-Aviv University, Tel-Aviv, Israel, in 2006 and 2009, respectively. From 2006 to 2013 he served as an algorithm Engineer in the Israeli Defense Forces. Currently, he is pursuing his Ph.D. degree at the Technion - Israel Institute of Technology, Haifa, Israel. His research interest is information theory, with emphasis on large deviations aspects in coding problems.
Neri Merhav (S’86–M’87–SM’93–F’99) was born in Haifa, Israel, on March 16, 1957. He received the B.Sc., M.Sc., and D.Sc. degrees from the Technion, Israel Institute of Technology, in 1982, 1985, and 1988, respectively, all in electrical engineering.

From 1988 to 1990 he was with AT&T Bell Laboratories, Murray Hill, NJ, USA. Since 1990 he has been with the Electrical Engineering Department of the Technion, where he is now the Irving Shepard Professor. During 1994-2000 he was also serving as a consultant to the Hewlett-Packard Laboratories - Israel (HPL-I). His research interests include information theory, statistical communications, and statistical signal processing. He is especially interested in the areas of lossless/lossy source coding and prediction/filtering, relationships between information theory and statistics, detection, estimation, as well as in the area of Shannon Theory, including topics in joint source-channel coding, source/channel simulation, and coding with side information with applications to information hiding and watermarking systems. Another recent research interest concerns the relationships between Information Theory and statistical physics.

Dr. Merhav was a co-recipient of the 1993 Paper Award of the IEEE Information Theory Society and he is a Fellow of the IEEE since 1999. He also received the 1994 American Technion Society Award for Academic Excellence and the 2002 Technion Henry Taub Prize for Excellence in Research. From 1996 until 1999 he served as an Associate Editor for Source Coding to the IEEE TRANSACTIONS ON INFORMATION THEORY. He also served as a co-chairman of the Program Committee of the 2001 IEEE International Symposium on Information Theory. He is currently on the Editorial Board of Foundations and Trends in Communications and Information Theory.