On the Duality Between Slepian–Wolf Coding and Channel Coding Under Mismatched Decoding

Jun Chen, *Member, IEEE*, Da-ke He, and Ashish Jagmohan

***Abstract*—In this paper, Slepian–Wolf coding with a mismatched decoding metric is studied. Two different dualities between Slepian–Wolf coding and channel coding under mismatched decoding are established. These two dualities provide a systematic framework for comparing linear Slepian–Wolf codes, nonlinear Slepian–Wolf codes, and variable-rate Slepian–Wolf codes. In contrast with the fact that linear codes suffice to achieve the Slepian–Wolf limit under matched decoding, the minimum rate achievable with nonlinear Slepian–Wolf codes under mismatched decoding can be strictly lower than that achievable with linear Slepian–Wolf codes.**

***Index Terms*—Belief propagation, channel coding, duality, mismatched decoding, Slepian–Wolf coding.**

I. Introduction

CONSIDER a joint stationary and memoryless process $\{(X_i, Y_i)\}$ over finite alphabet $\mathcal{X} \times \mathcal{Y}$ with zero-order probability distribution P_{XY}. Let P_X and P_Y be the marginal distributions induced by P_{XY}. Without loss of generality, we assume $P_X(x) > 0$ for all $x \in \mathcal{X}$.

In its most basic form, Slepian–Wolf coding refers to the following problem. The encoder compresses $X^n = (X_1, X_2, \ldots, X_n)$ and sends the compressed data to the decoder so that the decoder, given the side information $Y^n = (Y_1, Y_2, \ldots, Y_n)$, can recover X^n with asymptotically zero error probability as n goes to infinity. Let $f_n : \mathcal{X}^n \to \mathcal{M}_n$ be the encoding function. It is clear that given $m \in \mathcal{M}_n$ and $y^n \in \mathcal{Y}^n$, the optimal decoding rule, also known as the matched decoding rule, is

$$\arg \min_{x^n \in \mathcal{X}^n, f_n(x^n)=m} -\sum_{i=1}^{n} \log P_{XY}(x_i, y_i).$$

However, in practice, P_{XY} may not be known perfectly. Therefore, it is of considerable interest to study the performance of Slepian–Wolf coding with a mismatched decoder (i.e., a probability distribution Q_{XY}, different from P_{XY}, is used for decoding). In this case, the decoding rule becomes

$$\arg \min_{x^n \in \mathcal{X}^n, f_n(x^n)=m} -\sum_{i=1}^{n} \log Q_{XY}(x_i, y_i). \tag{1}$$

A decoding error is declared if the minimizer is not the correct x^n or the minimizer is not unique. We will refer to Q_{XY} as the decoding metric.[1] We assume that $Q_{XY}(x, y) = 0$ only if $P_{XY}(x, y) = 0$ for $x \in \mathcal{X}, y \in \mathcal{Y}$.

An interesting fact about Slepian–Wolf coding under matched decoding is that the minimum rates achievable with different types of codes (variable rate or fixed rate, nonlinear or linear) are the same. Therefore, it is of considerable interest to examine whether this fact continues to hold in the mismatched decoding scenario. We will show that under mismatched decoding, the minimum rate achievable with variable-rate Slepian–Wolf codes is the same as that achievable with fixed-rate Slepian–Wolf codes while the minimum rate achievable with nonlinear Slepian–Wolf codes can be strictly lower than that achievable with linear Slepian–Wolf codes. Therefore, linear Slepian–Wolf codes might not be as efficient and robust as nonlinear Slepian–Wolf codes under mismatched decoding.

It is well known that under matched decoding, Slepian–Wolf coding for source X with decoder side information Y is related to channel coding for channel $P_{Y|X}$ (i.e., the channel from X to Y induced by the joint distribution P_{XY}). This connection was already noticed in the seminal work by Slepian and Wolf [1] and was later formalized by Csiszár and Körner [2] and by Ahlswede and Dueck [3].

Wolf pointed out in [4] that their work [1] was inspired by the following simple example, which will be referred to as the binary source example. Consider the case where $X = Y \oplus Z$, where X, Y, and Z are binary, Z is independent of Y, and $\oplus$ is the modulo-2 addition. Assume that Z has a Bernoulli distribution with parameter $p \in (0, 0.5)$; therefore, the channel from Y to X (i.e., $P_{X|Y}$) is a binary symmetric channel with crossover probability p. Note that here X and Y do not need to be uniformly distributed. Slepian and Wolf observed that for this example the cosets induced by a good linear code for binary symmetric channel with crossover probability p yield a good Slepian–Wolf code for source X with decoder side information Y. This example was later popularized by Wyner [5] (see also [6] and [7]).

A careful reader might have noticed that in the binary source example the linear code is designed for channel $P_{X|Y}$, not

Manuscript received January 08, 2008; revised April 15, 2009. Current version published August 19, 2009.

J. Chen is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1 Canada (e-mail: junchen@ece.mcmaster.ca).

D.-k. He was with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA. He is now with SlipStream Data, Waterloo, ON N2L 5Z5 Canada (e-mail: dhe@rim.com).

A. Jagmohan is with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: ashishja@us.ibm.com).

Communicated by I. Kontoyiannis, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2009.2025527

[1]Through this paper, the term "metric" is used in a broad sense as any non-negative-valued function on $\mathcal{X} \times \mathcal{Y}$.

$P_{Y|X}$. Indeed, $P_{Y|X}$ is not a binary symmetric channel if Y (and consequently, X) is not uniformly distributed; furthermore, in this example, a good linear code for channel $P_{Y|X}$, in general, does not yield a good linear Slepian–Wolf code for source X with decoder side information Y. Therefore, it seems that there is an inconsistency with the aforementioned connection between Slepian–Wolf coding and channel coding. However, it should be noted that the connection between Slepian–Wolf coding for source X with decoder side information Y and channel coding for channel $P_{Y|X}$ is established [2], [3] using nonlinear codes. Therefore, there is no essential inconsistency; rather, it suggests that the connection between Slepian–Wolf coding and channel coding is quite subtle, depending on the type of codes used.

We will clarify this subtle issue in the general setting of mismatched decoding. To this end, it is worthwhile to first have a review of previous work on channel coding with a mismatched decoding metric. Notable results in this area include various lower bounds on the mismatch capacity, such as the Csiszár–Körner–Hui lower bound [8], [9] and the generalized mutual information [10]. In a remarkable paper [11], Balakirsky proved that the Csiszár–Körner–Hui lower bound is tight for binary-input channels. However, a computable characterization of the mismatch capacity is still unknown in general. Indeed, it was shown in [12] that a complete solution to this problem would resolve the long-standing problem of computing the zero-error capacity.

It will be seen that Slepian–Wolf coding under mismatched decoding is closely related to its counterpart in channel coding. However, the connection between these two coding problems depends critically on the type of codes used. Specifically, we will establish two different dualities between Slepian–Wolf coding and channel coding under mismatched decoding, namely, the type-level duality and the linear codebook-level duality. The type-level duality is relevant for variable-rate codes and fixed-rate nonlinear codes while the linear codebook-level duality is tailored for linear codes. The type-level duality strengthens and extends the duality results in [2] and [3]. The linear codebook-level duality generalizes the insight obtained from the binary source example by Slepian and Wolf as well as the related results in the case where the side information Y^n is absent [13], [14]. These two dualities together provide a systematic framework for comparing the performance of linear Slepian–Wolf codes, nonlinear Slepian–Wolf codes, and variable-rate Slepian–Wolf codes.

Although we mainly focus on obtaining a fundamental understanding of Slepian–Wolf coding under mismatched decoding by leveraging its duality with channel coding, our results also provide a new perspective on the existing results in channel coding. For example, 1) we show that the Csiszár–Körner–Hui lower bound and the generalized mutual information for a class of channels and decoding metrics can be naturally interpreted as the achievable rate of linear codes; and 2) a result by Csiszár and Narayan [12] regarding the equivalence between the positivity of the Csiszár–Körner–Hui lower bound and that of the mismatch capacity finds an interesting application to linear Slepian–Wolf codes.

The rest of the paper is divided into four sections. We introduce a few basic definitions in Section II. In Section III, we establish the type-level duality between Slepian–Wolf coding and channel coding, which is then leveraged to characterize the minimum rate achievable with fixed-rate Slepian–Wolf codes and variable-rate Slepian–Wolf codes under mismatched decoding. The linear codebook-level duality is established in Section IV. It is revealed that the minimum rate achievable with nonlinear Slepian–Wolf codes under mismatched decoding can be strictly lower than that achievable with linear Slepian–Wolf codes. It is also shown that the linear codebook-level duality continues to hold under mismatched belief propagation (BP) decoding. We conclude the paper in Section V. Throughout this paper, the logarithm function is to the base two unless specified otherwise.

II. Definitions

A fixed-rate Slepian–Wolf code $\phi_n(\cdot)$ is a mapping from $\mathcal{X}^n$ to a set $\mathcal{A}_n$. The rate of $\phi_n(\cdot)$ is defined as

$$R(\phi_n) = \frac{1}{n}\log|\mathcal{A}_n|.$$

Let $P_e(\phi_n, P_{XY}, Q_{XY})$ denote the decoding error probability of Slepian–Wolf code $\phi_n(\cdot)$ under decoding metric Q_{XY}. A variable-rate Slepian–Wolf code $\varphi_n(\cdot)$ is a mapping from $\mathcal{X}^n$ to a binary prefix code $\mathcal{B}_n$. Let $l(\varphi_n(x^n))$ denote the length of binary string $\varphi_n(x^n)$. The rate of variable-rate Slepian–Wolf code $\varphi_n(\cdot)$ is defined as

$$R(\varphi_n) = \frac{1}{n}\mathbb{E}l(\varphi_n(X^n)).$$

The decoding error probability of variable-rate Slepian–Wolf code $\varphi_n(\cdot)$ under decoding metric Q_{XY} is denoted as $P_e(\varphi_n, P_{XY}, Q_{XY})$.

We say rate R is achievable with fixed-rate Slepian–Wolf codes under decoding metric Q_{XY} if for any $\epsilon > 0$, there exists a sequence of fixed-rate Slepian–Wolf codes $\{\phi_n\}$ such that

$$\limsup_{n\to\infty} R(\phi_n) \leq R + \epsilon$$
$$\liminf_{n\to\infty} P_e(\phi_n, P_{XY}, Q_{XY}) = 0.$$

The minimum rate achievable with fixed-rate Slepian–Wolf codes under decoding metric Q_{XY} is denoted as $H_{M,f}(P_{XY}, Q_{XY})$. The minimum rate achievable with variable-rate Slepian–Wolf codes under decoding metric Q_{XY}, denoted as $H_{M,v}(P_{XY}, Q_{XY})$, is similarly defined (with ϕ_n replaced by φ_n in the above definition). It should be obvious from the definition that

$$H_{M,f}(P_{XY}, Q_{XY}) \geq H_{M,v}(P_{XY}, Q_{XY}).$$

It turns out that a good way to understand Slepian–Wolf coding with a mismatched decoding metric is to study the connection with its counterpart in channel coding. The method of types will be needed to establish such a connection. To set up the necessary background, we will introduce a few basic definitions from [2] and [15]. Let $\mathcal{P}(\mathcal{X})$ denote the set of all probability distributions on $\mathcal{X}$. For any $P \in \mathcal{P}(\mathcal{X})$, define $H(P) = -\sum_{x\in\mathcal{X}} P(x)\log P(x)$. The type of a sequence $x^n \in \mathcal{X}^n$, denoted as P_{x^n}, is the empirical probability distribution of x^n. Let $\mathcal{P}_n(\mathcal{X})$ denote the set consisting of the possible types of sequences $x^n \in \mathcal{X}^n$. For any $P \in \mathcal{P}_n(\mathcal{X})$, the type class $\mathcal{T}_n(P)$ is the set of sequences (in $\mathcal{X}^n$) of type P. We will frequently use the elementary results listed below

$$|\mathcal{P}_n(\mathcal{X})| \le (n+1)^{|\mathcal{X}|}$$
$$\frac{1}{(n+1)^{|\mathcal{X}|}} 2^{nH(P)} \le |\mathcal{T}_n(P)| \le 2^{nH(P)}, \qquad P \in \mathcal{P}_n(\mathcal{X})$$
$$\prod_{i=1}^{n} P_X(x_i) = 2^{-n[D(P\,\|\,P_X)+H(P)]}, \qquad x^n \in \mathcal{T}_n(P), \qquad P \in \mathcal{P}_n(\mathcal{X}).$$

In channel coding, given channel $W_{Y\,|\,X} : \mathcal{X} \to \mathcal{Y}$, codebook $\mathcal{C}_n \subseteq \mathcal{X}^n$, and channel output $y^n \in \mathcal{Y}^n$, the optimal decoding rule is

$$\arg \min_{x^n \in \mathcal{C}_n} -\sum_{i=1}^{n} \log W_{Y\,|\,X}(y_i\,|\,x_i).$$

In channel coding, mismatched decoding refers to the scenario in which the decoder uses a decoding metric $V_{Y\,|\,X}$ that is different from $W_{Y\,|\,X}$. In this case, the decoding rule becomes

$$\arg \min_{x^n \in \mathcal{C}_n} -\sum_{i=1}^{n} \log V_{Y\,|\,X}(y_i\,|\,x_i). \tag{2}$$

A decoding error is declared if the minimizer is not the transmitted codeword or the minimizer is not unique. We will assume that $V_{Y\,|\,X}(y\,|\,x) = 0$ only if $W_{Y\,|\,X}(y\,|\,x) = 0$ for $x \in \mathcal{X}, y \in \mathcal{Y}$. The rate of block channel code $\mathcal{C}_n$ is defined as

$$R(\mathcal{C}_n) = \frac{1}{n} \log |\mathcal{C}_n|.$$

Let $P_e(\mathcal{C}_n, W_{Y\,|\,X}, V_{Y\,|\,X})$ denote the maximum decoding error probability of block channel code $\mathcal{C}_n$ under decoding metric $V_{Y\,|\,X}$. We say rate R is achievable with block channel codes under decoding metric $V_{Y\,|\,X}$ if for any $\epsilon > 0$, there exists a sequence of block channel codes $\{\mathcal{C}_n\}$ such that

$$\liminf_{n\to\infty} R(\mathcal{C}_n) \ge R - \epsilon$$
$$\liminf_{n\to\infty} P_e(\mathcal{C}_n, W_{Y\,|\,X}, V_{Y\,|\,X}) = 0.$$

The maximum rate achievable with block channel codes under decoding metric $V_{Y\,|\,X}$, denoted as $C_M(W_{Y\,|\,X}, V_{Y\,|\,X})$, is referred to as the mismatch capacity.

For any $P \in \mathcal{P}(\mathcal{X})$, we say rate R is achievable with constant composition codes of type approximately P under decoding metric $V_{Y\,|\,X}$ if for any $\epsilon > 0$, there exists a sequence of channel codes $\{\mathcal{C}_n\}$ with $\mathcal{C}_n \subseteq \mathcal{T}_n(P_n)$ for some $P_n \in \mathcal{P}_n(\mathcal{X})$ such that

$$\lim_{n\to\infty} \|P_n - P\| = 0$$
$$\liminf_{n\to\infty} R(\mathcal{C}_n) \ge R - \epsilon$$
$$\liminf_{n\to\infty} P_e(\mathcal{C}_n, W_{Y\,|\,X}, V_{Y\,|\,X}) = 0$$

where $\|\cdot\|$ is the L_1 norm. The maximum rate achievable with constant composition codes of type approximately P under decoding metric $V_{Y\,|\,X}$ is denoted as $C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X})$.

Proposition 1: $C_M(W_{Y\,|\,X}, V_{Y\,|\,X}) = \sup_{P\in\mathcal{P}(\mathcal{X})} C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X})$.

Remark: This result is a simple consequence of the well-known fact that a block code can be reduced to a constant composition code with negligible loss of rate when the block length is long enough.

Proposition 2: The function $C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X})$ is continuous with respect to P in the sense that for any $P \in \mathcal{P}(\mathcal{X})$ and $\epsilon > 0$, there exists a $\delta(P,\epsilon) > 0$ such that $|C_M(P', W_{Y\,|\,X}, V_{Y\,|\,X}) - C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X})| \le \epsilon$ for all $P' \in \mathcal{P}(\mathcal{X})$ with $D(P'\,\|\,P) \le \delta(P,\epsilon)$.

Proof: Define $\mathcal{X}_+ = \{x \in \mathcal{X} : P(x) > 0\}$. For any $P \in \mathcal{X}$ and $\epsilon > 0$, we can find a $\delta(P,\epsilon) > 0$ such that

$$\sum_{x\in\mathcal{X}_+} P'(x) = 1$$
$$\max\left(\max_{x\in\mathcal{X}_+} \frac{P'(x)}{P(x)}, \max_{x\in\mathcal{X}_+} \frac{P(x)}{P'(x)}\right) \le 1+\epsilon$$

for all $P' \in \mathcal{P}(\mathcal{X})$ with $D(P'\,\|\,P) \le \delta(P,\epsilon)$.

Let $P' \in \mathcal{P}(\mathcal{X})$ be an arbitrary probability distribution satisfying $D(P'\,\|\,P) \le \delta(P,\epsilon)$. By the definition of $C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X})$, there exists a sequence of channel codes $\{\mathcal{C}_n\}$ with $\mathcal{C}_n \subseteq \mathcal{T}_n(P_n)$ for some $P_n \in \mathcal{P}_n(\mathcal{X})$ such that

$$\lim_{n\to\infty} \|P_n - P\| = 0$$
$$\liminf_{n\to\infty} R(\mathcal{C}_n) \ge C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X}) - \epsilon$$
$$\liminf_{n\to\infty} P_e(\mathcal{C}_n, W_{Y\,|\,X}, V_{Y\,|\,X}) = 0.$$

Given each $\mathcal{C}_n$, we can construct a constant composition code $\mathcal{C}'_{m_n}$ of length $m_n = \lceil (1+\epsilon)n \rceil$ and type P'_{m_n} for some $P'_{m_n} \in \mathcal{P}_{m_n}(\mathcal{X})$ by appending a fixed string in $\mathcal{X}_+^{m_n - n}$ to each codeword in $\mathcal{C}_n$ such that

$$\lim_{n\to\infty} \|P'_{m_n} - P'\| = 0.$$

It is easy to see that $P_e(\mathcal{C}'_{m_n}, W_{Y\,|\,X}, V_{Y\,|\,X}) = P_e(\mathcal{C}_n, W_{Y\,|\,X}, V_{Y\,|\,X})$ for all n. Moreover

$$\liminf_{n\to\infty} R(\mathcal{C}'_{m_n}) = \liminf_{n\to\infty} \frac{n}{m_n} R(\mathcal{C}_n) \ge \frac{C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X}) - \epsilon}{1+\epsilon}.$$

Now by the definition of $C_M(P', W_{Y\,|\,X}, V_{Y\,|\,X})$, we must have

$$C_M(P', W_{Y\,|\,X}, V_{Y\,|\,X}) \ge \frac{C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X}) - \epsilon}{1+\epsilon}.$$

It follows by symmetry that

$$C_M(P, W_{Y\,|\,X}, V_{Y\,|\,X}) \ge \frac{C_M(P', W_{Y\,|\,X}, V_{Y\,|\,X}) - \epsilon}{1+\epsilon}.$$

Since $\epsilon > 0$ is arbitrary, the proof is complete. □

III. Type-Level Duality

The key result of this section is Theorem 1, which reveals an intimate formula-level connection between Slepian–Wolf coding and channel coding under mismatched decoding. It will be clear from the proof that the formula-level connection can be viewed as a manifestation of the more fundamental type-level duality between Slepian–Wolf coding and channel coding.

Let $P_{Y|X}$ and $Q_{Y|X}$ be the conditional probability distributions of Y given X induced by P_{XY} and Q_{XY}, respectively. Roughly speaking, the type-level duality stands for the following connection between Slepian–Wolf coding and channel coding.

1) Given a sequence of constant composition codes of rate approximately R and type approximately P_X with diminishing maximum decoding error probability for channel $P_{Y|X}$ under decoding metric $Q_{Y|X}$, one can construct a sequence of Slepian–Wolf codes of rate approximately $H(P_X) - R$ with diminishing decoding error probability for source distribution P_{XY} under decoding metric Q_{XY}.
2) Given a sequence of Slepian–Wolf codes of rate approximately R with diminishing decoding error probability for source distribution P_{XY} under decoding metric Q_{XY}, one can lift out a sequence of constant composition codes of rate approximately $H(P_X) - R$ and type approximately P_X with diminishing maximum decoding error probability for channel $P_{Y|X}$ under decoding metric $Q_{Y|X}$.

Under matched decoding, the duality between Slepian–Wolf coding and channel coding is well known (e.g., see [2, Th. 1.2, p. 238], [2, Problem 2, p. 262], and [3, Th. 1]). In particular, a strong connection between Slepian–Wolf coding and channel coding was established in [3] using the covering lemma [16]. Here, we will also make use of the covering lemma to establish the type-level duality between Slepian–Wolf coding and channel coding in the context of mismatched decoding. However, our proof has several noteworthy differences from that in [3] even when specialized to the matched decoding scenario.

1) For the direct part (i.e., from channel code to Slepian–Wolf code) in [3], the Slepian–Wolf code is constructed using a polynomial number of channel codes (one for each type class) that have no explicit relationship. In contrast, the channel codes we use are lengthened versions of a single constant composition code.
2) The converse part (i.e., from Slepian–Wolf code to channel code) in [3] is for fixed-rate Slepian–Wolf codes while our converse part holds for variable-rate Slepian–Wolf codes (and consequently, holds for fixed-rate Slepian–Wolf codes as well).

Theorem 1: $H_{M,f}(P_{XY}, Q_{XY}) = H_{M,v}(P_{XY}, Q_{XY}) = H(P_X) - C_M(P_X, P_{Y|X}, Q_{Y|X})$.

Proof: We will first show that $H_{M,f}(P_{XY}, Q_{XY}) \leq H(P_X) - C_M(P_X, P_{Y|X}, Q_{Y|X})$. By the definition of $C_M(P_X, P_{Y|X}, Q_{Y|X})$, for any $\epsilon > 0$, there exists a sequence of channel codes $\{\mathcal{C}_n\}$ with $\mathcal{C}_n \subseteq \mathcal{T}_n(P_n)$ for some $P_n \in \mathcal{P}_n(\mathcal{X})$ such that

$$\lim_{n\to\infty} \|P_n - P_X\| = 0$$
$$\liminf_{n\to\infty} R(\mathcal{C}_n) \geq C(P_X, P_{Y|X}, Q_{Y|X}) - \epsilon$$
$$\liminf_{n\to\infty} P_e(\mathcal{C}_n, P_{Y|X}, Q_{Y|X}) = 0.$$

Define

$$\mathcal{E} = \left\{ P \in \mathcal{P}(\mathcal{X}) : \max_{x\in\mathcal{X}} \frac{P_X(x)}{P(x)} \leq 1+\epsilon, H(P) \leq H(P_X)+\epsilon \right\}.$$

Since $\lim_{n\to\infty} \|P_n - P_X\| = 0$ and $P_X(x) > 0$ for all $x \in \mathcal{X}$, there exists an $N(\epsilon)$ such that

$$\max_{x\in\mathcal{X}} \frac{P_n(x)}{P_X(x)} \leq 1 + \epsilon$$

for all $n \geq N(\epsilon)$, which further implies

$$\max_{P\in\mathcal{P}_n(\mathcal{X})\cap\mathcal{E}} \max_{x\in\mathcal{X}} \frac{P_n(x)}{P(x)} \leq (1+\epsilon)^2$$

for all $n \geq N(\epsilon)$. Let $k_n = \lceil (1+\epsilon)^2 n \rceil$. For any $n \geq N(\epsilon)$ and $P \in \mathcal{P}_{k_n}(\mathcal{X}) \cap \mathcal{E}$, we can construct a constant composition code $\mathcal{C}'_{k_n}(P)$ of length k_n and type P by appending a fixed string in $\mathcal{X}^{k_n - n}$ to each codeword in $\mathcal{C}_n$. Since the appended fixed string does not affect the decoding rule (2), it follows that

$$P_e(\mathcal{C}'_{k_n}(P), P_{Y|X}, Q_{Y|X}) = P_e(\mathcal{C}_n, P_{Y|X}, Q_{Y|X}) \quad (3)$$

for all $P \in \mathcal{P}_{k_n}(\mathcal{X}) \cap \mathcal{E}$. By the covering lemma [16], for each $P \in \mathcal{P}_{k_n}(\mathcal{X}) \cap \mathcal{E}$, there exist L permutations $\pi_1, \ldots, \pi_L$ of the integers $1, \ldots, k_n$ such that

$$\bigcup_{i=1}^{L} \pi_i(\mathcal{C}'_{k_n}(P)) = \mathcal{T}_{k_n}(P)$$

where

$$\begin{aligned} L &\triangleq \max_{P\in\mathcal{P}_{k_n}(\mathcal{X})\cap\mathcal{E}} \lfloor |\mathcal{C}'_{k_n}(P)|^{-1} |\mathcal{T}_{k_n}(P)| \log |\mathcal{T}_{k_n}(P)| + 1 \rfloor \\ &= \max_{P\in\mathcal{P}_{k_n}(\mathcal{X})\cap\mathcal{E}} \lfloor |\mathcal{C}_n|^{-1} |\mathcal{T}_{k_n}(P)| \log |\mathcal{T}_{k_n}(P)| + 1 \rfloor. \end{aligned}$$

Note that

$$P_e(\pi_i(\mathcal{C}'_{k_n}(P)), P_{Y|X}, Q_{Y|X}) = P_e(\mathcal{C}'_{k_n}(P), P_{Y|X}, Q_{Y|X}) \quad (4)$$

for all i, which is due to the fact that the decoding rule (2) is invariant under permutation. Intuitively, we can view $\pi_1(\mathcal{C}'_{k_n}(P)), \ldots, \pi_L(\mathcal{C}'_{k_n}(P))$ as a partition[2] of $\mathcal{T}_{k_n}(P)$. Now construct fixed-rate Slepian–Wolf code $\phi_{k_n}(\cdot) : \mathcal{X}^{k_n} \to \mathcal{P}_{k_n}(\mathcal{X}) \times \{1, 2, \ldots, L\}$ as follows.

1) The encoder sends the type of x^{k_n} (i.e., $P_{x^{k_n}}$) to the decoder.
2) If $P_{x^{k_n}} \notin \mathcal{E}$, the encoder sends an arbitrary symbol (say, 1) to the decoder.
3) If $x^{k_n} \in \mathcal{T}_{k_n}(P)$ for some $P \in \mathcal{E}$, the encoder finds the set $\pi_{i^*}(\mathcal{C}'_{k_n}(P))$ that contains x^{k_n} and sends the index i^* to the decoder.

Since the decoder knows type of x^{k_n}, the decoding rule (1) depends on Q_{XY} only through $Q_{Y|X}$. Now it is easy to see that conditioned on $X^{k_n} \in \pi_i(\mathcal{C}'_{k_n}(P))$, the decoding error probability is bounded from above by $P_e(\pi_i(\mathcal{C}'_{k_n}(P)), P_{Y|X}, Q_{Y|X})$. Therefore, in view of (3) and (4), we have

$$\begin{aligned} P_e(\phi_{k_n}, P_{XY}, Q_{XY}) \leq &\sum_{P\in\mathcal{P}_{k_n}(\mathcal{X}), P\notin\mathcal{E}} Pr\{X^{k_n} \in \mathcal{T}_{k_n}(P)\} \\ &+ P_e(\mathcal{C}_n, P_{Y|X}, Q_{Y|X}). \end{aligned}$$

[2]The overlapped part between $\pi_i(\mathcal{C}'_{k_n}(P))$ and $\pi_j(\mathcal{C}'_{k_n}(P))$ for different i and j can be split in an arbitrary manner.

Now it can be readily shown by invoking the weak law of large numbers that

$$\liminf_{n\to\infty} P_e(\phi_{k_n}, P_{XY}, Q_{XY}) = 0.$$

Since $|\mathcal{P}_{k_n}(\mathcal{X})| \le (k_n+1)^{|\mathcal{X}|}$ and $|\mathcal{T}_{k_n}(P)| \le 2^{k_n H(P)}$ for $P \in \mathcal{P}_{k_n}(\mathcal{X})$, it follows that

$$\begin{aligned}\limsup_{n\to\infty} & R(\phi_{k_n}) \\ &\le \limsup_{k_n\to\infty} \frac{1}{k_n}\log[(k_n+1)^{|\mathcal{X}|}L] \\ &\le \max_{P\in\mathcal{E}} H(P) - \frac{1}{(1+\epsilon)^2}\liminf_{n\to\infty} R(\mathcal{C}_n) \\ &\le H(P_X) + \epsilon - \frac{C_M(P_X, P_{Y|X}, Q_{Y|X}) - \epsilon}{(1+\epsilon)^2}.\end{aligned}$$

By the definition of $H_{M,f}(P_{XY}, Q_{XY})$, we must have

$$H_{M,f}(P_{XY}, Q_{XY}) \le H(P_X) + \epsilon - \frac{C_M(P_X, P_{Y|X}, Q_{Y|X}) - \epsilon}{(1+\epsilon)^2}.$$

Since ϵ can be arbitrarily close to zero, we obtain the desired result.

Now we proceed to show that $H_{M,v}(P_{XY}, Q_{XY}) \ge H(P_X) - C_M(P_X, P_{Y|X}, Q_{Y|X})$. By the definition of $H_{M,v}(P_{XY}, Q_{XY})$, for any $\epsilon > 0$, there exists a sequence of variable-rate Slepian–Wolf codes $\{\varphi_n\}$ such that

$$\limsup_{n\to\infty} R(\varphi_n) \le H_{M,v}(P_{XY}, Q_{XY}) + \epsilon \quad (5)$$

$$\liminf_{n\to\infty} P_e(\varphi_n, P_{XY}, Q_{XY}) = 0. \quad (6)$$

For each $P \in \mathcal{P}_n(\mathcal{X})$, suppose $\varphi_n(\cdot)$ partitions $\mathcal{T}_n(P)$ into $N_n(P)$ disjoint subsets $\mathcal{T}_n(P,1), \ldots, \mathcal{T}_n(P, N_n(P))$. Since $\varphi_n(\cdot)$ is a prefix code, it follows that conditioned on $X^n \in \mathcal{T}_n(P)$, the expected length of $\varphi_n(X^n)$ is bounded from below by the entropy of the partition of $\mathcal{T}_n(P)$ induced by $\varphi_n(\cdot)$, i.e.,

$$\begin{aligned}\mathbb{E}(l(\varphi_n(X^n)) \,|\, X^n \in \mathcal{T}_n(P)) \\ &\ge -\sum_{i=1}^{N_n(P)} Pr\{X^n \in \mathcal{T}_n(P,i) \,|\, X^n \in \mathcal{T}_n(P)\}\log \\ &\quad \times Pr\{X^n \in \mathcal{T}_n(P,i) \,|\, X^n \in \mathcal{T}_n(P)\} \\ &= \sum_{i=1}^{N_n(P)} \frac{|\mathcal{T}_n(P,i)|}{|\mathcal{T}_n(P)|}\log\frac{|\mathcal{T}_n(P)|}{|\mathcal{T}_n(P,i)|}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}R(\varphi_n) &= \frac{1}{n}\mathbb{E}(\mathbb{E}(l(\varphi_n(X^n)) \,|\, X^n \in \mathcal{T}_n(P))) \\ &= \frac{1}{n}\sum_{P\in\mathcal{P}_n(\mathcal{X})} Pr\{X^n \in \mathcal{T}_n(P)\} \\ &\quad \times \mathbb{E}(l(\varphi_n(X^n)) \,|\, X^n \in \mathcal{T}_n(P)) \\ &\ge \frac{1}{n}\sum_{P\in\mathcal{P}_n(\mathcal{X})}\sum_{i=1}^{N_n(P)} Pr\{X^n \in \mathcal{T}_n(P,i)\}\log\frac{|\mathcal{T}_n(P)|}{|\mathcal{T}_n(P,i)|}.\end{aligned}$$

It is also easy to see that

$$\begin{aligned}P_e(\varphi_n, P_{XY}, Q_{XY}) &= \sum_{P\in\mathcal{P}_n(\mathcal{X})}\sum_{i=1}^{N_n(P)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \\ &\quad \times Pr\{\text{error} \,|\, X^n \in \mathcal{T}_n(P,i)\}.\end{aligned}$$

Define

$$\begin{aligned}\mathcal{F}_n(\epsilon) &= \Big\{(P,i) : \frac{1}{n}\log\frac{|\mathcal{T}_n(P)|}{|\mathcal{T}_n(P,i)|} \le H_{M,v}(P_{XY}, Q_{XY}) + 2\epsilon, \\ &\qquad P \in \mathcal{P}_n(\mathcal{X}), i = 1, 2, \ldots, N_n(P)\Big\} \\ \mathcal{G}_n(\alpha) &= \Big\{(P,i) : Pr\{\text{error} \,|\, X^n \in \mathcal{T}_n(P,i)\} \\ &\qquad \le \alpha P_e(\varphi_n, P_{XY}, Q_{XY}), \\ &\qquad P \in \mathcal{P}_n(\mathcal{X}), i = 1, 2, \ldots, N_n(P)\Big\}, \\ \mathcal{S}_n(\epsilon) &= \Big\{P \in \mathcal{P}_n(\mathcal{X}) : H(P) \ge H(P_X) - \epsilon, \\ &\qquad \max_{x\in\mathcal{X}} \frac{P(x)}{P_X(x)} \le 1 + \epsilon\Big\}.\end{aligned}$$

Note that

$$\begin{aligned}R(\varphi_n) &\ge \frac{1}{n}\sum_{P\in\mathcal{P}_n(\mathcal{X})}\sum_{i=1}^{N_n(P)} Pr\{X^n \in \mathcal{T}_n(P,i)\}\log\frac{|\mathcal{T}_n(P)|}{|\mathcal{T}_n(P,i)|} \\ &= \frac{1}{n}\sum_{(P,i)\in\mathcal{F}_n(\epsilon)} Pr\{X^n \in \mathcal{T}_n(P,i)\}\log\frac{|\mathcal{T}_n(P)|}{|\mathcal{T}_n(P,i)|} \\ &\quad + \frac{1}{n}\sum_{(P,i)\in\mathcal{F}_n^c(\epsilon)} Pr\{X^n \in \mathcal{T}_n(P,i)\}\log\frac{|\mathcal{T}_n(P)|}{|\mathcal{T}_n(P,i)|} \\ &\ge \frac{1}{n}\sum_{(P,i)\in\mathcal{F}_n^c(\epsilon)} Pr\{X^n \in \mathcal{T}_n(P,i)\}\log\frac{|\mathcal{T}_n(P)|}{|\mathcal{T}_n(P,i)|} \\ &\ge [H_{M,v}(P_{XY}, Q_{XY}) + 2\epsilon] \\ &\quad \times \sum_{(P,i)\in\mathcal{F}_n^c(\epsilon)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \\ &= [H_{M,v}(P_{XY}, Q_{XY}) + 2\epsilon] \\ &\quad \times \left[1 - \sum_{(P,i)\in\mathcal{F}_n(\epsilon)} Pr\{X^n \in \mathcal{T}_n(P,i)\}\right]\end{aligned}$$

where $\mathcal{F}_n^c(\epsilon) = \{(P,i) \notin \mathcal{F}_n(\epsilon) : P \in \mathcal{P}_n, i = 1, 2, \ldots, N_n(P)\}$. In view of (5), we have

$$\begin{aligned}\liminf_{n\to\infty} \sum_{(P,i)\in\mathcal{F}_n(\epsilon)} & Pr\{X^n \in \mathcal{T}_n(P,i)\} \\ &\ge \frac{\epsilon}{H_{M,v}(P_{XY}, Q_{XY}) + 2\epsilon}. \quad (7)\end{aligned}$$

Similarly

$$\begin{aligned}P_e(\varphi_n, & P_{XY}, Q_{XY}) \\ &= \sum_{P\in\mathcal{P}_n(\mathcal{X})}\sum_{i=1}^{N_n(P)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \\ &\quad \times Pr\{\text{error} \,|\, X^n \in \mathcal{T}_n(P,i)\}\end{aligned}$$

$$
\begin{aligned}
&= \sum_{(P,i)\in\mathcal{G}_n(\alpha)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \\
&\quad \times Pr\{\text{error} \,|\, X^n \in \mathcal{T}_n(P,i)\} \\
&\quad + \sum_{(P,i)\in\mathcal{G}_n^c(\alpha)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \\
&\quad \times Pr\{\text{error} \,|\, X^n \in \mathcal{T}_n(P,i)\} \\
&\geq \sum_{(P,i)\in\mathcal{G}_n^c(\alpha)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \\
&\quad \times Pr\{\text{error} \,|\, X^n \in \mathcal{T}_n(P,i)\} \\
&\geq \alpha P_e(\varphi_n, P_{XY}, Q_{XY}) \\
&\quad \times \sum_{(P,i)\in\mathcal{G}_n^c(\alpha)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \\
&= \alpha P_e(\varphi_n, P_{XY}, Q_{XY}) \\
&\quad \times \left[1 - \sum_{(P,i)\in\mathcal{G}_n(\alpha)} Pr\{X^n \in \mathcal{T}_n(P,i)\}\right]
\end{aligned}
$$

where $\mathcal{G}_n^c(\alpha) = \{(P,i) \notin \mathcal{G}_n(\alpha) : P \in \mathcal{P}_n(\mathcal{X}), i = 1,2,\ldots,N_n(P)\}$. Therefore, for $\alpha > 0$, we have

$$\sum_{(P,i)\in\mathcal{G}_n(\alpha)} Pr\{X^n \in \mathcal{T}_n(P,i)\} \geq \frac{\alpha - 1}{\alpha}. \tag{8}$$

It is also easy to show by the weak law of large numbers that

$$\lim_{n\to\infty} \sum_{P\in\mathcal{S}_n(\epsilon)} Pr\{X^n \in \mathcal{T}_n(P)\} = 1. \tag{9}$$

Choose a positive number α^* such that

$$\frac{\alpha^* - 1}{\alpha^*} + \frac{\epsilon}{H_{M,v}(P_{XY}, Q_{XY}) + 2\epsilon} > 1.$$

Define $\mathcal{D}_n(\epsilon,\alpha^*) = \{(P,i) : (P,i) \in \mathcal{F}_n(\epsilon) \cap \mathcal{G}_n(\alpha^*), P \in \mathcal{S}_n(\epsilon)\}$. By (7)–(9), it is easy to see that $\mathcal{D}_n(\epsilon,\alpha^*)$ is nonempty for all sufficiently large n. Let $m_n = \lceil (1+\epsilon)n \rceil$. For each sufficiently large n and an arbitrary (P_n^*, i^*) from $\mathcal{D}_n(\epsilon,\alpha^*)$, we can construct a constant composition code $\mathcal{C}_{m_n}$ of length m_n and type P_{m_n} for some $P_{m_n} \in \mathcal{P}_{m_n}(\mathcal{X})$ by appending a fixed string in $\mathcal{X}^{m_n - n}$ to each sequence in $\mathcal{T}_n(P_n^*, i^*)$ such that

$$\lim_{n\to\infty} \|P_{m_n} - P_X\| = 0.$$

We retain the best half of the codewords in $\mathcal{C}_{m_n}$ and denote the resulting codebook by $\mathcal{C}'_{m_n}$. Note that

$$
\begin{aligned}
P_e(\mathcal{C}'_{m_n}, P_{Y|X}, Q_{Y|X}) &\leq 2Pr\{\text{error} \,|\, X^n \in \mathcal{T}_n(P_n^*, i^*)\} \\
&\leq 2\alpha^* P_e(\varphi_n, P_{XY}, Q_{XY}).
\end{aligned}
$$

Therefore, we have

$$\liminf_{n\to\infty} P_e(\mathcal{C}'_{m_n}, P_{Y|X}, Q_{Y|X}) = 0.$$

Moreover, it is easy to see that

$$
\begin{aligned}
&\liminf_{n\to\infty} R(\mathcal{C}'_{m_n}) \\
&\quad = \liminf_{n\to\infty} \frac{1}{m_n} \log\left[\frac{|\mathcal{T}_n(P_n^*, i^*)|}{2}\right] \\
&\quad \geq \liminf_{n\to\infty} \frac{n}{m_n}\left[\frac{1}{n}\log|\mathcal{T}_n(P_n^*)| - H_{M,v}(P_{XY}, Q_{XY}) - 2\epsilon\right] \\
&\quad \geq \frac{H(P_X) - H_{M,v}(P_{XY}, Q_{XY}) - 3\epsilon}{1+\epsilon}.
\end{aligned}
$$

Now by the definition of $C_M(P_X, P_{Y|X}, Q_{Y|X})$, we must have

$$C_M(P_X, P_{Y|X}, Q_{Y|X}) \geq \frac{H(P_X) - H_{M,v}(P_{XY}, Q_{XY}) - 3\epsilon}{1+\epsilon}.$$

Since $\epsilon > 0$ is arbitrary, this completes the proof. □

Remark: Although the minimum rate achievable with fixed-rate Slepian–Wolf codes is the same as that with variable-rate Slepian–Wolf codes, it does not mean that the attention can be restricted to fixed-rate codes with no loss of optimality. In fact, it has been shown [17] that variable-rate Slepian–Wolf codes can significantly outperform fixed-rate Slepian–Wolf codes in terms of rate-error tradeoff under matched decoding.

Theorem 1 implies that $H_{M,f}(P_{XY}, Q_{XY})$ and $H_{M,v}(P_{XY}, Q_{XY})$ depend on Q_{XY} only through $Q_{Y|X}$. Moreover, it is easy to see from the decoding rule (1) that $H_{M,f}(P_{XY}, Q_{XY})$ and $H_{M,v}(P_{XY}, Q_{XY})$ depend on Q_{XY} only through $Q_{X|Y}$, where $Q_{X|Y}$ is the conditional probability distribution of X given Y induced by Q_{XY}. Therefore, for any probability distributions $Q'_{XY} \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ with the property that $Q'_{XY}(x,y) = 0$ only if $P_{XY}(x,y) = 0$ for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, we have

$$
\begin{aligned}
H_{M,f}(P_{XY}, Q_{XY}) &= H_{M,v}(P_{XY}, Q_{XY}) \\
&= H_{M,f}(P_{XY}, Q'_{XY}) \\
&= H_{M,v}(P_{XY}, Q'_{XY})
\end{aligned}
$$

if $Q'_{Y|X} = Q_{Y|X}$ or $Q'_{X|Y} = Q_{X|Y}$, where $Q'_{Y|X}$ and $Q'_{X|Y}$ are the conditional probability distributions induced by Q'_{XY}.

Define $C_{LM}(P_X, P_{Y|X}, Q_{Y|X}) = \min_P D(P \,\|\, P_X P_Y)$, where the minimization is over the probability distributions $P \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ satisfying:

1) $\sum_{y\in\mathcal{Y}} P(x,y) = P_X(x)$ for all $x \in \mathcal{X}$;
2) $\sum_{x\in\mathcal{X}} P(x,y) = P_Y(y)$ for all $y \in \mathcal{Y}$;
3) $\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_X(x)P_{Y|X}(y|x)\log Q_{Y|X}(y|x) \leq \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P(x,y)\log Q_{Y|X}(y\,|\,x)$.

Note that $C_{LM}(P_X, P_{Y|X}, Q_{Y|X})$ is often referred to as the Csiszár–Körner–Hui lower bound on the mismatch capacity. Specifically, it was shown [8], [9] (cf., [12, Proposition 1]) that

$$C_M(P_X, P_{Y|X}, Q_{Y|X}) \geq C_{LM}(P_X, P_{Y|X}, Q_{Y|X}).$$

Moreover, for $|\mathcal{X}| = 2$, we have $C_M(P_X, P_{Y|X}, Q_{Y|X}) = C_{LM}(P_X, P_{Y|X}, Q_{Y|X})$ [11]. Therefore, it is straightforward to obtain the following result.

Corollary 1: $H_{M,f}(P_{XY}, Q_{XY}) = H_{M,v}(P_{XY}, Q_{XY}) \leq H(P_X) - C_{LM}(P_X, P_{Y\,|\,X}, Q_{Y\,|\,X})$, where the inequality can be replaced by equality if $|\mathcal{X}| = 2$.

Remark: It is easy to verify that

$$H(P_X) - C_{LM}(P_X, P_{Y\,|\,X}, Q_{Y\,|\,X}) = \max_P H(\tilde{X}\,|\,\tilde{Y})$$

where P is the joint probability distribution of $(\tilde{X}, \tilde{Y})$, and the maximization is over $P \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ satisfying:

1) $\sum_{y\in\mathcal{Y}} P(x,y) = P_X(x)$ for all $x \in \mathcal{X}$;
2) $\sum_{x\in\mathcal{X}} P(x,y) = P_Y(y)$ for all $y \in \mathcal{Y}$;
3) $\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_X(x) P_{Y\,|\,X}(y|x) \log Q_{Y|X}(y|x) \leq \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P(x,y)\log Q_{Y\,|\,X}(y\,|\,x)$.

IV. Linear Codebook-Level Duality

It is worth noting that the channel codes we used to construct Slepian–Wolf codes in the proof of Theorem 1 are constant composition codes, which are nonlinear; consequently, the resulting Slepian–Wolf codes are also nonlinear. Due to the wide applicability of linear codes in Slepian–Wolf coding and channel coding, it is natural to ask whether the duality established in the previous section continues to hold in the linear coding framework. The answer turns out to be negative. Indeed, even under matched decoding, the cosets induced by a good linear code for channel $P_{Y\,|\,X}$ may not be a good Slepian–Wolf code for source X with decoder side information Y; see [18] for a detailed discussion. It will be shown that in the linear coding framework, the type-level duality is replaced by a fundamentally different linear codebook-level duality.

Without loss of generality, we henceforth assume that $\mathcal{X} = \mathbb{Z}_K \triangleq \{0, 1, \ldots, K-1\}$. A linear channel code $\mathcal{C}_n$ of length n over $\mathbb{Z}_K$ is a subgroup of $\mathbb{Z}_K^n$, where $\mathbb{Z}_K^n$ is the group of n-tuples of elements of $\mathbb{Z}_K$ with componentwise modulo-K addition. Moreover, for any linear code $\mathcal{C}_n$ over $\mathbb{Z}_K$, we have $\mathcal{C}_n = \{c^n \in \mathbb{Z}_K^n : c^n\mathbf{H} = 0^n\}$ for some $n \times k$ parity check matrix $\mathbf{H}$. We will denote modulo-K addition, subtraction, and multiplication by $\oplus$, $\ominus$, and $\otimes$, respectively.

A channel with input alphabet $\mathcal{U} = \mathcal{X}$ and output alphabet $\mathcal{V} = \mathcal{X} \times \mathcal{Y}$ is cyclic symmetric if it can be written as $V = (U \oplus X, Y)$, where the channel input U is independent of (X, Y). Define two cyclic–symmetric channels $P_{V\,|\,U}$ and $Q_{V\,|\,U}$ by setting the joint probability of (X, Y) to be P_{XY} and Q_{XY}, respectively. Given a linear channel code $\mathcal{C}_n$ with parity check matrix $\mathbf{H}$, the induced linear Slepian–Wolf code is specified by the mapping $\phi_n : \mathbb{Z}_K^n \to \mathbb{Z}_K^k$ with $\phi_n(x^n) = x^n\mathbf{H}$ for all $x^n \in \mathbb{Z}_K^n$. Let $H_{M,l}(P_{XY}, Q_{XY})$ denote the minimum rate achievable with linear Slepian–Wolf codes over $\mathbb{Z}_K$ for source distribution P_{XY} under decoding metric Q_{XY}. Similarly, let $C_{M,l}(P_{V\,|\,U}, Q_{V\,|\,U})$ denote the maximum rate achievable with linear channel codes over $\mathbb{Z}_K$ for channel $P_{V\,|\,U}$ under decoding metric $Q_{U\,|\,V}$. Note that in the type-level duality, the channels of intrinsic importance are $P_{Y\,|\,X}$ and $Q_{Y\,|\,X}$; in contrast, the role of $P_{Y\,|\,X}$ and $Q_{Y\,|\,X}$, in a certain sense, is replaced by $P_{V\,|\,U}$ and $Q_{V\,|\,U}$ in the linear coding framework as shown by the following result.

Theorem 2: $H_{M,l}(P_{XY}, Q_{XY}) = \log K - C_{M,l}(P_{V\,|\,U}, Q_{V\,|\,U})$.

Proof: Given a linear channel code $\mathcal{C}_n$ over $\mathbb{Z}_K$ with parity check matrix $\mathbf{H}$, the rate of the induced Slepian–Wolf code $\phi_n(\cdot)$ is related to the rate of $\mathcal{C}_n$ by

$$R(\phi_n) = \log K - R(\mathcal{C}_n). \tag{10}$$

For any $u^n \in \mathcal{C}_n$, $x^n \in \mathcal{X}^n$, and $y^n \in \mathcal{Y}^n$, we have

$$\begin{aligned}
&\arg\min_{\hat{u}^n\in\mathcal{C}_n} -\sum_{i=1}^{n} Q_{V\,|\,U}(u_i \oplus x_i, y_i\,|\,\hat{u}_i)\\
&= \arg\min_{\hat{u}^n\in\mathcal{C}_n} -\sum_{i=1}^{n} Q_{XY}(u_i \oplus x_i \ominus \hat{u}_i, y_i)\\
&= u^n \oplus x^n \ominus \arg\min_{\hat{x}^n:\hat{x}^n\mathbf{H}=x^n\mathbf{H}} -\sum_{i=1}^{n} Q_{XY}(\hat{x}_i, y_i)
\end{aligned}$$

where the last equality follows from the fact that $\{u^n \oplus x^n \ominus \hat{u}^n : \hat{u}^n \in \mathcal{C}_n\} = \{\hat{x}^n : \hat{x}^n\mathbf{H} = x^n\mathbf{H}\}$. It is easy to see that a decoding error in Slepian–Wolf coding leads to a decoding error in channel coding, and *vice versa*; moreover, in this channel coding problem, the decoding error probability does not depend on the transmitted codeword. Therefore, we have

$$P_e(\phi_n, P_{XY}, Q_{XY}) = P_e(\mathcal{C}_n, P_{V\,|\,U}, Q_{V\,|\,U}). \tag{11}$$

The proof is complete by combining (10) and (11). □

Remark: For the binary source example considered by Slepian and Wolf (see Section I), it can be verified that under matched decoding, both $P_{V\,|\,U}$ and $Q_{V\,|\,U}$ degenerate to $P_{X\,|\,Y}$ (i.e., a binary symmetric channel with crossover probability p). Theorem 2 can also be viewed as a generalization of the duality results in [13] and [14] for the case where the side information Y^n is absent.

The purpose of introducing this linear codebook-level duality is twofold.

1) The linear codebook-level duality can be used as a technical tool. Specifically, it allows us to show that under mismatched decoding, the minimum rate achievable with nonlinear Slepian–Wolf codes can be strictly lower than that achievable with linear Slepian–Wolf codes.
2) The linear codebook-level duality can also be used as a design tool. It will be seen that under mismatched BP decoding, we can convert the Slepian–Wolf code design problem to the well-investigated channel code design problem by leveraging the linear codebook-level duality.

Now we proceed to derive the first main result of this section, namely, the suboptimality of linear Slepian–Wolf codes under mismatched decoding. Along the way, we will establish lower and upper bounds on $H_{M,l}(P_{XY}, Q_{XY})$, which are of interest by themselves.

Since $C_{M,l}(P_{V\,|\,U}, Q_{V\,|\,U}) \leq C_M(P_{V\,|\,U}, Q_{V\,|\,U})$, it follows from Theorem 2 that

$$H_{M,l}(P_{XY}, Q_{XY}) \geq \log K - C_M(P_{V\,|\,U}, Q_{V\,|\,U}) \tag{12}$$

which yields a lower bound on $H_{M,l}(P_{XY}, Q_{XY})$.

We will derive a computable upper bound on $H_{M,l}(P_{XY}, Q_{XY})$. For any $P_U \in \mathcal{P}(\mathcal{U})$, let P_V be the probability distribution of V induced by P_U and $P_{V|U}$, i.e., $P_V(v)=\sum_{u\in\mathcal{U}}P_U(u)P_{V|U}(v|u)$ for all $v\in\mathcal{V}$. Define $C_{GMI}(P_U,P_{V|U},Q_{V|U}) = \min_P D(P\|P_UP_V)$, where the minimization is over the distributions $P\in\mathcal{P}(\mathcal{U}\times\mathcal{V})$ satisfying:

1) $\sum_{u\in\mathcal{U}}P(u,v)=P_V(v)$ for all $v\in\mathcal{V}$;
2) $\sum_{u\in\mathcal{U},v\in\mathcal{V}}P_U(u)P_{V|U}(v|u)\log Q_{V|U}(v|u) \le \sum_{u\in\mathcal{U},v\in\mathcal{V}} P(v,u)\log Q_{V|U}(v|u)$.

Note that $C_{GMI}(P_U,P_{V|U},Q_{V|U})$ is known as the generalized mutual information. Moreover, we have [10]

$$C_{GMI}(P_U,P_{V|U},Q_{V|U})\le C_{LM}(P_U,P_{V|U},Q_{V|U}).$$

Theorem 3: Let P_U^* be the uniform distribution over $\mathcal{U}$.

1) $C_{GMI}(P_U^*,P_{V|U},Q_{V|U})=C_{LM}(P_U^*,P_{V|U},Q_{V|U})$.
2) $C_{LM}(P_U^*,P_{V|U},Q_{V|U}) = \log K - \max_P H(\tilde{X}|\tilde{Y})$, where P is the joint probability distribution of $(\tilde{X},\tilde{Y})$, and the maximization is over $P\in\mathcal{P}(\mathcal{X}\times\mathcal{Y})$ satisfying:
   a) $\sum_{x\in\mathcal{X}}P(x,y)=P_Y(y)$ for all $y\in\mathcal{Y}$;
   b) $\sum_{x\in\mathcal{X},y\in\mathcal{Y}}P_{XY}(x,y)\log Q_{XY}(x,y)\le\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P(x,y)\log Q_{XY}(x,y)$;

   specifically, $C_{LM}(P_U^*,P_{V|U},Q_{V|U})=0$ if and only if

$$\frac{1}{K}\sum_{x\in\mathcal{X},y\in\mathcal{Y}}P_Y(y)\log Q_{XY}(x,y) \ge \sum_{x\in\mathcal{X},y\in\mathcal{Y}}P_{XY}(x,y)\log Q_{XY}(x,y);$$

   if $C_{LM}(P_U^*,P_{V|U},Q_{V|U})>0$, then

$$C_{LM}(P_U^*,P_{V|U},Q_{V|U})=\log K+H(P_Y)-H(P^*)$$

   where

$$P^*(x,y)=\frac{P_Y(y)Q_{XY}^{\beta}(x,y)}{\sum_{x'\in\mathcal{X}}Q_{XY}^{\beta}(x',y)}$$

   and β ($\beta>0$) is determined by the constraint

$$\sum_{x\in\mathcal{X},y\in\mathcal{Y}}P^*(x,y)\log Q_{XY}(x,y) = \sum_{x\in\mathcal{X},y\in\mathcal{Y}}P_{XY}(x,y)\log Q_{XY}(x,y).$$

Proof: See the Appendix. □

To derive an upper bound on $H_{M,l}(P_{XY},Q_{XY})$, we need the following fundamental lemma from [14].

Lemma 1: Let $N_{\phi_n}(P_{\hat{X}\tilde{X}\hat{Y}})$ denote for each joint type $P_{\hat{X}\tilde{X}\hat{Y}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{X}\times\mathcal{Y})$ the number of pairs $(\hat{x}^n,\hat{y}^n)\in\mathcal{T}_n(P_{\hat{X}\hat{Y}})$ such that for some $\tilde{x}^n\ne\hat{x}^n$ with $P_{\hat{x}^n\tilde{x}^n\hat{y}^n}=P_{\hat{X}\tilde{X}\hat{Y}}$ the relation $\phi_n(\hat{x}^n)=\phi_n(\tilde{x}^n)$ holds.

If K is a prime number, then for arbitrary positive integers n and k, there exists a linear Slepian–Wolf code $\phi_n:\mathbb{Z}_K^n\to\mathbb{Z}_K^k$ such that for every joint type $P_{\hat{X}\tilde{X}\hat{Y}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{X}\times\mathcal{Y})$

$$\frac{1}{|\mathcal{T}_n(P_{\hat{X}\hat{Y}})|}N_{\phi_n}(P_{\hat{X}\tilde{X}\hat{Y}})\le 2^{-n|R(\phi_n)-H(\tilde{X}|\hat{X},\hat{Y})-\delta_n|^+}$$

if $\hat{X}\ne\tilde{X}$, where $\delta_n=\frac{\log(n+1)}{n}|\mathcal{X}|^2|\mathcal{Y}|$ and $|t|^+=\max(0,t)$.

Theorem 4: If K is a prime number, then

$$C_{M,l}(P_{V|U},Q_{V|U})\ge C_{GMI}(P_U^*,P_{V|U},Q_{V|U})$$
$$H_{M,l}(P_{XY},Q_{XY})\le\log K - C_{GMI}(P_U^*,P_{V|U},Q_{V|U}).$$

Remark: It follows from Theorem 3 that $\log K - C_{GMI}(P_U^*,P_{V|U},Q_{V|U}) = \max_P H(\tilde{X}|\tilde{Y})$, where P is the joint probability distribution of $(\tilde{X},\tilde{Y})$, and the maximization is over $P\in\mathcal{P}(\mathcal{X}\times\mathcal{Y})$ satisfying:

1) $\sum_{x\in\mathcal{X}}P(x,y)=P_Y(y)$ for all $y\in\mathcal{Y}$;
2) $\sum_{x\in\mathcal{X},y\in\mathcal{Y}}P_{XY}(x,y)\log Q_{XY}(x,y) \le \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P(x,y)\log Q_{XY}(x,y)$.

In view of the remark after Corollary 1, we have

$$\log K - C_{GMI}(P_U^*,P_{V|U},Q_{V|U}) \ge H(P_X)-C_{LM}(P_X,P_{Y|X},Q_{Y|X}).$$

Proof: By Theorem 2, it suffices to show that $H_{M,l}(P_{XY},Q_{XY})\le\log K - C_{GMI}(P_U^*,P_{V|U},Q_{V|U})$.

Note that

$$P_e(\phi_n,P_{XY},Q_{XY})\le\sum_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}N_{\phi_n}(P_{\hat{X}\tilde{X}\hat{Y}}) \times 2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+H(\hat{X},\hat{Y})]}$$

where Θ_n is the set of joint types $P_{\hat{X}\tilde{X}\hat{Y}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{X}\times\mathcal{Y})$ satisfying $\sum_{x\in\mathcal{X},y\in\mathcal{Y}}P_{\hat{X}\hat{Y}}(x,y)\log Q_{XY}(x,y)\le\sum_{x\in\mathcal{X},y\in\mathcal{Y}}P_{\tilde{X}\hat{Y}}(x,y)\log Q_{XY}(x,y)$. Using the linear Slepian–Wolf code $\phi_n(\cdot)$ as specified in Lemma 1, we have

$$\begin{aligned}
&\sum_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}N_{\phi_n}(P_{\hat{X}\tilde{X}\hat{Y}})2^{-n[D(P_{\hat{X}\hat{Y}}\|P_{XY})+H(\hat{X},\hat{Y})]}\\
&\le\sum_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}|\mathcal{T}_n(P_{\hat{X}\hat{Y}})|2^{-n|R(\phi_n)-H(\tilde{X}|\hat{X},\hat{Y})-\delta_n|^+}\\
&\quad\times 2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+H(\hat{X},\hat{Y})]}\\
&\le\sum_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}2^{nH(\hat{X},\hat{Y})}2^{-n|R(\phi_n)-H(\tilde{X}|\hat{X},\hat{Y})-\delta_n|^+}\\
&\quad\times 2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+H(\hat{X},\hat{Y})]}\\
&\le|\Theta_n|\\
&\quad\times\max_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+|R(\phi_n)-H(\tilde{X}|\hat{X},\hat{Y})-\delta_n|^+]}\\
&\le(n+1)^{|\mathcal{X}|^2|\mathcal{Y}|}\\
&\quad\times\max_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+|R(\phi_n)-H(\tilde{X}|\hat{X},\hat{Y})-\delta_n|^+]}\\
&\le(n+1)^{2|\mathcal{X}|^2|\mathcal{Y}|}\\
&\quad\times\max_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+|R(\phi_n)-H(\tilde{X}|\hat{X},\hat{Y})|^+]}\\
&\le(n+1)^{2|\mathcal{X}|^2|\mathcal{Y}|}\\
&\quad\times\max_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta_n}2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+|R(\phi_n)-H(\tilde{X}|\hat{Y})|^+]}.
\end{aligned}$$

Now it can be readily seen that for any $\epsilon>0$, if $R(\phi_n)\ge\log K - C_{GMI}(P_U^*,P_{V|U},Q_{V|U})+\epsilon$ for all sufficiently large n, then $\lim_{n\to\infty}P_e(\phi_n,P_{XY},Q_{XY})=0$. The proof is complete. □

By the definition of $H_{M,l}(P_{XY}, Q_{XY})$ and $H_{M,f}(P_{XY}, Q_{XY})$, it is obvious that

$$H_{M,l}(P_{XY}, Q_{XY}) \geq H_{M,f}(P_{XY}, Q_{XY}).$$

Note that under matched decoding, linear codes suffice to achieve the Slepian–Wolf limit, i.e.,

$$H_{M,l}(P_{XY}, P_{XY}) = H_{M,f}(P_{XY}, P_{XY}).$$

Therefore, it is natural to ask whether the equality continues to hold under mismatched decoding. The answer turns out to be negative as shown below.

It is known [12] that $C_M(P_{V\,|\,U}, Q_{V\,|\,U}) > 0$ if and only if $C_{LM}(P_U, P_{V\,|\,U}, Q_{V\,|\,U}) > 0$ for some $P_U \in \mathcal{P}(\mathcal{U})$ such that $P_U(u_1) = P_U(u_2) = \frac{1}{2}$ for some $u_1, u_2 \in \mathcal{U}$. Note that by (12), $C_M(P_{V\,|\,U}, Q_{V\,|\,U}) = 0$ implies $H_{M,l}(P_{XY}, Q_{XY}) = \log K$. On the other hand, we have $H_{M,f}(P_{XY}, Q_{XY}) \leq H(P_X) \leq \log K$, where the second inequality is strict unless P_X is the uniform distribution over $\mathcal{X}$. Therefore, the following result follows immediately.

Corollary 2: We have $H_{M,f}(P_{XY}, Q_{XY}) < H_{M,l}(P_{XY}, Q_{XY})$ under the conditions:

1) $\sum_{u\in\mathcal{U},v\in\mathcal{V}} P_U(u)P_{V|U}(v|u)\log Q_{V|U}(v|u) \leq \sum_{u\in\mathcal{U},v\in\mathcal{V}} P_U(u)(\sum_{u'\in\mathcal{U}} P_U(u')P_{V\,|\,U}(v\,|\,u'))\log Q_{V\,|\,U}(v\,|\,u)$ for all $P_U \in \mathcal{P}(\mathcal{U})$ such that $P_U(u_1) = P_U(u_2) = \frac{1}{2}$ for some $u_1, u_2 \in \mathcal{U}$;
2) $H(P_X) < \log K$.

It is worth emphasizing that different from $H_{M,f}(P_{XY}, Q_{XY})$ and $H_{M,v}(P_{XY}, Q_{XY})$, $H_{M,l}(P_{XY}, Q_{XY})$ does depend on Q_X. Note that for the case $K = 2$, we have

$$\begin{aligned} 1 - C_M(P_{V\,|\,U}, Q_{V\,|\,U}) &\leq H_{M,l}(P_{XY}, Q_{XY}) && (13)\\ &\leq 1 - C_{GMI}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) && (14)\\ &= 1 - C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) && (15)\end{aligned}$$

where (13) is due to (12), (14) is due to Theorem 4, and (15) is due to Theorem 3. Since $C_M(P_{V\,|\,U}, Q_{V\,|\,U}) > 0$ if and only if $C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V|U}) > 0$ [12], it follows that $H_{M,l}(P_{XY}, Q_{XY}) = 1$ if and only if $C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) = 0$. In view of Theorem 3, we have the following result.

Corollary 3: For the case $K = 2$, given $Q_{Y\,|\,X}$, we have $H_{M,l}(P_{XY}, Q_{XY}) = 1$ if

$$\begin{aligned}&\frac{1}{2}\sum_{x\in\mathcal{X}} \log Q_X(x) + \frac{1}{2}\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_Y(y)\log Q_{Y\,|\,X}(y\,|\,x)\\ &\quad\geq \sum_{x\in\mathcal{X}} P_X(x)\log Q_X(x) + \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y)\log Q_{Y|X}(y|x)\end{aligned}$$

and $H_{M,l}(P_{XY}, Q_{XY}) < 1$ if

$$\begin{aligned}&\frac{1}{2}\sum_{x\in\mathcal{X}} \log Q_X(x) + \frac{1}{2}\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_Y(y)\log Q_{Y\,|\,X}(y\,|\,x)\\ &\quad< \sum_{x\in\mathcal{X}} P_X(x)\log Q_X(x) + \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y)\log Q_{Y|X}(y|x).\end{aligned}$$

The reason why $H_{M,l}(P_{XY}, Q_{XY})$ can be greater than $H(P_X)$ is that linear Slepian–Wolf codes do not carry the type information of X^n. This problem can be remedied by adding an overhead that indicates the type of X^n to linear Slepian–Wolf codes. Since $\frac{1}{n}\log|\mathcal{P}_n(\mathcal{X})| \leq \frac{\log(n+1)}{n}|\mathcal{X}|$, the impact of this overhead on the code rate is negligible as $n \to \infty$. We will refer to the resulting codes as pseudolinear Slepian–Wolf codes.

Theorem 5: The rate $H(P_X) - C_{LM}(P_X, P_{Y\,|\,X}, Q_{Y\,|\,X})$ is achievable with pseudolinear Slepian–Wolf codes under decoding metric Q_{XY} if K is a prime number.

Proof: Let $\phi'_n(\cdot)$ be a pseudolinear Slepian–Wolf code consisting of a linear Slepian–Wolf code $\phi_n(\cdot)$ as specified in Lemma 1 and an overhead that indicates the type of X^n. By the argument similar to that in the proof of Theorem 4, we get

$$\begin{aligned}P_e(\phi'_n, P_{XY}, Q_{XY}) &\leq (n+1)^{2|\mathcal{X}|^2|\mathcal{Y}|}\\ &\quad\times \max_{P_{\hat{X}\tilde{X}\hat{Y}}\in\Theta'_n} 2^{-n[D(P_{\hat{X}\hat{Y}}\,\|\,P_{XY})+|R(\phi_n)-H(\tilde{X}|\hat{Y})|^+]}\end{aligned}$$

where Θ'_n is the set of joint types $P_{\hat{X}\tilde{X}\hat{Y}} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{X}\times\mathcal{Y})$ satisfying $\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{\tilde{X}\hat{Y}}(x,y)\log Q_{XY}(x,y) \leq \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{\hat{X}\hat{Y}}(x,y)\log Q_{XY}(x,y)$ and $P_{\hat{X}} = P_{\tilde{X}}$. Note that $R(\phi'_n) \leq R(\phi_n) + \frac{\log(n+1)}{n}|\mathcal{X}|$. Now it can be readily seen that for any $\epsilon > 0$, if $R(\phi'_n) \geq H(P_X) - C_{LM}(P_X, P_{Y\,|\,X}, Q_{Y\,|\,X}) + \epsilon$ for all sufficiently large n, then $\lim_{n\to\infty} P_e(\phi'_n, P_{XY}, Q_{XY}) = 0$. The proof is complete. □

Now we proceed to show that the linear codebook-level duality between Slepian–Wolf coding for source distribution P_{XY} under decoding metric Q_{XY} and channel coding for channel $P_{V\,|\,U}$ under decoding metric $Q_{V\,|\,U}$ continues to hold in the scenario where the belief propagation algorithm is used for decoding. A weaker version of this result under matched BP decoding was established in [18] for the special case $|\mathcal{X}| = 2$ using density evolution.

It is well known that each $n\times k$ parity check matrix $\mathbf{H}$ can be represented by a Tanner graph with n variable nodes and k check nodes. Let $\mathcal{C}_i$ denote the set of check nodes that are connected to variable node i, and $\mathcal{V}_j$ denote the set of variable nodes that are connected to check node j. We use h_{ij}, the (i,j) entry of $\mathbf{H}$, to label the edge connecting variable node i and check node j. Let x^n be the realization of X^n, and define $s^n = x^n\mathbf{H}$. Also, let u^n be the transmitted codeword, and y^n be the realization of Y^n.

In Slepian–Wolf coding, the initial message at variable node i, denoted as $M_{v=i}^{(0)}$, is given by

$$M_{v=i}^{(0)} = \left[m_{v=i}^{(0)}(0), \ldots, m_{v=i}^{(0)}(K-1)\right]$$

where $m_{v=i}^{(0)}(k) = Q_{XY}(k, y_i)$ for all $k \in \mathbb{Z}_K$. In channel coding, the initial message at variable node i, denoted as $\bar{M}_{v=i}^{(0)}$, is given by

$$\bar{M}_{v=i}^{(0)} = \left[\bar{m}_{v=i}^{(0)}(0), \ldots, \bar{m}_{v=i}^{(0)}(K-1)\right]$$

where $\bar{m}^{(0)}_{v=i}(k) = Q_{V\,|\,U}(u_i \oplus x_i, y_i\,|\,k)$ for all $k \in \mathbb{Z}_K$. It is easy to see that

$$\bar{m}^{(0)}_{v=i}(k) = m^{(0)}_{v=i}(u_i \oplus x_i \ominus k)$$

for all $k \in \mathbb{Z}_K$.

Now consider the message from check node j to variable node i in the first iteration. In Slepian–Wolf coding, the message, denoted as $M^{(1)}_{c=j,v=i}$, is given by

$$M^{(1)}_{c=j,v=i} = \left[m^{(1)}_{c=j,v=i}(0), \dots, m^{(1)}_{c=j,v=i}(K-1)\right]$$

where

$$m^{(1)}_{c=j,v=i}(k) = \sum_{\mathcal{A}_k} \prod_{l \in \mathcal{C}_j \setminus \{i\}} m^{(0)}_{v=l}(\tilde{x}_l)$$
$$\mathcal{A}_k = \left\{ (\tilde{x}_l)_{l \in \mathcal{C}_j \setminus \{i\}} : h_{ij} \otimes k \oplus \bigoplus_{l \in \mathcal{C}_j \setminus \{i\}} h_{lj} \otimes \tilde{x}_l = s_j \right\}.$$

In channel coding, the message, denoted as $\overline{M}^{(1)}_{c=j,v=i}$, is given by

$$\bar{M}^{(1)}_{c=j,v=i} = [\bar{m}^{(1)}_{c=j,v=i}(0), \dots, \bar{m}^{(1)}_{c=j,v=i}(K-1)]$$

where

$$\bar{m}^{(1)}_{c=j,v=i}(k) = \sum_{\bar{\mathcal{A}}_k} \prod_{l \in \mathcal{C}_j \setminus \{i\}} \bar{m}^{(0)}_{v=l}(\tilde{u}_l)$$
$$\bar{\mathcal{A}}_k = \left\{ (\tilde{u}_l)_{l \in \mathcal{C}_j \setminus \{i\}} : h_{ij} \otimes k \oplus \bigoplus_{l \in \mathcal{C}_j \setminus \{i\}} h_{lj} \otimes \tilde{u}_l = 0 \right\}.$$

Note that

$$\begin{aligned}
\bar{m}^{(1)}_{c=j,v=i}(k) &= \sum_{\bar{\mathcal{A}}_k} \prod_{l \in \mathcal{C}_j \setminus \{i\}} \bar{m}^{(0)}_{v=l}(\tilde{u}_l) \\
&= \sum_{\bar{\mathcal{A}}_k} \prod_{l \in \mathcal{C}_j \setminus \{i\}} m^{(0)}_{v=l}(u_l \oplus x_l \ominus \tilde{u}_l) \\
&= \sum_{\mathcal{A}_{u_i \oplus x_i \ominus k}} \prod_{l \in \mathcal{C}_j \setminus \{i\}} m^{(0)}_{v=l}(\tilde{x}_l) \\
&= m^{(1)}_{c=j,v=i}(u_i \oplus x_i \ominus k), \qquad k \in \mathbb{Z}_K.
\end{aligned}$$

Then, consider the message from variable node i to check node j in the first iteration. In Slepian–Wolf coding, the message, denoted as $M^{(1)}_{v=i,c=j}$, is given by

$$M^{(1)}_{v=i,c=j} = \left[m^{(1)}_{v=i,c=j}(0), \dots, m^{(1)}_{v=i,c=j}(K-1)\right]$$

where

$$m^{(1)}_{v=i,c=j}(k) = m^{(0)}_{v=i}(k) \prod_{l \in \mathcal{V}_i \setminus \{j\}} m^{(1)}_{c=l,v=i}(k).$$

In channel coding, the message, denoted as $\bar{M}^{(1)}_{v=i,c=j}$, is given by

$$\bar{M}^{(1)}_{v=i,c=j} = \left[\bar{m}^{(1)}_{v=i,c=j}(0), \dots, \bar{m}^{(1)}_{v=i,c=j}(K-1)\right]$$

where

$$\bar{m}^{(1)}_{v=i,c=j}(k) = \bar{m}^{(0)}_{v=i}(k) \prod_{l \in \mathcal{V}_i \setminus \{j\}} \bar{m}^{(1)}_{c=l,v=i}(k).$$

Clearly, we have

$$\begin{aligned}
&\bar{m}^{(1)}_{v=i,c=j}(k) \\
&\quad = m^{(0)}_{v=i}(u_i \oplus x_i \ominus k) \prod_{l \in \mathcal{V}_i \setminus \{j\}} m^{(1)}_{c=l,v=i}(u_i \oplus x_i \ominus k) \\
&\quad = m^{(1)}_{v=i,c=j}(u_i \oplus x_i \ominus k), \qquad k \in \mathbb{Z}_K.
\end{aligned}$$

By induction, for any iteration number t, variable node i, and check node j, we have

$$\begin{aligned}
\bar{m}^{(t)}_{v=i,c=j}(k) &= m^{(t)}_{v=i,c=j}(u_i \oplus x_i \ominus k) \\
\bar{m}^{(t)}_{c=j,v=i}(k) &= m^{(t)}_{c=j,v=i}(u_i \oplus x_i \ominus k)
\end{aligned}$$

for all $k \in \mathbb{Z}_K$.

Suppose at the t^*th iteration, a decision is to be made at variable node i by forming a decision vector. In Slepian–Wolf coding, the decision vector, denoted as $D^{(t^*)}_i$, is given by

$$D^{(t^*)}_i = [d^{(t^*)}_i(0), \dots, d^{(t^*)}_i(K-1)]$$

where

$$d^{(t^*)}_i(k) = m^{(0)}_{v=i}(k) \prod_{l \in \mathcal{V}_i} m^{(t^*)}_{c=l,v=i}(k).$$

In channel coding, the decision vector, denoted as $\bar{D}^{(t^*)}_i$, is given by

$$\bar{D}^{(t^*)}_i = \left[\bar{d}^{(t^*)}_i(0), \dots, \bar{d}^{(t^*)}_i(K-1)\right]$$

where

$$\bar{d}^{(t^*)}_i(k) = \bar{m}^{(0)}_{v=i}(k) \prod_{l \in \mathcal{V}_i} \bar{m}^{(t^*)}_{c=l,v=i}(k).$$

It is easy to see that

$$\begin{aligned}
\bar{d}^{(t^*)}_i(k) &= m^{(0)}_{v=i}(u_i \oplus x_i \ominus k) \prod_{l \in \mathcal{V}_i} m^{(t^*)}_{c=l,v=i}(u_i \oplus x_i \ominus k) \\
&= d^{(t^*)}_i(u_i \oplus x_i \ominus k), \qquad k \in \mathbb{Z}_K.
\end{aligned}$$

Note that if $\bar{d}^{(t^*)}_i(k)$ is maximized at $k = u_i$, then $d^{(t^*)}_i(k')$ is maximized at $k' = x_i$, and *vice versa*, which implies that the decoding error probabilities in these two problems must be the same. One can readily develop the density evolution algorithm for the channel coding problem,[3] which, by the linear codebook-level duality, yields the density evolution algorithm for Slepian–Wolf coding under mismatched BP decoding. Moreover, it can be verified that the decoding metric determines the set of possible values of the initial message while the source

[3]Note that the all-zero codeword assumption is valid in the current setting.

distribution or the channel transition probability distribution determines the distribution of the initial message over that set of possible values. In view of the fact that channel code design using density evolution is a well-studied subject, the linear codebook-level duality effectively provides a tool for designing Slepian–Wolf codes.

V. Conclusion

Two different dualities between Slepian–Wolf coding and channel coding under mismatched decoding are established. It is shown that under mismatched decoding, the minimum rate achievable with variable-rate Slepian–Wolf codes is the same as that achievable with fixed-rate Slepian–Wolf codes while the minimum rate achievable with nonlinear Slepian–Wolf codes can be strictly lower than that achievable with linear Slepian–Wolf codes.

Appendix
Proof of Theorem 3

Let P_V^* be the probability distribution of V induced by P_U^* and $P_{V\,|\,U}$. For any $v \in \mathcal{V}$, we can write $v = (v_1, v_2)$ with $v_1 \in \mathcal{X}$ and $v_2 \in \mathcal{Y}$. Recall that P_U^* is the uniform distribution over $\mathcal{U}$. Therefore, we have

$$P_V^*(v) = P_V^*(v_1, v_2) = \frac{1}{K} P_Y(v_2).$$

1) Consider the optimization problem

$$C_{GMI}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) = \min_P D(P \,\|\, P_U^* P_V^*) \tag{16}$$

where the minimization is over the distributions $P \in \mathcal{P}(\mathcal{U} \times \mathcal{V})$ satisfying:

a) $\sum_{u\in\mathcal{U}} P(u,v) = P_V^*(v)$ for all $v \in \mathcal{V}$;

b) $\sum_{u\in\mathcal{U},v\in\mathcal{V}} P_U^*(u) P_{V\,|\,U}(v\,|\,u) \log Q_{V\,|\,U}(v\,|\,u) \le \sum_{u\in\mathcal{U},v\in\mathcal{V}} P(v,u) \log Q_{V\,|\,U}(v\,|\,u)$.

Suppose the minimum in (16) is achieved at $\tilde{P}$. Define

$$P^{(x)}(u, v_1, v_2) = \tilde{P}(u \oplus x, v_1 \oplus x, v_2)$$
$$\bar{P}(u, v_1, v_2) = \frac{1}{K} \sum_{x\in\mathcal{X}} P^{(x)}(u, v_1, v_2)$$

for all $x \in \mathcal{X}$, $u \in \mathcal{U}$, $v_1 \in \mathcal{X}$, and $v_2 \in \mathcal{Y}$. It is easy to verify that

$$D(P^{(x)} \,\|\, P_U^* P_V^*) = C_{GMI}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U})$$

for all $x \in \mathcal{X}$. Moreover, we have

$$\sum_{u\in\mathcal{U}} \bar{P}(u, v_1, v_2) = \frac{1}{K} \sum_{u\in\mathcal{U}} \sum_{x\in\mathcal{X}} \tilde{P}(u \oplus x, v_1 \oplus x, v_2) = P_V^*(v)$$

and

$$\begin{aligned}
&\sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} \bar{P}(u, v_1, v_2) \log Q_{V\,|\,U}(v_1, v_2\,|\,u) \\
&= \frac{1}{K} \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} \sum_{x\in\mathcal{X}} \tilde{P}(u \oplus x, v_1, v_2) \times \log Q_{V\,|\,U}(v_1, v_2\,|\,u) \\
&= \frac{1}{K} \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} \sum_{x\in\mathcal{X}} \tilde{P}(u \oplus x, v_1, v_2) \times \log Q_{V\,|\,U}(v_1 \oplus x, v_2\,|\,u \oplus x) \\
&\ge \frac{1}{K} \sum_{x\in\mathcal{X}} \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} P_U^*(u) P_{V\,|\,U}(v\,|\,u) \times \log Q_{V\,|\,U}(v\,|\,u) \\
&= \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} P_U^*(u) P_{V\,|\,U}(v\,|\,u) \times \log Q_{V\,|\,U}(v\,|\,u).
\end{aligned}$$

In view of the fact that $D(P \,\|\, P_U^* P_V^*)$ is a convex function of P, the minimum in (16) is attained at $\bar{P}$. Since

$$\begin{aligned}
&\sum_{v_1\in\mathcal{X},v_2\in\mathcal{Y}} \bar{P}(u, v_1, v_2) \\
&= \frac{1}{K} \sum_{v_1\in\mathcal{X},v_2\in\mathcal{Y}} \sum_{x\in\mathcal{X}} \tilde{P}(u \oplus x, v_1 \oplus x, v_2) \\
&= \frac{1}{K}
\end{aligned}$$

by the definition of $C_{GMI}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U})$ and $C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U})$, we must have

$$C_{GMI}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) = C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}).$$

2) For any conditional probability distribution $W_{V\,|\,U}$, let $P_U^* \circ W_{V\,|\,U}$ be the joint probability distribution over $\mathcal{U} \times \mathcal{V}$ induced by P_U^* and $W_{V\,|\,U}$. Consider the optimization problem

$$C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) = \min_{W_{U\,|\,V}} D(P_U^* \circ W_{V\,|\,U} \,\|\, P_U^* P_V^*) \tag{17}$$

where the minimization is over $W_{U\,|\,V}$ subject to the constraints:

a) $\sum_u P_U^*(u \in \mathcal{U}) W_{V\,|\,U}(v\,|\,u) = P_V^*(v)$ for all $v \in \mathcal{V}$;

b) $\sum_{u\in\mathcal{U},v\in\mathcal{V}} P_U^*(u) W_{V\,|\,U}(v\,|\,u) \log Q_{V\,|\,U}(v\,|\,u) \ge \sum_{u\in\mathcal{U},v\in\mathcal{V}} P_U^*(u) P_{V\,|\,U}(v\,|\,u) \log Q_{V\,|\,U}(v\,|\,u)$

.

Suppose the minimizer to (17) is given by $\tilde{W}_{V\,|\,U}$. Define

$$W_{V\,|\,U}^{(x)}(v_1, v_2\,|\,u) = \tilde{W}_{V\,|\,U}(v_1 \oplus x, v_2\,|\,u \oplus x)$$
$$W_{V\,|\,U}^*(v_1, v_2\,|\,u) = \frac{1}{K} \sum_{x\in\mathcal{X}} W_{V\,|\,U}^{(x)}(v_1, v_2\,|\,u)$$

for all $x \in \mathcal{X}$, $u \in \mathcal{U}$, $v_1 \in \mathcal{X}$, and $v_2 \in \mathcal{Y}$. It is easy to verify that

$$D\left(P_U^* \circ W_{V\,|\,U}^{(x)} \,\|\, P_U^* P_V^*\right) = C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U})$$

for all $x \in \mathcal{X}$. Moreover, we have the equation shown at the bottom of the page. Since $D(P_U^* \circ W_{V\,|\,U} \,\|\, P_U^* P_V^*)$ is a convex function of $W_{V\,|\,U}$, it follows that the minimum in (17) is achieved at $W_{V\,|\,U}^*$. Note that $W_{V\,|\,U}^*$ is a cyclic symmetric channel; therefore, it can be written as $V = (U \oplus X, Y)$ with (X, Y) specified by some probability distribution P^*. It is easy to verify that

$$D(P_U^* \circ W_{V\,|\,U}^* \,\|\, P_U^* P_V^*) = \log K + H(P_Y) - H(P^*).$$

Therefore, we can rewrite the minimization problem (17) in the following equivalent form:

$$\begin{aligned} C_{LM}&(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) \\ &= \min_P \log K - H(\tilde{X}\,|\,\tilde{Y}) \\ &= \min_P \log K + H(P_Y) - H(P) \end{aligned} \tag{18}$$

where P is the joint probability distribution of $(\tilde{X}, \tilde{Y})$, and the minimization is over $P \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ with the constraints:

a) $\sum_{x\in\mathcal{X}} P(x,y) = P_Y(y)$ for all $y \in \mathcal{Y}$;
b) $\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P(x,y) \log Q_{XY}(x,y) \geq \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y) \log Q_{XY}(x,y)$.

It can be seen that the minimum in (18) is attained at P^*. Introducing Lagrangian multipliers α_y $(y \in \mathcal{Y})$ and β $(\beta \geq 0)$, we define

$$G = -H(P) + \sum_{y\in\mathcal{Y}} \alpha_y \sum_{x\in\mathcal{X}} P(x,y) - \beta \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P(x,y) \log Q_{XY}(x,y).$$

The Karush–Kuhn–Tucker conditions yield

$$\log P^*(x,y) + \log e + \alpha_y - \beta \log Q_{XY}(x,y) = 0 \qquad x \in \mathcal{X}, y \in \mathcal{Y},$$

$$\sum_{x\in\mathcal{X}} P^*(x,y) = P_Y(y), \qquad y \in \mathcal{Y}$$

$$\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P^*(x,y) \log Q_{XY}(x,y) \geq \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y) \log Q_{XY}(x,y)$$

$$\beta \left[\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P^*(x,y) \log Q_{XY}(x,y) - \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y) \log Q_{XY}(x,y) \right] = 0.$$

Therefore

$$P^*(x,y) = \frac{P_Y(y) Q_{XY}^{\beta}(x,y)}{\sum_{x'\in\mathcal{X}} Q_{XY}^{\beta}(x',y)}.$$

If $\beta > 0$, then we must have

$$\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P^*(x,y) \log Q_{XY}(x,y) = \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y) \log Q_{XY}(x,y).$$

If $\beta = 0$, then $P^*(x,y) = \frac{1}{K} P_Y(y)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, which implies $C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) = 0$. Conversely, if $C_{LM}(P_U^*, P_{V\,|\,U}, Q_{V\,|\,U}) = 0$, then $P^*(x,y) = \frac{1}{K} P_Y(y)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

$$\sum_{u\in\mathcal{U}} P_U^*(u) W_{V\,|\,U}^*(v_1, v_2\,|\,u) = \frac{1}{K} \sum_{x\in\mathcal{X}} \sum_{u\in\mathcal{U}} P_U^*(u) \tilde{W}_{V\,|\,U}(v_1 \oplus x, v_2\,|\,u \oplus x) = \frac{1}{K} \sum_{x\in\mathcal{X}} P_V^*(v_1 \oplus x, v_2) = \frac{1}{K} P_Y(v_2)$$

and

$$\begin{aligned} &\sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} P_U^*(u) W_{V\,|\,U}^*(v_1,v_2\,|\,u) \log Q_{V\,|\,U}(v_1,v_2\,|\,u) \\ &= \frac{1}{K} \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} P_U^*(u) \sum_{x\in\mathcal{X}} \tilde{W}_{V\,|\,U}(v_1 \oplus x, v_2\,|\,u \oplus x) \log Q_{V\,|\,U}(v_1,v_2\,|\,u) \\ &= \frac{1}{K^2} \sum_{x\in\mathcal{X}} \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} \tilde{W}_{V\,|\,U}(v_1 \oplus x, v_2\,|\,u \oplus x) \log Q_{V\,|\,U}(v_1,v_2\,|\,u) \\ &= \frac{1}{K^2} \sum_{x\in\mathcal{X}} \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} \tilde{W}_{V\,|\,U}(v_1 \oplus x, v_2\,|\,u \oplus x) \log Q_{V\,|\,U}(v_1 \oplus x, v_2\,|\,u \oplus x) \\ &\geq \frac{1}{K} \sum_{x\in\mathcal{X}} \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} P_U^*(u) P_{V\,|\,U}(v_1,v_2\,|\,u) \log Q_{V\,|\,U}(v_1,v_2\,|\,u) \\ &= \sum_{u\in\mathcal{U},v_1\in\mathcal{X},v_2\in\mathcal{Y}} P_U^*(u) P_{V\,|\,U}(v_1,v_2\,|\,u) \log Q_{V\,|\,U}(v_1,v_2\,|\,u). \end{aligned}$$

Therefore, the necessary and sufficient condition for $C_{LM}(P_U^*, P_{V|U}, Q_{V|U}) = 0$ is

$$\frac{1}{K}\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_Y(y)\log Q_{XY}(x,y) \geq \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y)\log Q_{XY}(x,y).$$

The proof is complete.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the associate editor Prof. Kontoyiannis for their constructive comments that helped to improve the presentation of this paper.

REFERENCES

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[2] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[3] R. Ahlswede and G. Dueck, "Good codes can be produced by a few permutations," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 3, pp. 430–443, May 1982.

[4] J. K. Wolf, "Tribute to David Slepian," *IEEE Inf. Theory Soc. Newsletter*, vol. 58, pp. 3–4, Mar. 2008.

[5] A. D. Wyner, "Recent results in Shannon theory," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 2–10, Jan. 1974.

[6] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.

[7] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.

[8] I. Csiszár and J. Körner, "Graph decomposition: A new key to coding theorems," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 5–12, Jan. 1981.

[9] J. Y. N. Hui, "Fundamental issues of multiple accessing," Ph.D. dissertation, Dept. Electr. Eng. Comput. Sci., Mass. Inst. Technol., Cambridge, MA, 1983.

[10] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai (Shitz), "On information rates for mismatched decoders," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[11] V. B. Balakirsky, "A converse coding theorem for mismatched decoding at the output of binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1889–1902, Nov. 1995.

[12] I. Csiszár and P. Narayan, "Channel capacity for a given decoding metric," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 35–43, Jan. 1995.

[13] T. C. Ancheta, Jr., "Syndrome-source-coding and its universal generalizations," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 4, pp. 432–436, Jul. 1976.

[14] I. Csiszár, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 4, pp. 585–592, Jul. 1982.

[15] I. Csiszár, "The method of types," *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 2505–2523, Oct. 1998.

[16] R. Ahlswede, "Coloring hypergraphs: A new approach to multi-user source coding—II," *J. Combinatorics Inf. Syst. Sci.*, vol. 5, pp. 220–268, 1980.

[17] J. Chen, D.-k. He, A. Jagmohan, and L. Lastras, "On the reliability function of variable-rate Slepian-Wolf coding," in *Proc. 45th Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, Sep. 26–28, 2007, pp. 292–299.

[18] J. Chen, D.-k. He, and A. Jagmohan, "The equivalence between Slepian-Wolf coding and channel coding under density evolution," *IEEE Trans. Commun.*, to be published.

Jun Chen (S'03–M'06) received the B.E. degree with honors in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2001 and the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, in 2003 and 2006, respectively.

He was a Postdoctoral Research Associate in the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, from 2005 to 2006, and a Josef Raviv Memorial Postdoctoral Fellow at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, from 2006 to 2007. He is currently an Assistant Professor of Electrical and Computer Engineering at McMaster University, Hamilton, ON, Canada. He holds the Barber-Gennum Chair in Information Technology. His research interests include information theory, wireless communications, and signal processing.

Da-ke He received the B.S. and M.S. degrees in electrical engineering from Huazhong University of Science and Technology, Wuhan, Hubei, China, in 1993 and 1996, respectively, and the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 2003.

From 1996 to 1998, he worked at Apple Technology China, Zhuhai, China, as a Software Engineer. From 2003 to 2004, he worked in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, as a Postdoctoral Research Fellow in the Leitch-University of Waterloo Multimedia Communications Lab. From 2005 to 2008, he was a Research Staff Member in the Department of Multimedia Technologies, IBM T. J. Watson Research Center, Yorktown Heights, NY. Since 2008, he has been a technical manager in Slipstream Data, a subsidiary of Research In Motion, in Waterloo, ON, Canada. His research interests are in source coding theory and algorithm design, multimedia data compression and transmission, multiterminal source coding theory and algorithms, and digital communications.

Ashish Jagmohan received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, India, in 1999 and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, in 2002 and 2004, respectively.

Since 2004, he has worked at the Departments of Multimedia Technologies and Memory Systems in the IBM TJ Watson Research Center, Yorktown Heights, NY. His research interests include memory system technologies, video compression, multimedia communication, signal processing, and information theory.