Improved Finite Blocklength Converses for Slepian-Wolf Coding via Linear Programming

Sharu Theresa Jose Ankur A. Kulkarni

arXiv:1801.08693v1 [cs.IT] 26 Jan 2018

***Abstract*—A new finite blocklength converse for the Slepian-Wolf coding problem is presented which significantly improves on the best known converse for this problem, due to Miyake and Kanaya [2]. To obtain this converse, an extension of the linear programming (LP) based framework for finite blocklength point-to-point coding problems from [3] is employed. However, a direct application of this framework demands a complicated analysis for the Slepian-Wolf problem. An analytically simpler approach is presented wherein LP-based finite blocklength converses for this problem are synthesized from point-to-point lossless source coding problems with perfect side-information at the decoder. New finite blocklength metaconverses for these point-to-point problems are derived by employing the LP-based framework, and the new converse for Slepian-Wolf coding is obtained by an appropriate combination of these converses.**

I. Introduction

The intractability of evaluating the nonasymptotic or finite blocklength fundamental limit of communication has put the onus on discovering finite blocklength achievability and converses that sandwich tightly the nonasymptotic fundamental limit. Accordingly, recent years have witnessed a surge of tight finite blocklength achievability and converses ([4], [5], [6], [3]), particularly for coding problems in the point-to-point setting.

Eventhough many sharp and asymptotically tight finite blocklength converses have been obtained in the point-to-point setting employing tools like hypothesis testing [4] and information spectrum [7], deriving tight finite blocklength converses for multiterminal coding problems still remains particularly challenging. Part of this challenge could be attributed to the difficulty in extending the techniques in the point-to-point setting to the network setting. In this paper, we consider the classical multiterminal source coding problem – the Slepian-Wolf coding problem and show that the extension of the linear programming (LP) based framework we introduced for the point-to-point setting in [3], in fact results in new and improved finite blocklength converses for this problem. Moreover, it yields a framework via a hierarchy of relaxations in which classical converses can be recovered, and converses for the networked problem can be synthesized using a combination of point-to-point converses.

Consider the finite blocklength Slepian-Wolf distributed lossless source coding problem (in Figure 1) posed as the

Sharu and Ankur are with the Systems and Control Engineering group at the Indian Institute of Technology Bombay, Mumbai, 400076, India. email: sharutheresa@iitb.ac.in, kulkarni.ankur@iitb.ac.in. This work was presented in part at the IEEE Information Theory Workshop, in Kaohsiung, Taiwan, 2017 [1].

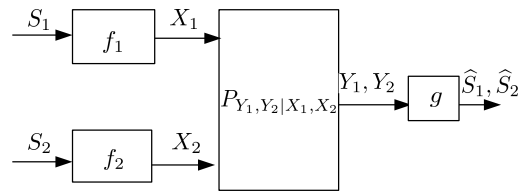

Fig. 1: Two-User Joint Source-Channel Coding Problem

following optimization problem,

$$\begin{array}{ll} \text{SW} & \min\limits_{f_1,f_2,g} \ \mathbb{E}[\mathbb{I}\{(S_1,S_2) \neq (\widehat{S}_1,\widehat{S}_2)\}] \\ & X_1 = f_1(S_1), \\ \text{s.t.} & X_2 = f_2(S_2), \\ & (\widehat{S}_1,\widehat{S}_2) = g(Y_1,Y_2), \end{array}$$

where $S_1, S_2, X_1, X_2, Y_1, Y_2, \widehat{S}_1, \widehat{S}_2$ are discrete random variables taking values in fixed, finite spaces $\mathcal{S}_1, \mathcal{S}_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, \widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2$, respectively. Notice that these spaces could be Cartesian products of smaller spaces, and hence could be sets of finite length strings. Here, S_1 and S_2 represent the two *correlated* sources distributed according to a known joint probability distribution P_{S_1,S_2}. The source signals are *seperately* encoded by functions $f_1 : \mathcal{S}_1 \to \mathcal{X}_1$ and $f_2 : \mathcal{S}_2 \to \mathcal{X}_2$ to produce signals $X_1 = f_1(S_1)$ and $X_2 = f_2(S_2)$, respectively. The encoded signals are sent through a deterministic channel with conditional distribution $P_{Y_1,Y_2|X_1,X_2} = \mathbb{I}\{(Y_1,Y_2) = (X_1,X_2)\}$ to get the signal (Y_1,Y_2), where $\mathbb{I}\{\cdot\}$ represents the indicator function which equals unity when '$\cdot$' is true and is zero otherwise. (Y_1,Y_2) is then *jointly* decoded by $g : \mathcal{Y}_1 \times \mathcal{Y}_2 \to \widehat{\mathcal{S}}_1 \times \widehat{\mathcal{S}}_2$ to obtain the output signal $(\widehat{S}_1,\widehat{S}_2) = g(Y_1,Y_2)$. For the finite blocklength Slepian-Wolf coding problem, we note that spaces $\mathcal{S}_1 = \widehat{\mathcal{S}}_1$, $\mathcal{S}_2 = \widehat{\mathcal{S}}_2$, $\mathcal{X}_1 = \mathcal{Y}_1 = \{1,\ldots,M_1\}$ and $\mathcal{X}_2 = \mathcal{Y}_2 = \{1,\ldots,M_2\}$, $M_1, M_2 \in \mathbb{N}$. An error in transmission occurs when $(S_1,S_2) \neq (\widehat{S}_1,\widehat{S}_2)$. Hence, the objective of the finite blocklength Slepian-Wolf coding problem SW is to minimize the probability of error over all *codes*, i.e. over all encoder-decoder functions (f_1, f_2, g).

Our interest in this paper is in obtaining finite blocklength converses (or lower bounds) on the optimal value of SW and our approach is via the linear programming (LP) based framework introduced in [3]. In [3], we showed that this framework recovers and improves on most of the well-known finite blocklength converses for point-to-point coding problems. In particular, the LP framework is shown to imply the metaconverse of Polyanskiy-Poor-Verdú [4] for finite blocklength channel coding. For lossy source coding and lossy joint source-channel coding with the probability of excess

distortion as the loss criterion, the LP framework results in two levels of improvements on the asymptotically tight tilted-information based converses of Kostina and Verdú in [5] and [6], respectively.

Fundamental to this framework is the observation that the finite blocklength coding problem can be posed equivalently as a nonconvex optimization problem over joint probability distributions. A natural optimizer's approach [8] to obtain lower bounds would then be via a convex relaxation of the nonconvex optimization problem. In particular, resorting to the "lift-and-project" technique due to Lovasz, Schrijver, Sherali, Adams and others [9], we obtain a LP relaxation of the problem. From linear programming duality, we then get that the objective value of *any feasible point of the dual of this LP relaxation yields a lower bound* on the optimal loss in the finite blocklength problem. As a result of this observation, the problem of obtaining converses reduces to constructing feasible points for the dual linear program.

The converses in [3] stated above for various point-to-point settings emerge as special cases of this LP-based framework, implied by the construction of specific dual feasible points. This tightness of the LP relaxation shows that there is an alternative, asymptotically tight way of thinking about optimal finite blocklength coding – as the optimal packing of a pair of *source* and *channel flows* satisfying a certain error density bottleneck. The flows here are the variables of the dual program and the bottleneck, its constraint.

In this paper, we further this theme towards the Slepian-Wolf coding problem. In the present paper we observe that our LP relaxation has an operational interpretation based on optimal transport [10], wherein one designs not only codes, but also couplings between them to minimize the resulting 'error'. Using the LP relaxation, we first establish new, clean, meta-converses in the point-to-point setting for lossy source-coding problems with side-information at the decoder; these converses are stronger than our earlier converses in [3], they imply the hypothesis testing and tilted information based converses of Kostina and Verdú [5, Theorem 7,Theorem 8] and the converse of Palzer and Timo [11, Theorem 1], and are, to the best of our knowledge, the strongest known. Subsequently, we analyse the dual LP of the finite blocklength Slepian-Wolf coding problem. When extended to the networked Slepian-Wolf coding problem, the LP-based framework results in a large number of dual variables and constraints, which makes it quite challenging to analyse and interpret. Consequently, we devise an analytically simpler approach to construct feasible points of the dual program using feasible points of simpler point-to-point problems. This yields tight finite blocklength converses that improve on the hitherto best known converse for this problem, due to Miyake and Kanaya [2].

The dual variables of the LP relaxation of SW also have a structure of 'source flows' and 'channel flows'. Though, as yet, we do not have physical or operational interpretations for these 'flows', they serve as useful analytical devices for synthesizing converse expressions for SW. We find that source and channel flows for problem SW follow a hierarchy such that flows at the highest level satisfy the error density bottleneck, whereas the flows at the next levels have to meet a bottleneck, dictated by the flows at the level above, along various paths in the network. We show that the well-known information spectrum-based converse of Miyake and Kanaya [2] results from a particular way of constructing these flows. Improvements on this converse follow by synthesizing these flows in a more sophisticated manner. Specifically, by synthesizing flows for the networked problem using flows from the following point-to-point problems: (a) lossless source coding of jointly encoded correlated sources (S_1, S_2), (b) lossless source coding of S_1 with perfect side-information of S_2 available at the decoder, and (c) lossless source coding of S_2 with perfect side-information of S_1 at the decoder, we show that a new finite blocklength meta-converse results, which improves on the converse of Miyake and Kanaya.

The paper is organized as follows. In Section II, we consider the point-to-point lossy source coding problem with side-information. By the LP framework and an appropriate choice of source and channel flows, we derive new tight finite blocklength converses for these problems. In Section III, we discuss the extension of the LP relaxation to problem SW and establish the duality based framework. In Section IV, we illustrate how to synthesize new finite blocklength converses for SW from point-to-point sub-problems and present a new finite blocklength converse which improves on the converse of Miyake and Kanaya. Lastly, in Section V, we discuss the structure of the constraints of the dual program corresponding to SW and possible avenues for further strengthening of the bound.

A. Notation

Throughout this paper, we consider only discrete random variables. We make use of the following notation. Upper case letters A, B represent random variables taking values in finite spaces represented by calligraphic letters, $\mathcal{A}, \mathcal{B}$ respectively; lower case letters a, b represent the specific values these random variables take. $\mathbb{I}\{\cdot\}$ represents the indicator function which is equal to one when '$\cdot$' is true and is zero otherwise. $\mathcal{P}(\cdot)$ denotes the set of all probability distributions on '$\cdot$' and $Q \in \mathcal{P}(\cdot)$ represents a specific distribution. If Q is a joint probability distribution, let $Q_\bullet$ denote the marginal distribution of '$\bullet$'. For example, $Q_{X|S}$ represents the vector with $Q_{X|S}(x|s)$ for $x \in \mathcal{X}, s \in \mathcal{S}$ as its components. Let $P_{A|B}P_{C|D}(a,b,c,d)$ stand for $P_{A|B}(a|b)P_{C|D}(c|d)$. If P represents an optimization problem, then $\mathrm{OPT}(P)$ represents its optimal value and $\mathrm{FEA}(P)$ represents its feasible region. LHS stands for Left Hand Side and RHS stands for Right Hand Side. The notation $a \perp\!\!\!\perp b$ denotes that a is independent of b.

II. Finite Blocklength Point-to-Point Source Coding

In this section, we consider the point-to-point lossy source coding problem and the lossless source coding problem with side information at the decoder. We employ the LP relaxation framework to obtain finite blocklength converses for these problems.

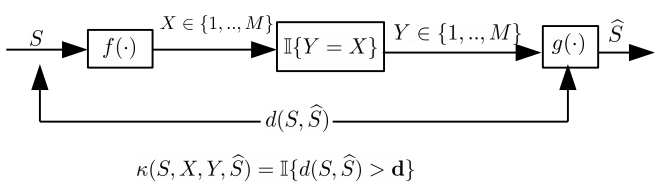

Fig. 2: Lossy Source Coding

A. Point-to-Point Lossy Source Coding

We begin with point-to-point lossy source coding. The finite blocklength lossy source coding problem (Figure 2) with probability of excess distortion as the loss criterion can be posed as the following optimization problem,

$$\begin{array}{ll} \text{SC} & \min\limits_{f,g} \ \mathbb{E}[\mathbb{I}\{d(S,\widehat{S}) > \mathbf{d}\}] \\ \text{s.t.} & X = f(S) \\ & \widehat{S} = g(Y). \end{array}$$

Here, S, X, Y and $\widehat{S}$ are discrete random variables taking values in fixed, finite spaces $\mathcal{S}, \mathcal{X}, \mathcal{Y}, \widehat{\mathcal{S}}$ respectively, with $\mathcal{X} = \mathcal{Y} = \{1, \ldots, M\}$, $M \in \mathbb{N}$ and $\mathcal{S} = \widehat{\mathcal{S}}$. S represents the source message distributed according to a known distribution P_S. The source message is encoded according to $f : \mathcal{S} \to \mathcal{X}$ to get the signal X which is transmitted across a deterministic channel with conditional probability distribution $P_{Y|X} = \mathbb{I}\{Y = X\}$. Y represents the channel output which is decoded according to $g : \mathcal{Y} \to \widehat{\mathcal{S}}$ to get the message $\widehat{S}$ at the destination. $d : \mathcal{S} \times \widehat{\mathcal{S}} \to [0, +\infty]$ represents the distortion measure and $\mathbf{d} \geq 0$ represents the distortion level. The optimization problem SC, thus, seeks to find a code (f, g) which minimizes $\mathbb{E}[\mathbb{I}\{d(S,\widehat{S}) > \mathbf{d}\}] = \mathbb{P}[d(S,\widehat{S}) > \mathbf{d}]$, the probability of excess distortion under the measure $\mathbb{P}$ induced by f, g.

SC can be posed equivalently as the following optimization problem over joint probability distributions,

$$\begin{array}{ll} \text{SC} & \min\limits_{Q, Q_{X|S}, Q_{\widehat{S}|Y}} \sum\limits_{s,\widehat{s}} \mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\} \sum\limits_{x,y} Q(s,x,y,\widehat{s}) \\ & Q(\widehat{z}) \equiv P_S Q_{X|S} P_{Y|X} Q_{\widehat{S}|Y}(\widehat{z}), \\ & \sum_x Q_{X|S}(x|s) = 1 \quad \forall s \in \mathcal{S}, \\ \text{s.t.} & \sum_{\widehat{s}} Q_{\widehat{S}|Y}(\widehat{s}|y) = 1 \quad \forall y \in \mathcal{Y}, \\ & Q_{X|S}(x|s) \geq 0 \quad \forall s \in \mathcal{S}, x \in \mathcal{X}, \\ & Q_{\widehat{S}|Y}(\widehat{s}|y) \geq 0 \quad \forall \widehat{s} \in \widehat{\mathcal{S}}, y \in \mathcal{Y}, \end{array}$$

where $\widehat{z} := (s, x, y, \widehat{s}) \in \widehat{\mathcal{Z}}$, $\widehat{\mathcal{Z}} := \mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \widehat{\mathcal{S}}$ and $P_S Q_{X|S} P_{Y|X} Q_{\widehat{S}|Y}(\widehat{z}) \equiv P_S(s) Q_{X|S}(x|s) P_{Y|X}(y|x) Q_{\widehat{S}|Y}(\widehat{s}|y)$. Here, $Q_{X|S}$ represents a randomized encoder and $Q_{\widehat{S}|Y}$ represents a randomized decoder. We refer the readers to [3] for details on this formulation.

To obtain lower bounds on the optimal value of SC, we adopt the LP relaxation detailed in [3]. Towards this, we introduce a new variable $W(s, x, y, \widehat{s}) \equiv Q_{X|S}(x|s) Q_{\widehat{S}|Y}(\widehat{s}|y)$ and obtain valid constraints involving W through the constraints of the problem. Specifically, multiply both sides of the constraint $\sum_x Q_{X|S}(x|s) \equiv 1$ by $Q_{\widehat{S}|Y}(\widehat{s}|y)$ for all $s, y, \widehat{s}$ and multiply both sides of $\sum_{\widehat{s}} Q_{\widehat{S}|Y}(\widehat{s}|y) \equiv 1$ by $Q_{X|S}(x|s)$ for all s, x, y. Replacing the bilinear product terms in the resulting set of constraints and in the objective of SC with W, gives new linear constraints in the variables $Q_{X|S}, Q_{\widehat{S}|Y}, W$, which together with $Q_{X|S} \in \mathcal{P}(\mathcal{X}|\mathcal{S})$ and $Q_{\widehat{S}|Y} \in \mathcal{P}(\widehat{\mathcal{S}}|\mathcal{Y})$ give the following LP relaxation.

$$\begin{array}{lll} \text{LP} & \min\limits_{Q_{X|S}, Q_{\widehat{S}|Y}, W} \sum\limits_{\widehat{z}} \mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\} P_S(s) P_{Y|X}(y|x) W(\widehat{z}) & \\ & \sum_x Q_{X|S}(x|s) = 1 & : \gamma^a(s) \quad \forall s \\ & \sum_{\widehat{s}} Q_{\widehat{S}|Y}(\widehat{s}|y) = 1 & : \gamma^b(y) \quad \forall y \\ \text{s.t.} & \sum_x W(\widehat{z}) - Q_{\widehat{S}|Y}(\widehat{s}|y) = 0 & : \lambda_{\mathsf{s}}(s,\widehat{s},y) \ \forall s, \widehat{s}, y \\ & \sum_{\widehat{s}} W(\widehat{z}) - Q_{X|S}(x|s) = 0 & : \lambda_{\mathsf{c}}(x,s,y) \ \forall x, s, y \\ & Q_{X|S}, Q_{\widehat{S}|Y}, W \geq 0. & \end{array}$$

Above $\gamma^a, \gamma^b, \lambda_{\mathsf{s}}$ and λ_{c} are Lagrange multipliers corresponding to the respective constraints.

1) An operational interpretation via optimal transport: The above LP relaxation can be explained operationally by relating it to the optimal transport problem [10]. Note that for each $s \in \mathcal{S}$ and $y \in \mathcal{Y}$, $W(s, \cdot, y, \cdot)$ is a *coupling* on $\mathcal{X} \times \widehat{\mathcal{S}}$ between the marginals $Q_{X|S}(\cdot|s)$ and $Q_{\widehat{S}|Y}(\cdot|y)$; let the set of such W be denoted by $\Xi(Q_{X|S}, Q_{\widehat{S}|Y})$. The LP relaxation of SC is a *nested* minimization – the inner minimization is over all couplings $W \in \Xi(Q_{X|S}, Q_{\widehat{S}|Y})$ and the outer minimization is over all randomized codes $(Q_{X|S}, Q_{\widehat{S}|Y})$:

$$\min_{Q_{X|S}, Q_{\widehat{S}|Y}} \ \min_{W \in \Xi(Q_{X|S}, Q_{\widehat{S}|Y})} \sum_{\widehat{z}} \mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\} P_S(s) P_{Y|X}(y|x) W(\widehat{z}).$$

The original problem SC has the outer minimization over codes, but in place of the inner minimization over W it employs the product $Q_{X|S}(\cdot|\cdot) Q_{\widehat{S}|Y}(\cdot|\cdot) \in \Xi(Q_{X|S}, Q_{\widehat{S}|Y})$ to obtain the distribution. Thus the LP relaxation is arrived at by considering the term $Q_{X|S}(\cdot|\cdot) Q_{\widehat{S}|Y}(\cdot|\cdot)$ in SC as an element of $\Xi(Q_{X|S}, Q_{\widehat{S}|Y})$ and minimizing the resulting cost over all elements of $\Xi(Q_{X|S}, Q_{\widehat{S}|Y})$. Operationally speaking, the LP relaxation seeks to design codes *and* couplings between them that minimize the error under the joint distribution induced by the coupling.

We caution the readers that for multiterminal problems, one must apply this interpretation with additional caveats. We discuss this in Section III-A1.

2) Duality and bounds: Employing the Lagrange multipliers corresponding to the constraints of LP, we obtain the following dual of LP,

$$\begin{array}{llll} \text{DP} & \max\limits_{\gamma^a, \gamma^b, \lambda_{\mathsf{s}}, \lambda_{\mathsf{c}}} \sum\limits_s \gamma^a(s) + \sum\limits_y \gamma^b(y) & & \\ & \gamma^a(s) - \sum_y \lambda_{\mathsf{c}}(x,s,y) \leq 0 & \forall x, s & \text{(P1)} \\ \text{s.t.} & \gamma^b(y) - \sum_s \lambda_{\mathsf{s}}(s,\widehat{s},y) \leq 0 & \forall \widehat{s}, y & \text{(P2)} \\ & \lambda_{\mathsf{s}}(s,\widehat{s},y) + \lambda_{\mathsf{c}}(x,s,y) \leq \Sigma(\widehat{z}) & \forall \widehat{z} & \text{(P3)} \end{array}$$

where $\Sigma(\widehat{z}) = \mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\} P_S(s) \mathbb{I}\{y = x\}$ for all $\widehat{z}$, since $P_{Y|X}(y|x) = \mathbb{I}\{y = x\}$.

In problem DP, it is optimal to choose $\gamma^a(s)$ and $\gamma^b(y)$ such that (P1) and (P2) hold with equality, i.e., $\gamma^a(s) \equiv \min_x \sum_y \lambda_{\mathsf{c}}(x,s,y)$ and $\gamma^b(y) \equiv \min_{\widehat{s}} \sum_s \lambda_{\mathsf{s}}(s,\widehat{s},y)$. Then

the optimal value of DP with $\Sigma(\widehat{z})$ as the RHS of (P3) evaluates to,

$$\begin{aligned}&\mathrm{OPT}(\mathrm{DP}, \Sigma(\widehat{z}))\\ &= \max_{\lambda_{\mathsf{s}}, \lambda_{\mathsf{c}}} \left\{ \sum_s \min_x \sum_y \lambda_{\mathsf{c}}(x,s,y) + \sum_y \min_{\widehat{s}} \sum_s \lambda_{\mathsf{s}}(s,\widehat{s},y) \right\}\\ &\text{s.t} \qquad \lambda_{\mathsf{c}}(x,s,y) + \lambda_{\mathsf{s}}(s,\widehat{s},y) \leq \Sigma(\widehat{z}) \quad \forall \widehat{z} \in \widehat{\mathcal{Z}}. \quad (1)\end{aligned}$$

It follows that if we construct functions $\lambda_{\mathsf{s}} : \mathcal{S} \times \widehat{\mathcal{S}} \times \mathcal{Y} \to \mathbb{R}$ and $\lambda_{\mathsf{c}} : \mathcal{S} \times \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ satisfying (1), then linear programming duality implies the following lower bound on $\mathrm{OPT}(\mathrm{SC})$,

$$\begin{aligned}&\mathrm{OPT}(\mathrm{SC}) \geq \mathrm{OPT}(\mathrm{LP}) = \mathrm{OPT}(\mathrm{DP})\\ &\geq \sum_s \min_x \sum_y \lambda_{\mathsf{c}}(x,s,y) + \sum_y \min_{\widehat{s}} \sum_s \lambda_{\mathsf{s}}(s,\widehat{s},y). \quad (2)\end{aligned}$$

Notice that λ_{s} and λ_{c} are functions on subspaces of $\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \widehat{\mathcal{S}}$. λ_{c} is a function of the source signal s, the channel input x and channel output y; we call this function a *channel flow*. On the other hand, λ_{s} is a function of the source signal s, the decoder input y and decoder output $\widehat{s}$. We refer to it as a *source flow*. Hence, for the point-to-point finite blocklength source coding problem, our LP-based framework reduces to constructing a *source flow* and a *channel flow* such that they satisfy the bottleneck imposed by the constraint (P3). The RHS of (P3) is the "error density", $\Sigma(\widehat{z}) = P_S(s)\mathbb{I}\{y = x\}\mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\}$, and hence the challenge is to *optimally pack* a source flow and a channel flow so as to not exceed the error density.

It was shown in [3] that by an appropriate construction of these source and channel flows, a new finite blocklength converse for lossy source coding results which improves on the tilted information based converse of Kostina and Verdú [5]. Modifying and generalizing this construction of flows, we now present a new metaconverse for lossy source coding, which *implies* our improvement on the Kostina-Verdú converse and the hypothesis testing based converse in [5, Theorem 8]. To the best of our knowledge, the metaconverse below is the strongest known.

Theorem 2.1 (Metaconverse for Lossy Source Coding): Consider problem SC. For any code,

$$\begin{aligned}&\mathbb{E}[\mathbb{I}\{d(S,\widehat{S}) > \mathbf{d}\}] \geq \mathrm{OPT}(\mathrm{SC}) \geq \mathrm{OPT}(\mathrm{LP}) = \mathrm{OPT}(\mathrm{DP})\\ &\geq \sup_{0 \leq \phi(s) \leq P_S(s)} \left\{ \sum_s \phi(s) - M \max_{\widehat{s}} \sum_s \phi(s)\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\} \right\}, \quad (3)\end{aligned}$$

where the supremum is over all functions $\phi : \mathcal{S} \to [0,1]$ such that $0 \leq \phi(s) \leq P_S(s)$ for all $s \in \mathcal{S}$.

Proof : Consider the following values of source and channel flows,

$$\begin{aligned}\lambda_{\mathsf{c}}(x,s,y) &\equiv \mathbb{I}\{y = x\}\phi(s), \quad (4)\\ \lambda_{\mathsf{s}}(s,\widehat{s},y) &\equiv -\phi(s)\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\}.\end{aligned}$$

We now check if the above choice of flows satisfy constraint (P3). For this, consider the following two cases.

Case 1: $\mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\} = 1$.

In this case, $\lambda_{\mathsf{s}}(s,\widehat{s},y) = 0$ and $\lambda_{\mathsf{c}}(x,s,y) = \mathbb{I}\{y = x\}\phi(s) \leq P_S(s)\mathbb{I}\{y = x\}$, which is the RHS of (P3).

Case 2: $\mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\} = 0$.

In this case, the RHS of (P3) is zero and LHS becomes, $\mathbb{I}\{y = x\}\phi(s) - \phi(s) \leq 0$, thereby satisfying (P3). Hence, the considered choice of flows satisfy constraint (P3). Consequently, the required lower bound follows from (2) by taking supremum over ϕ such that $0 \leq \phi(s) \leq P_S(s)\ \forall s \in \mathcal{S}$. ■

In particular, choosing $\phi(s) = \min\{P_S(s), z(s)\}$ in (3) where $z : \mathcal{S} \to [0,\infty)$, and taking the supremum over such z, we get the following bound,

$$\begin{aligned}&\mathbb{E}[\mathbb{I}\{d(S,\widehat{S}) > \mathbf{d}\}] \geq \mathrm{OPT}(\mathrm{SC}) \geq \mathrm{OPT}(\mathrm{LP})\\ &\geq \sup_{z \geq 0} \Bigg\{ \sum_s \min\{P_S(s), z(s)\}\\ &\quad - M \max_{\widehat{s}} \sum_s \min\{P_S(s), z(s)\}\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\} \Bigg\}. \quad (5)\end{aligned}$$

Remark II.1. (Choice of Flows) An easy way of motivating the choice of flows is as follows. Observe that if $0 \leq \phi(s) \leq P_S(s)$ for all s, we have that,

$$\begin{aligned}\Sigma(\widehat{z}) &\geq \phi(s)\mathbb{I}\{y = x\}\mathbb{I}\{d(s,\widehat{s}) > \mathbf{d}\}\\ &= \phi(s)\mathbb{I}\{y = x\} - \phi(s)\mathbb{I}\{y = x\}\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\}.\end{aligned}$$

An obvious choice of the flows would thus be $\lambda_{\mathsf{c}}(x,s,y) = \phi(s)\mathbb{I}\{y = x\}$ and $\lambda_{\mathsf{s}}(s,\widehat{s},y) \leq -\phi(s)\mathbb{I}\{y = x\}\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\}$, which results in our metaconverse in (3). □

The following results are corollaries to the metaconverse. $j_S(\cdot, \mathbf{d})$ below is the d-tilted information; we refer the reader to [5] for details.

Corollary 2.2: (Metaconverse Recovers Improvement on Kostina-Verdú Converse from [3, Corollary 5.9])

The following converse follows from (5):

$$\begin{aligned}&\mathbb{E}[\mathbb{I}\{d(S,\widehat{S}) > \mathbf{d}\}] \geq \mathrm{OPT}(\mathrm{SC}) \geq \mathrm{OPT}(\mathrm{LP}) = \mathrm{OPT}(\mathrm{DP})\\ &\geq \sup_{\gamma} \Bigg\{ \mathbb{P}[j_S(S,\mathbf{d}) \geq \gamma + \log M] + \frac{1}{M} \times\\ &\quad \sum_s P_S(s) \exp(j_S(s,\mathbf{d}) - \gamma)\mathbb{I}\{j_S(s,\mathbf{d}) < \log M + \gamma\}\\ &\quad - \max_{\widehat{s}} \exp(-\gamma) \sum_s P_S(s) \exp(j_S(s,\mathbf{d}))\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\} \Bigg\},\end{aligned}$$

which is the improvement on the Kostina-Verdú tilted information based converse in [3, Corollary 5.9].

Proof : To see this, take $z(s) = P_S(s)\frac{1}{M}\exp(j_S(s,\mathbf{d}) - \gamma)$ for any scalar γ and lower bound $-M \sup_{\widehat{s}} \sum_s \min\{P_S(s), z(s)\}\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\}$ with $-M \sup_{\widehat{s}} \sum_s z(s)\mathbb{I}\{d(s,\widehat{s}) \leq \mathbf{d}\}$. Subsequently, take supremum over γ to get the required bound. ■

Consider a binary hypothesis testing problem between distributions P_S and Q_S. Let $\alpha_\theta(P_S, Q_S)$ represent the minimum type-I error, $\sum_s P_S(s)T(s)$ over all tests T such that the type-II error, $\sum_s Q_S(s)(1 - T(s))$ is at most θ. The following corollary shows that the metaconverse in Theorem 2.1 in fact recovers the hypothesis testing based converse of Kostina and Verdú [5, Theorem 8].

Corollary 2.3: (Metaconverse Recovers Kostina-Verdú Hypothesis testing based Converse from [5, Theorem 8])
The following converse follows from (5),

$$\begin{aligned}&\mathbb{E}[\mathbb{I}\{d(S,\widehat{S})>\mathbf{d}\}]\geq \mathrm{OPT(SC)}\geq \mathrm{OPT(LP)}=\mathrm{OPT(DP)}\\&\overset{(a)}{\geq}\sup_{Q_S\in\mathcal{P}(\mathcal{S})}\sup_{\beta\geq 0}\Big\{\sum_s \min\{P_S(s),\beta Q_S(s)\}-\beta M^*\Big\}\\&\overset{(b)}{=}\sup_{Q_S\in\mathcal{P}(\mathcal{S})}\alpha_{M^*}(P_S,Q_S),\end{aligned}\tag{6}$$

which is equivalent to the hypothesis testing based converse of Kostina and Verdú in [5, Theorem 8]. Here, $M^*=M\max_{\widehat{s}}\mathbb{Q}[d(S,\widehat{s})\leq\mathbf{d}]$ and $\mathbb{Q}$ is the measure on $\mathcal{S}$ induced by Q_S.

Proof : To recover the converse in (a) from (5), take $z(s)=\beta Q_S(s)$ where $\beta\geq 0$ and lower bound $-M\sup_{\widehat{s}}\sum_s\min\{P_S(s),z(s)\}\mathbb{I}\{d(s,\widehat{s})\leq\mathbf{d}\}$ with $-M\sup_{\widehat{s}}\sum_s z(s)\mathbb{I}\{d(s,\widehat{s})\leq\mathbf{d}\}=-\beta M^*$. Subsequently, take the supremum over $\beta\geq 0$ and $Q_S\in\mathcal{P}(\mathcal{S})$ to get the required bound. The proof of the relation in (b) is included in Corollary A.1 in Appendix A. ■

Corollary 2.4: (Metaconverse Recovers Palzer-Timo Converse [11, Theorem 1])
The following converse follows from (5),

$$\begin{aligned}&\mathbb{E}[\mathbb{I}\{d(S,\widehat{S})>\mathbf{d}\}]\geq \mathrm{OPT(SC)}\geq \mathrm{OPT(LP)}=\mathrm{OPT(DP)}\\&\geq\sup_{\beta\in\mathbb{R}}\Big\{\mathbb{P}[j_S(S,\mathbf{d})\geq\beta]\\&-M\max_{\widehat{s}}\mathbb{P}[j_S(S,\mathbf{d})\geq\beta,d(S,\widehat{s})\leq\mathbf{d}]\Big\},\end{aligned}$$

which is the converse of Palzer and Timo in [11, Theorem 1].

Proof : To obtain this converse from (5), take $z(s)=P_S(s)\mathbb{I}\{j_S(s,\mathbf{d})\geq\beta\}$ where $\beta\geq 0$ in (3) and take supremum over $\beta\geq 0$. Notice that in this case, $\min\{z(s),P_S(s)\}=z(s)$ since $z(s)=P_S(s)\mathbb{I}\{j_S(s,\mathbf{d})\geq\beta\}\leq P_S(s)$. ■

The finite blocklength lossless data compression problem results from SC by setting $d(S,\widehat{S})=\mathbb{I}\{S\neq\widehat{S}\}$ and $\mathbf{d}=0$. The following corollary particularizes the metaconverse in (3) to lossless data compression. In this case, the metaconverse takes a particularly simple form.

Corollary 2.5 (Metaconverse for Lossless Source Coding): For lossless data compression, consider the setting of Theorem 2.1 with $d(S,\widehat{S})=\mathbb{I}\{S\neq\widehat{S}\}$ and $\mathbf{d}=0$. Consequently, for any code, the following converse follows from Theorem 2.1,

$$\begin{aligned}&\mathbb{E}[\mathbb{I}\{S\neq\widehat{S}\}]\geq\mathrm{OPT(SC)}\geq\mathrm{OPT(LP)}=\mathrm{OPT(DP)}\\&\geq\sup_{0\leq\phi\leq P_S}\Big\{\|\phi\|_1-M\|\phi\|_\infty\Big\}.\end{aligned}\tag{7}$$

In the above converse we have viewed ϕ as a vector in $\mathbb{R}^{|\mathcal{S}|}$ that is nonnegative and dominated by $P_S\in\mathbb{R}^{|\mathcal{S}|}$. The maximization in (7) is a tradeoff between increasing the ℓ_1 norm of ϕ on the one hand, and decreasing the ℓ_∞ norm of ϕ on the other. One plausible strategy for this tradeoff is to take $\phi(s)=P_S(s)$ for those $s\in\mathcal{S}$ for which $P_S(s)$ is not too large, and zero otherwise. Specifically, one may take $\phi(s)=P_S(s)\mathbb{I}\{P_S(s)\leq\frac{\exp(-\gamma)}{M}\}$ for some $\gamma\geq 0$. Then the RHS of (7) is lower bounded by

$$\sup_{\gamma\geq 0}\big\{\mathbb{P}[h(S)\geq\log M+\gamma]-\exp(-\gamma)\big\},$$

where $h(S)=-\log P_S(S)$. The above converse is [5, Theorem 7] specialized to the lossless case.

Having outlined the LP based framework for point-to-point lossless source coding, we now consider three problems that will serve as sub-problems for analysing problem SW.

3) Lossless Coding of Jointly Encoded Correlated Sources (S_1,S_2): In this sub-problem of Slepian-Wolf coding problem, the correlated sources S_1,S_2 are *jointly encoded* by $f:\mathcal{S}_1\times\mathcal{S}_2\to\mathcal{X}_1\times\mathcal{X}_2$ to get (X_1,X_2). (X_1,X_2) is sent through the channel $P_{Y_1,Y_2|X_1,X_2}=\mathbb{I}\{(Y_1,Y_2)=(X_1,X_2)\}$ to get (Y_1,Y_2) which is then decoded according to $g:\mathcal{Y}_1\times\mathcal{Y}_2\to\widehat{\mathcal{S}}_1\times\widehat{\mathcal{S}}_2$. The objective, as for SW problem, is to losslessly recover (S_1,S_2) at the destination.

It is easy to see that the above joint encoding problem is equivalent to the point-to-point lossless source coding problem SC with $S:=(S_1,S_2)$, $X:=(X_1,X_2)$, $Y:=(Y_1,Y_2)$, $\widehat{S}:=(\widehat{S}_1,\widehat{S}_2)$ and $d(S,\widehat{S})=\mathbb{I}\{S\neq\widehat{S}\}$ with $\mathbf{d}=0$. Consequently, to obtain finite blocklength converses for the joint encoding problem of correlated sources, we resort to the following generalized version of DP for lossless source coding problem,

$$\boxed{\begin{aligned}&\text{DPJE}\quad\max_{\widehat{\gamma}^a,\widehat{\gamma}^b,\widehat{\lambda}_{\mathsf{s}},\widehat{\lambda}_{\mathsf{c}}}\sum_{s_1,s_2}\widehat{\gamma}^a(s_1,s_2)+\sum_{y_1,y_2}\widehat{\gamma}^b(y_1,y_2)\\&\text{s.t.}\quad\begin{aligned}&\widehat{\gamma}^a(s_1,s_2)-\textstyle\sum_{y_1,y_2}\widehat{\lambda}_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2)\leq 0\quad\forall x_1,x_2,s_1,s_2\ \text{(A1)}\\&\widehat{\gamma}^b(y_1,y_2)-\textstyle\sum_{s_1,s_2}\widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)\leq 0\quad\forall\widehat{s}_1,\widehat{s}_2,y_1,y_2\ \text{(A2)}\\&\widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)+\widehat{\lambda}_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2)\leq\Upsilon(\widetilde{z})\quad\forall\widetilde{z}\ \text{(A3)}\end{aligned}\end{aligned}}$$

where $\widetilde{z}:=(s_1,s_2,x_1,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2)$, $\Upsilon(\widetilde{z})=\mathbb{I}\{(s_1,s_2)\neq(\widehat{s}_1,\widehat{s}_2)\}\ P_{S_1,S_2}(s_1,s_2)\mathbb{I}\{(y_1,y_2)=(x_1,x_2)\}$ for all $\widetilde{z}$. Though this is a straightforward generalization of DP, we will need this later and hence we have included it here.

As in the case of DP, taking $\widehat{\gamma}^a$ and $\widehat{\gamma}^b$ such that (A1) and (A2) hold with equality, OPT(DPJE) can be written in terms of the channel flow $\widehat{\lambda}_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2)$ and source flow $\widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)$. The metaconverse for lossless source coding problem in Corollary 2.5 then readily implies the following corollary.

Corollary 2.6 (Metaconverse for Jointly Encoded Sources): Consider problem SC with $S:=(S_1,S_2)$, $X:=(X_1,X_2)$, $Y:=(Y_1,Y_2)$, $\widehat{S}:=(\widehat{S}_1,\widehat{S}_2)$ and $d(S,\widehat{S})=\mathbb{I}\{S\neq\widehat{S}\}$ with $\mathbf{d}=0$. Consequently for any code, we have from Corollary 2.5,

$$\begin{aligned}&\mathbb{E}[\mathbb{I}\{(S_1,S_2)\neq(\widehat{S}_1,\widehat{S}_2)\}]\geq\mathrm{OPT(DPJE)}\\&\geq\sup_{0\leq\widehat{\phi}(s_1,s_2)\leq P_{S_1,S_2}(s_1,s_2)}\Big\{\sum_{s_1,s_2}\widehat{\phi}(s_1,s_2)\\&\qquad-M_1M_2\max_{\widehat{s}_1,\widehat{s}_2}\widehat{\phi}(\widehat{s}_1,\widehat{s}_2)\Big\},\end{aligned}\tag{8}$$

where the supremum is over $\widehat{\phi} : \mathcal{S}_1 \times \mathcal{S}_2 \to [0,1]$ such that $0 \leq \widehat{\phi}(s_1, s_2) \leq P_{S_1,S_2}(s_1, s_2)$ for all $s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2$.

Proof : To obtain the required converse, we consider the following choice for the flows in DPJE, which generalizes the ones adopted in (4).

$$\widehat{\lambda}_{\mathsf{c}}(x_1, x_2, s_1, s_2, y_1, y_2) \equiv \mathbb{I}\{(y_1, y_2) = (x_1, x_2)\}\widehat{\phi}(s_1, s_2) \quad (9)$$
$$\widehat{\lambda}_{\mathsf{s}}(s_1, s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) \equiv -\widehat{\phi}(s_1, s_2)\mathbb{I}\{(s_1, s_2) = (\widehat{s}_1, \widehat{s}_2)\},$$

The feasibility of these flows with respect to (A3) can be verified as in the proof of Theorem 2.1 and we skip the proof here. ■

B. Lossless Source Coding of S_1 with S_2 as the Side-Information

We now consider the following sub-problem of Slepian-Wolf coding: S_1 is to be recovered losslessly at the destination with S_2 available as side-information at the decoder (Figure 3). Towards this, S_1 is encoded according to $f_1 : \mathcal{S}_1 \to \mathcal{X}_1$ to get X_1, which is transmitted through the channel $P_{Y_1|X_1} = \mathbb{I}\{Y_1 = X_1\}$ to get Y_1. S_2 is the side information available at the decoder which decodes according to $g : \mathcal{S}_2 \times \mathcal{Y}_1 \to \widehat{\mathcal{S}}_1$ to get $\widehat{S}_1$. The finite blocklength source coding of S_1 given S_2 as the side information can be then posed as the following optimization problem,

$$\begin{array}{ll} \mathrm{SID}_{1|2} & \min\limits_{f_1, g} \mathbb{E}[\mathbb{I}\{S_1 \neq \widehat{S}_1\}] \\ \text{s.t.} & X_1 = f_1(S_1), \quad \widehat{S}_1 = g(S_2, Y_1). \end{array}$$

Thus, $\mathrm{SID}_{1|2}$ seeks to obtain a code (f_1, g) which minimizes $\mathbb{E}[\mathbb{I}\{S_1 \neq \widehat{S}_1\}] = \mathbb{P}[S_1 \neq \widehat{S}_1]$, the average probability of error.

To obtain finite blocklength converses, we employ the LP relaxation approach in [3] to obtain the following LP relaxation of the problem $\mathrm{SID}_{1|2}$.

$$\begin{array}{ll} \mathrm{LPSI}_{1|2} & \min\limits_{Q_{X_1|S_1}, Q_{\widehat{S}_1|Y_1,S_2}, W} \sum\limits_{\bar{z}} \Psi(\bar{z}) W(\bar{z}) \\ & \sum_{x_1} Q_{X_1|S_1}(x_1|s_1) \equiv 1 : \bar{\gamma}^a(s_1) \\ & \sum_{\widehat{s}_1} Q_{\widehat{S}_1|Y_1,S_2}(\widehat{s}_1|y_1, s_2) \equiv 1 : \bar{\gamma}^b(y_1, s_2) \\ \text{s.t.} & \sum_{x_1} W(\bar{z}) - Q_{\widehat{S}_1|Y_1,S_2}(\widehat{s}_1|y_1, s_2) = 0 : \bar{\lambda}_{\mathsf{s}}^{(1|2)}(\mathbf{s}, \widehat{s}_1, y_1) \\ & \sum_{\widehat{s}_1} W(\bar{z}) - Q_{X_1|S_1}(x_1|s_1) \equiv 0 : \bar{\lambda}_{\mathsf{c}}^{(1|2)}(\mathbf{s}, x_1, y_1) \\ & Q_{X_1|S_1}, Q_{\widehat{S}_1|Y_1,S_2}, W \geq 0, \end{array}$$

where $\mathbf{s} := (s_1, s_2)$, $\bar{z} := (x_1, s_1, s_2, \widehat{s}_1, y_1)$ and $\Psi(\bar{z}) = P_{S_1,S_2}(s_1, s_2)\mathbb{I}\{s_1 \neq \widehat{s}_1\}\mathbb{I}\{y_1 = x_1\}$.

Employing the Lagrange multipliers $\bar{\gamma}^a, \bar{\gamma}^b, \bar{\lambda}_{\mathsf{s}}^{(1|2)}$ and $\bar{\lambda}_{\mathsf{c}}^{(1|2)}$ corresponding to the constraints of $\mathrm{LPSI}_{1|2}$, we obtain the following dual of $\mathrm{LPSI}_{1|2}$.

$$\begin{array}{ll} \mathrm{DPSI}_{1|2} & \max\limits_{\bar{\gamma}^a, \bar{\gamma}^b, \bar{\lambda}_{\mathsf{s}}^{(1|2)}, \bar{\lambda}_{\mathsf{c}}^{(1|2)}} \sum\limits_{s_1} \bar{\gamma}^a(s_1) + \sum\limits_{y_1, s_2} \bar{\gamma}^b(y_1, s_2) \\ & \bar{\gamma}^a(s_1) - \sum_{y_1,s_2} \bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1, s_2, x_1, y_1) \leq 0 \quad \forall x_1, s_1 \quad \text{(B1)} \\ \text{s.t.} & \bar{\gamma}^b(s_2, y_1) - \sum_{s_1} \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1, s_2, \widehat{s}_1, y_1) \leq 0 \quad \forall s_2, \widehat{s}_1, y_1 \quad \text{(B2)} \\ & \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1, s_2, \widehat{s}_1, y_1) + \bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1, s_2, x_1, y_1) \leq \Psi(\bar{z}) \quad \forall \bar{z} \quad \text{(B3)} \end{array}$$

Choosing $\bar{\gamma}^a(s_1)$ and $\bar{\gamma}^b(s_2, y_1)$ such that (B1) and (B2) hold with equality, $\mathrm{OPT}(\mathrm{DPSI}_{1|2})$ can be written in terms of $\bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1, s_2, x_1, y_1)$ and $\bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1, s_2, \widehat{s}_1, y_1)$. Notice that $\bar{\lambda}_{\mathsf{c}}^{(1|2)}$ is a function of x_1, s_1, s_2, y_1. Thus, for each $s_2 \in \mathcal{S}_2$, it is akin to a channel flow of the point-to-point source coding problem with S_1 as the source. Likewise, for each $s_2 \in \mathcal{S}_2$, $\bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1, s_2, \widehat{s}_1, y_1)$ is akin to a *source flow* for this problem. Following these observations, we now show that an appropriate construction of these source and channel flows results in the following finite blocklength converse for $\mathrm{SID}_{1|2}$.

Theorem 2.7: Consider the problem $\mathrm{SID}_{1|2}$. For any code, the following lower bound holds,

$$\begin{aligned} &\mathbb{E}[\mathbb{I}\{S_1 \neq \widehat{S}_1\}] \geq \mathrm{OPT}(\mathrm{SID}_{1|2}) \geq \\ &\mathrm{OPT}(\mathrm{LPSI}_{1|2}) \geq \sup_{0 \leq \phi^{(1|2)} \leq P_{S_1,S_2}} \Bigg\{ \sum_{s_1,s_2} \phi^{(1|2)}(s_1, s_2) \\ &\qquad - M_1 \sum_{s_2} \max_{\widehat{s}_1} \phi^{(1|2)}(\widehat{s}_1, s_2) \Bigg\}, \end{aligned} \quad (10)$$

where the supremum is over $\phi^{(1|2)} : \mathcal{S}_1 \times \mathcal{S}_2 \to [0,1]$ such that $\phi^{(1|2)}(s_1, s_2) \leq P_{S_1,S_2}(s_1, s_2)$ for all $s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2$.

Proof : To obtain the required converse, we consider the following values for the source flow and channel flow,

$$\begin{aligned} \bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1, s_2, x_1, y_1) &= \mathbb{I}\{x_1 = y_1\}\phi^{(1|2)}(s_1, s_2), \\ \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1, s_2, \widehat{s}_1, y_1) &= -\phi^{(1|2)}(s_1, s_2)\mathbb{I}\{s_1 = \widehat{s}_1\}. \end{aligned} \quad (11)$$

The feasibility of these flows with respect to (B3) can be verified as in the proof of Theorem 2.1. Consequently, employing linear programming duality and taking supremum over $\phi^{(1|2)}$ gives the required bound. ■

Notice that choosing $\phi^{(1|2)}(s_1, s_2) = \min\{P_{S_1,S_2}(s_1, s_2), \eta_2(s_1, s_2)\}$, $\eta_2 : \mathcal{S}_1 \times \mathcal{S}_2 \to [0,1]$, yields the following bound, $\mathbb{E}[\mathbb{I}\{S_1 \neq \widehat{S}_1\}] \geq \mathrm{OPT}(\mathrm{SID}_{1|2}) \geq$

$$\begin{aligned} &\mathrm{OPT}(\mathrm{LPSI}_{1|2}) \geq \sup_{\eta_2 \geq 0} \Bigg\{ \sum_{s_1,s_2} \min\{P_{S_1,S_2}(s_1, s_2), \eta_2(s_1, s_2)\} \\ &- M_1 \sum_{s_2} \max_{\widehat{s}_1} \min\{P_{S_1,S_2}(s_1, s_2), \eta_2(s_1, s_2)\} \Bigg\}. \end{aligned} \quad (12)$$

When particularized to $\eta_2(s_1, s_2) = P_{S_2}(s_2)\frac{\exp(-\beta)}{M_1}$ for $\beta \geq 0$, where $P_{S_2}(s_2) = \sum_{s_1} P_{S_1,S_2}(s_1, s_2)$ and taking supremum over β, the converse in (12) becomes,

$$\begin{aligned} &\mathbb{E}[\mathbb{I}\{S_1 \neq \widehat{S}_1\}] \geq \mathrm{OPT}(\mathrm{LPSI}_{1|2}) \geq \mathrm{OPT}(\mathrm{DPSI}_{1|2}) \\ &\overset{(a)}{\geq} \sup_{\beta \geq 0} \Bigg\{ \mathbb{P}[h_{S_1|S_2}(S_1|S_2) \geq \log M_1 + \beta] \\ &+ \frac{\exp(-\beta)}{M_1} \sum_{s_1,s_2} P_{S_2}(s_2)\mathbb{I}\{h_{S_1|S_2}(s_1|s_2) < \log M_1 + \beta\} \\ &- M_1 \sum_{s_2} \sup_{\widehat{s}_1} \min\Bigg\{ P_{S_1,S_2}(\widehat{s}_1, s_2), P(s_2)\frac{\exp(-\beta)}{M_1} \Bigg\} \Bigg\} \quad (13) \\ &\overset{(b)}{\geq} \sup_{\beta \geq 0} \Bigg\{ \mathbb{P}[h_{S_1|S_2}(S_1|S_2) \geq \log M_1 + \beta] - \exp(-\beta) \Bigg\}. \quad (14) \end{aligned}$$

Here $h_{A|B}(a|b) \equiv -\log P_{A|B}(a|b)$ is the conditional entropy density. The inequality in (a) follows from the definition

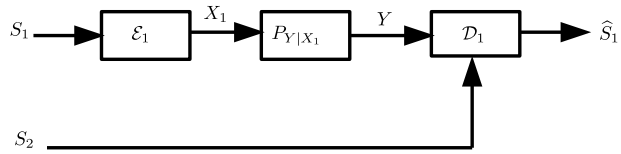

Fig. 3: Lossless source coding with side information available only at the decoder

of conditional entropy density. The inequality in (b) follows by lower bounding the non-negative term corresponding to $\mathbb{I}\{h_{S_1|S_2}(s_1|s_2) < \log M_1 + \beta\}$ in (13) by zero and upper bounding $\min\{P_{S_1,S_2}(\widehat{s}_1, s_2), P_{S_2}(s_2)\frac{\exp(-\beta)}{M_1}\}$ with $P_{S_2}(s_2)\frac{\exp(-\beta)}{M_1}$. Notice that the converse in (14) is the well-known converse for lossless source-coding problem with side-information at the decoder. The converse in (13) provides a new improvement on the standard converse.

C. Lossless Source Coding of S_2 with S_1 as the Side-Information

Analogous to $\mathrm{SID}_{1|2}$, the finite blocklength lossless source coding problem of S_2 with S_1 as the side information can be posed as the following optimization problem,

$$\begin{array}{ll} \mathrm{SID}_{2|1} & \min\limits_{f,g} \mathbb{E}[\mathbb{I}\{S_2 \neq \widehat{S}_2\}] \\ \text{s.t.} & \begin{array}{l} X_2 = f_2(S_2) \\ \widehat{S}_2 = g(S_1, Y_2). \end{array} \end{array}$$

The dual of the corresponding LP relaxation of $\mathrm{SID}_{2|1}$ then becomes,

$$\begin{array}{ll} \mathrm{DPSI}_{2|1} & \max\limits_{\tilde{\gamma}^a, \tilde{\gamma}^b, \tilde{\lambda}_{\mathsf{s}}^{(2|1)}, \tilde{\lambda}_{\mathsf{c}}^{(2|1)}} \sum\limits_{s_2} \tilde{\gamma}^a(s_2) + \sum\limits_{y_2, s_1} \tilde{\gamma}^b(y_2, s_1) \\ \text{s.t.} & \begin{array}{l} \tilde{\gamma}^a(s_2) - \sum_{y_2,s_1} \tilde{\lambda}_{\mathsf{c}}^{(2|1)}(s_1, s_2, x_2, y_2) \leq 0 \\ \qquad \forall x_2, s_2 \quad (C1) \\ \tilde{\gamma}^b(s_1, y_2) - \sum_{s_2} \tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1, s_2, \widehat{s}_2, y_2) \leq 0 \\ \qquad \forall s_1, \widehat{s}_2, y_2 \quad (C2) \\ \tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1, s_2, \widehat{s}_2, y_2) + \tilde{\lambda}_{\mathsf{c}}^{(2|1)}(s_1, s_2, x_2, y_2) \leq \Delta(z') \\ \qquad \forall z' \quad (C3), \end{array} \end{array}$$

where $z' := (s_1, s_2, x_2, y_2, \widehat{s}_2)$, and $\Delta(z') = P_{S_1,S_2}(s_1, s_2)\mathbb{I}\{y_2 = x_2\}\mathbb{I}\{s_2 \neq \widehat{s}_2\}$. The following choice of source flow and channel flow results in a converse similar to the one in Theorem 2.7 for the problem $\mathrm{SID}_{2|1}$.

$$\begin{aligned} \tilde{\lambda}_{\mathsf{c}}^{(2|1)}(s_1, s_2, x_2, y_2) &= \phi^{(2|1)}(s_1, s_2)\mathbb{I}\{x_2 = y_2\} \\ \tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1, s_2, \widehat{s}_2, y_2) &= -\phi^{(2|1)}(s_1, s_2)\mathbb{I}\{s_2 = \widehat{s}_2\}, \end{aligned} \tag{15}$$

where $\phi^{(2|1)} : \mathcal{S}_1 \times \mathcal{S}_2 \to [0, 1]$ is such that $0 \leq \phi^{(2|1)}(s_1, s_2) \leq P_{S_1,S_2}(s_1, s_2)$ for all $s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2$.

Theorem 2.8: Consider the problem $\mathrm{SID}_{2|1}$. Consequently, for any code, the following lower bound holds,

$$\begin{aligned} &\mathbb{E}[\mathbb{I}\{S_2 \neq \widehat{S}_2\}] \geq \mathrm{OPT}(\mathrm{SID}_{2|1}) \geq \mathrm{OPT}(\mathrm{DPSI}_{2|1}) \\ &\geq \sup_{0 \leq \phi^{(2|1)}(s_1,s_2) \leq P_{S_1,S_2}(s_1,s_2)} \Bigg\{ \sum_{s_1,s_2} \phi^{(2|1)}(s_1, s_2) \\ &\qquad - M_2 \sum_{s_1} \max_{\widehat{s}_2} \phi^{(2|1)}(s_1, \widehat{s}_2) \Bigg\}. \end{aligned} \tag{16}$$

In the next section, we extend the LP based framework to finite blocklength Slepian-Wolf coding problem and establish the duality based framework.

III. Linear Programming Based Framework for the Slepian-Wolf Problem

In this section, we discuss the extension of the linear programming (LP) based framework in Section II to the finite blocklength Slepian-Wolf coding problem SW. Towards this, consider the joint probability distribution $Q : \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}_1 \times \mathcal{Y}_2 \times \widehat{\mathcal{S}}_1 \times \widehat{\mathcal{S}}_2 \to [0, 1]$ which can be factored as,

$$Q(z) \equiv P_{S_1,S_2} Q_{X_1|S_1} Q_{X_2|S_2} P_{Y_1Y_2|X_1X_2} Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(z),$$

where $\mathcal{Z} := \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}_1 \times \mathcal{Y}_2 \times \widehat{\mathcal{S}}_1 \times \widehat{\mathcal{S}}_2$ and $z := (s_1, s_2, x_1, x_2, y_1, y_2, \widehat{s}_1, \widehat{s}_2) \in \mathcal{Z}$. Employing Q, we obtain the following optimization problem over joint probability distributions,

$$\begin{array}{ll} \mathrm{SW} & \min\limits_{\substack{Q_{X_1|S_1}, Q_{X_2|S_2}, \\ Q, Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}}} \sum\limits_{z \in \mathcal{Z}} \mathbb{I}\{(s_1, s_2) \neq (\widehat{s}_1, \widehat{s}_2)\} Q(z) \\ \text{s.t.} & \begin{array}{r} P_{S_1S_2} Q_{X_1|S_1} Q_{X_2|S_2} P_{Y_1,Y_2|X_1,X_2} Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(z) \equiv Q(z), \\ Q_{X_1|S_1} \in \mathcal{P}(\mathcal{X}_1|\mathcal{S}_1), \\ Q_{X_2|S_2} \in \mathcal{P}(\mathcal{X}_2|\mathcal{S}_2), \\ Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2} \in \mathcal{P}(\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2|\mathcal{Y}_1, \mathcal{Y}_2), \end{array} \end{array}$$

where, $\mathcal{P}(A|B) := \{Q_{A|B} | \sum_a Q_{A|B}(a|b) = 1, Q_{A|B}(a|b) \geq 0, \forall\, a, b\}$. Here, $Q_{X_1|S_1}$ and $Q_{X_2|S_2}$ represent the two randomized encoders, and $Q_{\widehat{S}_1,\widehat{S}|Y_1,Y_2}$ represents a randomized decoder. It is easy to argue as in [3] that the above formulation is in fact equivalent to problem SW stated in the introduction.

As in the case of the point-to-point problems, the presence of the multilinear constraint renders the feasible region of SW nonconvex. Notice that the degree of the multilinear term is three since there are three decision makers, whereas in the point-to-point problems the degree was two. To obtain converses or lower bounds on the optimal value of SW, we will again derive a linear programming (LP) relaxation of the nonconvex feasible region of SW, as shown in the next section.

A. LP Relaxation

For obtaining a linear programming relaxation of SW, we resort to the "lift-and-project" technique in integer programming. Towards this, we define the following new variables,

$$W(z) \equiv Q_{X_1|S_1} Q_{X_2|S_2} Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(z), \tag{17}$$

$$U(z_1) \equiv Q_{X_1|S_1}(x_1|s_1) Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1, \widehat{s}_2|y_1, y_2), \tag{18}$$

$$V(z_2) \equiv Q_{X_2|S_2}(x_2|s_2) Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1, \widehat{s}_2|y_1, y_2), \tag{19}$$

$$T(z_3) \equiv Q_{X_1|S_1}(x_1|s_1) Q_{X_2|S_2}(x_2|s_2), \tag{20}$$

where recall that $z := (s_1, s_2, x_1, x_2, y_1, y_2, \widehat{s}_1, \widehat{s}_2)$, $z_1 := (s_1, x_1, y_1, y_2, \widehat{s}_1, \widehat{s}_2)$, $z_2 := (s_2, x_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2)$ and $z_3 := (s_1, s_2, x_1, x_2)$. Using these variables, we first lift the problem SW to a higher dimensional space and then impose additional valid constraints involving these new variables. To obtain these constraints, we adopt the following procedure.

For each $s_1 \in \mathcal{S}_1$, we multiply both sides of the constraint $\sum_{x_1} Q_{X_1|S_1}(x_1|s_1) = 1$ by $Q_{X_2|S_2}(x_2|s_2)$ for all x_2, s_2, by $Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1, \widehat{s}_2|y_1, y_2)$ for all $\widehat{s}_1, \widehat{s}_2, y_1, y_2$ and by $Q_{X_2|S_2}(x_2|s_2)Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1, \widehat{s}_2|y_1, y_2)$ for all $x_2, s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2$, to obtain three new sets of multilinear equality constraints. Replace the resulting multilinear product terms by the newly defined variables in (17)–(20). This results in the following new valid linear constraints in the lifted space,

$$\sum_{x_1} T(x_1, x_2, s_1, s_2) \equiv Q_{X_2|S_2}(x_2|s_2)$$
$$\sum_{x_1} U(x_1, s_1, \widehat{s}_1, \widehat{s}_2, y_1, y_2) \equiv Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1, \widehat{s}_2|y_1, y_2)$$
$$\sum_{x_1} W(z) \equiv V(x_2, s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2).$$

Similar set of linear constraints can be obtained corresponding to $\sum_{x_2} Q_{X_2|S_2}(x_2|s_2) = 1$ for all s_2 and $\sum_{\widehat{s}_1,\widehat{s}_2} Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1, \widehat{s}_2|y_1, y_2) = 1$ for all y_1, y_2. Subsequently, add these new sets of linear constraints to the original constraints of SW. Further, replace Q in the objective function of SW with the first constraint written in terms of W, drop the multilinear equalities in (17)-(20) and we have the LP relaxation of SW, LPSW as given in the next page. Notice that the constraints of LPSW are implied by the constraints of SW whereby LPSW is a relaxation of SW.

Here, $\gamma^a(s_1)$, $\gamma^b(s_2)$, $\gamma^c(y_1, y_2)$, $\mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2)$, $\mu_{\mathsf{s}}^{(1)}(s_1, \widehat{s}_1, \widehat{s}_2, y_1, y_2)$, $\lambda_{\mathsf{s}}^{(2|1)}(s_1, s_2, x_2, y_1, y_2, \widehat{s}_1, \widehat{s}_2)$, $\mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2)$, $\mu_{\mathsf{s}}^{(2)}(s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2)$, $\lambda_{\mathsf{s}}^{(1|2)}(s_1, s_2, x_1, y_1, y_2, \widehat{s}_1, \widehat{s}_2)$, $\lambda_{\mathsf{c}}(s_1, s_2, x_1, x_2, y_1, y_2)$, $\mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2)$ and $\mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2)$ represent the Lagrange multipliers corresponding to the constraints of LPSW in that order.

1) An stronger optimal transport interpretation: The LP relaxation of SW, LPSW, also admits an interpretation via a "multiterminal" optimal transport problem. As in the point-to-point LP relaxation, we note that for each $s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, y_1 \in \mathcal{Y}_1, y_2 \in \mathcal{Y}_2$, $W(s_1, s_2, \cdot, \cdot, y_1, y_2, \cdot, \cdot)$ is a coupling on $\mathcal{X}_1 \times \mathcal{X}_2 \times (\widehat{\mathcal{S}}_1 \times \widehat{\mathcal{S}}_2)$ between all three marginals $Q_{X_1|S_1}(\cdot|s_1)$, $Q_{X_2|S_2}(\cdot|s_2)$ and $Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\cdot, \cdot|y_1, y_2)$; denote the set of all such W by Ξ'. However, note that $W \in \Xi'$ does not automatically imply that, for instance,

$$\sum_{x_1} W(z) \in \Xi(Q_{X_2|S_2}, Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}), \tag{21}$$

and likewise, that

$$\sum_{x_2} W(z) \in \Xi(Q_{X_1|S_1}, Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}), \tag{22}$$

$$\sum_{\widehat{s}_1,\widehat{s}_2} W(z) \in \Xi(Q_{X_1|S_1}, Q_{X_2|S_2}). \tag{23}$$

LPSW is obtained by imposing not only that $W \in \Xi'$, but also (21)-(23). Skipping the latter requirements would evidently lead to a looser relaxation which would perhaps not suffice for our purpose of obtaining tight converses. As in the point-to-point problem, LPSW is a nested minimization where the relaxation arises from replacing the product of kernels of randomized codes by any coupling in Ξ' that is constrained by (21)-(23), and then minimizing over all codes.

We note that the variables U, V, T in (18)-(20) are introduced in LPSW only to express the constraints (21)-(23) on W in a clearer manner. They could be eliminated in a straightforward manner and the entire problem could be expressed only in terms of W and the randomized code $(Q_{X_1|S_1}, Q_{X_2|S_2}, Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2})$.

B. Duality and Converses

Employing the Lagrange multipliers corresponding to the constraints of LPSW, we now obtain the dual of LPSW, denoted DPSW, and shown on the next page. Here, $\Pi(z) \equiv \mathbb{I}\{(s_1, s_2) \neq (\widehat{s}_1, \widehat{s}_2)\} \times P_{S_1,S_2}(s_1, s_2)\ \mathbb{I}\{(y_1, y_2) = (x_1, x_2)\}$. Let $\Theta := (\lambda_{\mathsf{s}}^{(1|2)}, \lambda_{\mathsf{s}}^{(2|1)}, \lambda_{\mathsf{c}}, \gamma^a, \gamma^b, \gamma^c, \mu_{\mathsf{s}}^{(1)}, \mu_{\mathsf{c}}^{(1)}, \mu_{\mathsf{s}}^{(2)}, \mu_{\mathsf{c}}^{(2)}, \mu_{\mathsf{c}}^{(1|2)}, \mu_{\mathsf{c}}^{(2|1)})$ represent the collection of all these Lagrange multipliers or dual variables.

To evaluate the optimal value of DPSW, it suffices to take $\gamma^a(s_1)$, $\gamma^b(s_2)$ and $\gamma^c(y_1, y_2)$ such that the constraints (D1), (D2), (D3) hold with equality. Thus, at optimality,

$$\gamma^a(s_1) \equiv \min_{x_1}\Big\{\sum_{y_1,y_2} \mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2) + \sum_{s_2} \mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2)\Big\}$$
$$\gamma^b(s_2) \equiv \min_{x_2}\Big\{\sum_{y_1,y_2} \mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2) + \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2)\Big\}$$
$$\gamma^c(y_1, y_2) \equiv \min_{\widehat{s}_1,\widehat{s}_2}\Big\{\sum_{s_2} \mu_{\mathsf{s}}^{(2)}(s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) + \sum_{s_1} \mu_{\mathsf{s}}^{(1)}(s_1, \widehat{s}_1, \widehat{s}_2, y)\Big\}.$$

Let $\bar{\Theta} := (\lambda_{\mathsf{s}}^{(1|2)}, \lambda_{\mathsf{s}}^{(2|1)}, \lambda_{\mathsf{c}}, \mu_{\mathsf{s}}^{(1)}, \mu_{\mathsf{c}}^{(1)}, \mu_{\mathsf{s}}^{(2)}, \mu_{\mathsf{c}}^{(2)}, \mu_{\mathsf{c}}^{(1|2)}, \mu_{\mathsf{c}}^{(2|1)})$ represent the collection of remaining dual variables. From the duality of linear programming, the following lemma then outlines our framework for obtaining lower bounds.

Lemma 3.1: Any collection of functions $\bar{\Theta}$ satisfying constraints (D4)-(D7) yields the following lower bound on the optimal value of SW, i.e.,

$$\begin{aligned}
&\mathrm{OPT(SW)} \overset{(a)}{\geq} \mathrm{OPT(LPSW)} \overset{(b)}{=} \mathrm{OPT(DPSW)} \\
&\geq \sum_{s_1} \min_{x_1}\Big\{\sum_{y_1,y_2} \mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2) + \sum_{s_2} \mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2)\Big\} \\
&+ \sum_{s_2} \min_{x_2}\Big\{\sum_{y_1,y_2} \mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2) + \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2)\Big\} \\
&+ \sum_{y_1,y_2} \min_{\widehat{s}_1,\widehat{s}_2}\Big\{\sum_{s_2} \mu_{\mathsf{s}}^{(2)}(s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) \\
&+ \sum_{s_1} \mu_{\mathsf{s}}^{(1)}(s_1, \widehat{s}_1, \widehat{s}_2, y)\Big\}.
\end{aligned} \tag{24}$$

The inequality in (a) follows since LPSW is a relaxation of SW and (b) results from the duality of linear programming.

$$
\begin{aligned}
\text{LPSW} \quad & \min_{\substack{W,T,V,U,\\ Q_{X_1|S_1},Q_{X_2|S_2},\\ Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}}} \sum_{z\in\mathcal{Z}} \mathbb{I}\{(s_1,s_2)\neq(\widehat{s}_1,\widehat{s}_2)\}\; P_{S_1,S_2}(s_1,s_2)\; \mathbb{I}\{(y_1,y_2)=(x_1,x_2)\}W(z) \\
& \sum_{x_1} Q_{X_1|S_1}(x_1|s_1) = 1 : \; \gamma^a(s_1) && \forall s_1 \\
& \sum_{x_2} Q_{X_2|S_2}(x_2|s_2) = 1 : \; \gamma^b(s_2) && \forall s_2 \\
& \sum_{\widehat{s}_1,\widehat{s}_2} Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1,\widehat{s}_2|y_1,y_2) = 1 : \; \gamma^c(y_1,y_2) && \forall y_1,y_2 \\
& \sum_{x_1} T(z_3) - Q_{X_2|S_2}(x_2|s_2) = 0 : \; \mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2) && \forall x_2,s_1,s_2 \\
& \sum_{x_1} U(z_1) - Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1,\widehat{s}_2|y_1,y_2) = 0 : \; \mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2) && \forall s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2 \\
& \sum_{x_1} W(z) - V(z_2) = 0 : \; \lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2) && \forall s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2 \\
\text{s.t.} \quad & \sum_{x_2} T(z_3) - Q_{X_1|S_1}(x_1|s_1) = 0 : \; \mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2) && \forall x_1,s_1,s_2 \\
& \sum_{x_2} V(z_2) - Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}(\widehat{s}_1,\widehat{s}_2|y_1,y_2) = 0 : \; \mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) && \forall s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2 \\
& \sum_{x_2} W(z) - U(z_1) = 0 : \; \lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2) && \forall s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2 \\
& \sum_{\widehat{s}_2,\widehat{s}_1} W(z) - T(z_3) = 0 : \; \lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2) && \forall s_1,s_2,x_1,x_2,y_1,y_2 \\
& \sum_{\widehat{s}_2,\widehat{s}_1} V(z_2) - Q_{X_2|S_2}(x_2|s_2) = 0 : \; \mu_{\mathsf{c}}^{(2)}(x_2,s_2,y_1,y_2) && \forall x_2,s_2,y_1,y_2 \\
& \sum_{\widehat{s}_1,\widehat{s}_2} U(z_1) - Q_{X_1|S_1}(x_1|s_1) = 0 : \; \mu_{\mathsf{c}}^{(1)}(x_1,s_1,y_1,y_2) && \forall x_1,s_1,y_1,y_2 \\
& Q_{X_1|S_1}, Q_{X_2|S_2}, Q_{\widehat{S}_1,\widehat{S}_2|Y_1,Y_2}, V,U,W,T \geq 0.
\end{aligned}
$$

$$
\begin{aligned}
\text{DPSW} \quad & \max_{\Theta} \sum_{s_1}\gamma^a(s_1) + \sum_{s_2}\gamma^b(s_2) + \sum_{y_1,y_2}\gamma^c(y_1,y_2) \\
& \gamma^a(s_1) - \textstyle\sum_{y_1,y_2}\mu_{\mathsf{c}}^{(1)}(s_1,x_1,y_1,y_2) - \sum_{s_2}\mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2) \leq 0 && \forall x_1,s_1 && (D1) \\
& \gamma^b(s_2) - \textstyle\sum_{y_1,y_2}\mu_{\mathsf{c}}^{(2)}(s_2,x_2,y_1,y_2) - \sum_{s_1}\mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2) \leq 0 && \forall x_2,s_2 && (D2) \\
& \gamma^c(y_1,y_2) - \textstyle\sum_{s_2}\mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) - \sum_{s_1}\mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2) \leq 0 && \forall \widehat{s}_1,\widehat{s}_2,y_1,y_2 && (D3) \\
\text{s.t.} \quad & \lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2) + \lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2) + \lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2) \leq \Pi(z) && \forall z && (D4) \\
& \mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) + \mu_{\mathsf{c}}^{(2)}(s_2,x_2,y_1,y_2) - \textstyle\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2) \leq 0 && \forall s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2 && (D5) \\
& \mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2) + \mu_{\mathsf{c}}^{(1)}(s_1,x_1,y_1,y_2) - \textstyle\sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2) \leq 0 && \forall s_1,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2 && (D6) \\
& \mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2) + \mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2) - \textstyle\sum_{y_1,y_2}\lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2) \leq 0 && \forall x_1,x_2,s_1,s_2 && (D7)
\end{aligned}
$$

The last inequality follows from $\bar{\Theta}$ being feasible for DPSW and using that (D1)-(D3) hold with equality.

Thus, to obtain finite blocklength lower bounds on SW, it suffices to construct functions,

$$
\begin{aligned}
\lambda_{\mathsf{s}}^{(2|1)} &: \mathcal{S}_1\times\mathcal{S}_2\times\mathcal{X}_2\times\mathcal{Y}_1\times\mathcal{Y}_2\times\widehat{\mathcal{S}}_1\times\widehat{\mathcal{S}}_2\to\mathbb{R}, \\
\lambda_{\mathsf{s}}^{(1|2)} &: \mathcal{S}_1\times\mathcal{S}_2\times\mathcal{X}_1\times\mathcal{Y}_1\times\mathcal{Y}_2\times\widehat{\mathcal{S}}_1\times\widehat{\mathcal{S}}_2\to\mathbb{R}, \\
\lambda_{\mathsf{c}} &: \mathcal{S}_1\times\mathcal{S}_2\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Y}_1\times\mathcal{Y}_2\to\mathbb{R}, \\
\mu_{\mathsf{s}}^{(1)} &: \mathcal{S}_1\times\widehat{\mathcal{S}}_1\times\widehat{\mathcal{S}}_2\times\mathcal{Y}_1\times\mathcal{Y}_2\to\mathbb{R}, \\
\mu_{\mathsf{c}}^{(1)} &: \mathcal{S}_1\times\mathcal{X}_1\times\mathcal{Y}_1\times\mathcal{Y}_2\to\mathbb{R}, \\
\mu_{\mathsf{s}}^{(2)} &: \mathcal{S}_2\times\widehat{\mathcal{S}}_1\times\widehat{\mathcal{S}}_2\times\mathcal{Y}_1\times\mathcal{Y}_2\to\mathbb{R}, \\
\mu_{\mathsf{c}}^{(2)} &: \mathcal{S}_2\times\mathcal{X}_2\times\mathcal{Y}_1\times\mathcal{Y}_2\to\mathbb{R}, \\
\mu_{\mathsf{c}}^{(2|1)} &: \mathcal{X}_2\times\mathcal{S}_1\times\mathcal{S}_2\to\mathbb{R}, \\
\mu_{\mathsf{c}}^{(1|2)} &: \mathcal{X}_1\times\mathcal{S}_1\times\mathcal{S}_2\to\mathbb{R},
\end{aligned}
\tag{25}
$$

such that the point-wise inequalities in (D4)-(D7) are satisfied.

We call the above collection of functions $\bar{\Theta}$, a feasible point of DPSW. As is evident, construction of such a feasible point of DPSW is challenging and probably cumbersome at first glance. Another hindrance is the difficulty in interpreting these variables so as to develop any intuitions on construction of these variables.

Consequently, in this paper, we present a systematic method to construct feasible points of DPSW and thereby, obtain finite blocklength converses for SW coding. We show that a combination of the source and channel flows of the problems DPJE, $\mathrm{DPSI}_{1|2}$ and $\mathrm{DPSI}_{2|1}$, yields a new feasible point of DPSW and thereby, a new finite blocklength converse. We discuss this in the next section.

IV. From Point-to-Point Converses to Slepian-Wolf Converses

In this section, we present a systematic synthesis of finite blocklength converses for the Slepian-Wolf coding problem from the source and channel flows of the dual programs DPJE, $\mathrm{DPSI}_{1|2}$ and $\mathrm{DPSI}_{2|1}$ discussed in Section II.

We begin by discussing the structure of DPSW. Since constraints (D1)-(D3) can be assumed to hold with equality, our main concern is with the variables $\lambda_{\mathsf{s}}^{(1|2)}, \lambda_{\mathsf{s}}^{(2|1)}, \lambda_{\mathsf{c}}$ and $\mu_{\mathsf{s}}^{(1)}, \mu_{\mathsf{s}}^{(2)}, \mu_{\mathsf{c}}^{(1)}, \mu_{\mathsf{c}}^{(2)}, \mu_{\mathsf{c}}^{(2|1)}, \mu_{\mathsf{c}}^{(1|2)}$. We will refer to these variables (recall that these are functions, as stated in (25)) also as *flows*. Our approach for interpreting and classifying these flows is based on relating these flows to flows of problems DPJE, $\mathrm{DPSI}_{1|2}$ and $\mathrm{DPSI}_{2|1}$. We remark that there may be other approaches that would yield a more refined understanding.

We begin with the λ's. Consider the flow $\lambda_{\mathsf{s}}^{(1|2)}$. Observe that $\lambda_{\mathsf{s}}^{(1|2)}$ is a function of $s_1, s_2, x_2, y_1, y_2, \widehat{s}_1, \widehat{s}_2$ but is independent of x_1. Hence, for each fixed s_2, x_2, y_2 and $\widehat{s}_2$, $\lambda_{\mathsf{s}}^{(1|2)}$ may be likened to a source flow from S_1 to $\widehat{S}_1$ (recall that the

source flow in the point-to-point problem was a function of the source, the channel output and the destination, but *not* of the channel input). The dependence of $\lambda_{\mathsf{s}}^{(1|2)}$ on $s_2, x_2, y_2, \widehat{s}_2$ hints at (coded or uncoded) side-information about S_2 through the path $S_2 \to X_2 \to Y_2 \to \widehat{S}_2$. This leads one to surmise that the flow $\lambda_{\mathsf{s}}^{(1|2)}$ would have a close relation to the source flow of the problem $\mathrm{DPSI}_{1|2}$. Thus we refer to $\lambda_{\mathsf{s}}^{(1|2)}$ as a source flow for DPSW. Note though, that this is not the only heuristic one can apply. If $\lambda_{\mathsf{s}}^{(1|2)}$ is assumed to be also independent of x_2, then $\lambda_{\mathsf{s}}^{(1|2)}(s_1, s_2, x_2, y_1, y_2, \widehat{s}_1, \widehat{s}_2)$ can be also interpreted to be the source flow in the problem DPJE where S_1, S_2 are jointly encoded. Thus one surmises that a value for $\lambda_{\mathsf{s}}^{(1|2)}$ could probably be arrived at by a combination of the source flows of DPJE and $\mathrm{DPSI}_{1|2}$. A similar heuristic can be applied to surmise that $\lambda_{\mathsf{s}}^{(2|1)}$ could be arrived at through the source flows of DPJE and $\mathrm{DPSI}_{2|1}$. The flow $\lambda_{\mathsf{c}}(s_1, s_2, x_1, x_2, y_1, y_2)$ which depends on correlated sources and the channel inputs and outputs, appears to be related to the channel flows of all three problems $\mathrm{DPJE}, \mathrm{DPSI}_{1|2}$ and $\mathrm{DPSI}_{2|1}$, and should therefore be a function of the latter flows. We refer to it as the channel flow. Thus in problem DPSW, there are *two* source flows and one channel flow that satisfy an error density bottleneck (D4).

We now come to the μ's. Notice that the flows of problem DPSW fall into a hierarchy wherein the λ's are constrained by the error density bottleneck (constraint (D4)), whereas the μ's are constrained by a bottleneck determined by the λ's. Arguing as in the case of the λ's we see that $\mu_{\mathsf{s}}^{(1)}$ and $\mu_{\mathsf{c}}^{(1)}$ are akin to source and channel flows of a coding problem along the path $S_1 \to X_1 \to Y_1 \to \widehat{S}_1$. Note though the *objective* of the problem (source coding, or something else) would depend on $\lambda_{\mathsf{s}}^{(2|1)}$, the RHS of constraint (D6). Likewise $\mu_{\mathsf{s}}^{(2)}$ and $\mu_{\mathsf{c}}^{(2)}$ resemble source channel flows for a coding problem along $S_2 \to X_2 \to Y_2 \to \widehat{S}_2$ whose objective is determined by $\lambda_{\mathsf{s}}^{(1|2)}$. The final set of dual variables $\mu_{\mathsf{c}}^{(1|2)}$ and $\mu_{\mathsf{c}}^{(2|1)}$ are somewhat distinct from the rest, since they do not seem to be analogous to any flows from point-to-point problems. We will interpret these later.

The following two propositions distill these heuristics into a formal relationship between the feasible regions of problems DPSW and problems DPJE, $\mathrm{DPSI}_{1|2}$ and $\mathrm{DPSI}_{2|1}$.

Proposition 4.1: Let $\bar{\Theta}_{1|2} := (\bar{\gamma}^a, \bar{\gamma}^b, \bar{\lambda}_{\mathsf{s}}^{(1|2)}, \bar{\lambda}_{\mathsf{c}}^{(1|2)}) \in \mathrm{FEA}(\mathrm{DPSI}_{1|2})$ with its corresponding objective value, $\mathrm{OBJ}(\mathrm{DPSI}_{1|2}) = \sum_{s_1} \bar{\gamma}^a(s_1) + \sum_{s_2, y_1} \bar{\gamma}^b(s_2, y_1)$. Then the following choice of values for the variables of DPSW is feasible.

$$\begin{aligned}
\lambda_{\mathsf{s}}^{(1|2)}(s_1, s_2, x_2, y_1, y_2, \widehat{s}_1, \widehat{s}_2) &\equiv \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1, s_2, \widehat{s}_1, y_1)\mathbb{I}\{x_2 = y_2\} \\
\lambda_{\mathsf{c}}(s_1, s_2, x_1, x_2, y_1, y_2) &\equiv \bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1, s_2, x_1, y_1)\mathbb{I}\{x_2 = y_2\}, \\
\mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2) &\equiv \bar{\gamma}^b(s_2, y_1)\mathbb{I}\{x_2 = y_2\}, \quad (26) \\
\mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2) &\equiv \sum_{y_1} \bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1, s_2, x_1, y_1),
\end{aligned}$$

$$\begin{aligned}
\lambda_{\mathsf{s}}^{(2|1)}(s_1, s_2, x_1, y_1, y_2, \widehat{s}_1, \widehat{s}_2) &\equiv 0, \quad \mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2) \equiv 0, \\
\mu_{\mathsf{s}}^{(1)}(s_1, \widehat{s}_1, \widehat{s}_2, y_1, y_2) &\equiv 0, \quad \mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2) \equiv 0, \\
\mu_{\mathsf{s}}^{(2)}(s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) &\equiv 0, \quad \gamma^c(y_1, y_2) \equiv 0,
\end{aligned}$$

$$\gamma^b(s_2) \equiv \sum_{y_1} \bar{\gamma}^b(s_2, y_1), \qquad \gamma^a(s_1) \equiv \bar{\gamma}^a(s_1).$$

Consequently, $\mathrm{OPT}(\mathrm{DPSW}) \geq \sum_{s_1} \gamma^a(s_1) + \sum_{s_2} \gamma^b(s_2) + \sum_{y_1, y_2} \gamma^c(y_1, y_2)$ which is equal to the objective of $\mathrm{DPSI}_{1|2}$ under $\bar{\Theta}_{1|2}$. In particular, considering $\bar{\Theta}_{1|2}$ as the optimal solution of $\mathrm{DPSI}_{1|2}$ gives $\mathrm{OPT}(\mathrm{DPSW}) \geq \mathrm{OPT}(\mathrm{DPSI}_{1|2})$.

Proof : The proof is included in Appendix B. ■

It thus becomes clear that given any feasible point $(\bar{\gamma}^a, \bar{\gamma}^b, \bar{\lambda}_{\mathsf{s}}^{(1|2)}, \bar{\lambda}_{\mathsf{c}}^{(1|2)}) \in \mathrm{FEA}(\mathrm{DPSI}_{1|2})$, one can construct a feasible point of DPSW as given in (26). Moreover, the objective value of the resulting feasible point gives a lower bound on $\mathrm{OPT}(\mathrm{DPSW})$. Similarly, it can be shown that given any feasible point of $\mathrm{DPSI}_{2|1}$, one can construct a feasible point of DPSW with the same cost, thereby implying

$$\mathrm{OPT}(\mathrm{DPSW}) \geq \mathrm{OPT}(\mathrm{DPSI}_{2|1}).$$

As with the problems with side-information, the following proposition illustrates a relation between the feasible regions of DPJE and DPSW.

Proposition 4.2: The following relationship between the feasible region of DPJE and DPSW holds. Let $\hat{\Theta} := (\hat{\gamma}^a, \hat{\gamma}^b, \hat{\lambda}_{\mathsf{s}}, \hat{\lambda}_{\mathsf{c}}) \in \mathrm{FEA}(\mathrm{DPJE})$ with its corresponding objective value, $\mathrm{OBJ}(\mathrm{DPJE}) = \sum_{s_1, s_2} \widehat{\gamma}^a(s_1, s_2) + \sum_{y_1, y_2} \widehat{\gamma}^b(y_1, y_2)$. Then the following choice of values of the variables of DPSW are feasible.

$$\begin{aligned}
\lambda_{\mathsf{s}}^{(1|2)}(s_1, s_2, x_2, y_1, y_2, \widehat{s}_1, \widehat{s}_2) &\equiv \widehat{\lambda}_{\mathsf{s}}(s_1, s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) \\
\lambda_{\mathsf{c}}(s_1, s_2, x_1, x_2, y_1, y_2) &\equiv \widehat{\lambda}_{\mathsf{c}}(s_1, s_2, x_1, x_2, y_1, y_2), \\
\mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2) &\equiv 0, \quad \mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2) \equiv 0, \\
\lambda_{\mathsf{s}}^{(2|1)}(s_1, s_2, x_1, y_1, y_2, \widehat{s}_1, \widehat{s}_2) &\equiv 0, \\
\mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2) &\equiv \widehat{\gamma}^a(s_1, s_2), \\
\mu_{\mathsf{s}}^{(1)}(s_1, \widehat{s}_1, \widehat{s}_2, y_1, y_2) &\equiv 0, \quad \mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2) \equiv 0,
\end{aligned}$$

$$\begin{aligned}
\gamma^a(s_1) &\equiv 0, \quad \gamma^b(s_2) \equiv \sum_{s_1} \widehat{\gamma}^a(s_1, s_2), \\
\gamma^c(y_1, y_2) &\equiv \widehat{\gamma}^b(y_1, y_2), \\
\mu_{\mathsf{s}}^{(2)}(s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) &\equiv \sum_{s_1} \widehat{\lambda}_{\mathsf{s}}(s_1, s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2).
\end{aligned}$$

Consequently, $\mathrm{OPT}(\mathrm{DPSW}) \geq \sum_{s_1} \gamma^a(s_1) + \sum_{s_2} \gamma^b(s_2) + \sum_{y_1, y_2} \gamma^c(y_1, y_2)$ which is the objective of DPJE under $\hat{\Theta}$. Moreover, considering $\hat{\Theta}$ to be the optimizing feasible point of DPJE yields that $\mathrm{OPT}(\mathrm{DPSW}) \geq \mathrm{OPT}(\mathrm{DPJE})$.

Proof : The proof is similar to the proof of Proposition 4.1 and we skip the proof here. ■

The relationships between the feasible regions of DPSW with that of DPJE, $\mathrm{DPSI}_{1|2}$ and $\mathrm{DPSI}_{2|1}$ established through Propositions 4.1 and 4.2 help in establishing a formal interpretation for the roles of the dual variables of DPSW. From Proposition 4.1, we see that the dual variable $\lambda_{\mathsf{s}}^{(1|2)}$ that we considered as akin to a source flow for $\mathrm{DPSI}_{1|2}$ and also as a source flow for the DPJE, has a somewhat more complex interpretation. Specifically, while the latter interpretation holds thanks to Proposition 4.2, the former holds only along the

diagonal $x_2 = y_2$, as seen in (26). A similar caveat holds for the channel flow λ_{c} and the other source flow $\lambda_{\mathsf{s}}^{(2|1)}$.

We remark here that the choices of the μ's in Propositions 4.2 and 4.1 are not necessarily optimal and hence, one may not obtain a sharp interpretation for these flows. Nonetheless, it can be seen that when the sum $\sum_{s_1} \lambda_{\mathsf{s}}^{(1|2)}$ is considered as in constraint (D5), the source flow from S_1 to destination $\widehat{S}_1$ is averaged out, and what is left, is the influence of side-information of S_2. From Proposition 4.2 it can be seen that $\mu_{\mathsf{s}}^{(2)}$ accounts for the point-to-point like source flow through the path from S_2 to the destination node $(\widehat{S}_1, \widehat{S}_2)$. Further, (26) implies that $\mu_{\mathsf{c}}^{(2)}$ accounts for the channel flow through this path. Similarly, $\mu_{\mathsf{s}}^{(1)}$ represents the point-to-point like source flow through the path from S_1 to $(\widehat{S}_1, \widehat{S}_2)$ via the decoder and $\mu_{\mathsf{c}}^{(1)}$ represents the corresponding channel flow for this path. Finally, thanks to the relation $\mu_{\mathsf{c}}^{(1|2)} = \sum_{y_1} \bar{\lambda}_{\mathsf{c}}^{(1|2)}$ in (26), we can interpret that $\mu_{\mathsf{c}}^{(1|2)}$ represents an average channel flow from S_1 to $(\widehat{S}_1, \widehat{S}_2)$ given the side-information about S_2. Similarly, $\mu_{\mathsf{c}}^{(2|1)}$ represents an average channel flow from S_2 given information about S_1.

A. Synthesizing Converses for SW from Point-to-Point Duals

As an immediate consequence of Proposition 4.1 and Proposition 4.2, we get that the point-to-point metaconverses in (10), (16) and (8) are all lower bounds on OPT(DPSW). Consequently, the following is a straightforward lower bound on OPT(DPSW),

$$\begin{aligned}
\mathrm{OPT(DPSW)} \geq \max\Bigg\{ & \sup_{0\leq \widehat{\phi}(s_1,s_2)\leq P_{S_1,S_2}(s_1,s_2)} \Bigg\{ \sum_{s_1,s_2} \widehat{\phi}(s_1,s_2) \\
& - M_1 M_2 \max_{\widehat{s}_1,\widehat{s}_2} \widehat{\phi}(\widehat{s}_1,\widehat{s}_2) \Bigg\}, \\
& \sup_{0\leq \phi^{(1|2)}(s_1,s_2)\leq P_{S_1,S_2}(s_1,s_2)} \Bigg\{ \sum_{s_1,s_2} \phi^{(1|2)}(s_1,s_2) \\
& - M_1 \sum_{s_2} \max_{\widehat{s}_1} \phi^{(1|2)}(\widehat{s}_1,s_2) \Bigg\}, \\
& \sup_{0\leq \phi^{(2|1)}(s_1,s_2)\leq P_{S_1,S_2}(s_1,s_2)} \Bigg\{ \sum_{s_1,s_2} \phi^{(2|1)}(s_1,s_2) \\
& - M_2 \sum_{s_1} \max_{\widehat{s}_2} \phi^{(2|1)}(s_1,\widehat{s}_2) \Bigg\} \Bigg\}. \qquad (27)
\end{aligned}$$

Convex analytically speaking, the above bound considers a *convex combination* of feasible points of DPSW obtained via Propositions 4.1 and 4.2. In the following theorem we synthesize a new feasible point for DPSW by a *nonlinear* combination of the source and channel flows in the point-to-point dual programs DPJE, $\mathrm{DPSI}_{1|2}$ and $\mathrm{DPSI}_{2|1}$. We will subsequently apply specific metaconverses from Corollary 2.6, Theorem 2.7 and Theorem 2.8 to get our new metaconverse.

Theorem 4.3: Let $(\bar{\gamma}^a, \bar{\gamma}^b, \bar{\lambda}_{\mathsf{s}}^{(1|2)}, \bar{\lambda}_{\mathsf{c}}^{(1|2)}) \in \mathrm{FEA}(\mathrm{DPSI}_{1|2})$, $(\tilde{\gamma}^a, \tilde{\gamma}^b, \tilde{\lambda}_{\mathsf{s}}^{(2|1)}, \tilde{\lambda}_{\mathsf{c}}^{(2|1)}) \in \mathrm{FEA}(\mathrm{DPSI}_{2|1})$ and $(\widehat{\gamma}^a, \widehat{\gamma}^b, \widehat{\lambda}_{\mathsf{s}}, \widehat{\lambda}_{\mathsf{c}}) \in \mathrm{FEA}(\mathrm{DPJE})$. Then, any choice of values for the variables of DPSW satisfying the following equations is feasible for DPSW.

$$\begin{aligned}
\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2) = & \Big[\bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1,s_2,\widehat{s}_1,y_1) \times \\
& \mathbb{I}\{x_2 = y_2\} + \alpha \widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) \Big] \mathbb{I}\{(s_1,s_2) = (\widehat{s}_1,\widehat{s}_2)\}, \\
\lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2) = & \Big[\tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1,s_2,\widehat{s}_2,y_2) \times \\
& \mathbb{I}\{x_1 = y_1\} + (1-\alpha) \widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) \Big] \times \\
& \mathbb{I}\{(s_1,s_2) = (\widehat{s}_1,\widehat{s}_2)\}, \\
\lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2) = & \min \Big\{ P(s_1,s_2) \mathbb{I}\{(y_1,y_2) = (x_1,x_2)\}, \\
& \widehat{\lambda}_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2) + \bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1,s_2,x_1,y_1) \mathbb{I}\{x_2 = y_2\} \\
& + \tilde{\lambda}_{\mathsf{c}}^{(2|1)}(s_1,s_2,x_2,y_2) \mathbb{I}\{x_1 = y_1\} \Big\}, \qquad (28) \\
\mu_{\mathsf{c}}^{(2)}(s_2,x_2,y_1,y_2) \leq & \Big[\bar{\gamma}^b(y_1,s_2) - \sum_{s_1 \neq \widehat{s}_1} \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1,s_2,\widehat{s}_1,y_2) \Big] \\
& \times \mathbb{I}\{x_2 = y_2\} \mathbb{I}\{s_2 = \widehat{s}_2\}, \\
\mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) = & \alpha \widehat{\lambda}_{\mathsf{s}}(\widehat{s}_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) \mathbb{I}\{s_2 = \widehat{s}_2\}, \\
\mu_{\mathsf{c}}^{(1)}(s_1,x_1,y_1,y_2) \leq & \Big[\tilde{\gamma}^b(s_1,y_2) - \sum_{s_2 \neq \widehat{s}_2} \tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_2,s_2,\widehat{s}_2,y_2) \Big] \\
& \times \mathbb{I}\{x_1 = y_1\} \mathbb{I}\{s_1 = \widehat{s}_1\}, \\
\mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2) = & (1-\alpha) \widehat{\lambda}_{\mathsf{s}}(s_1,\widehat{s}_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) \\
& \times \mathbb{I}\{s_1 = \widehat{s}_1\}, \\
\mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2) + & \mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2) \\
& \leq \sum_{y_1,y_2} \lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2),
\end{aligned}$$

where $\alpha \in (0,1)$ and $\gamma^a(s_1), \gamma^b(s_2)$ and $\gamma^c(y_1,y_2)$ are chosen such that (D1), (D2) and (D3) hold with equality.

Proof : The proof is included in Appendix B. ■

Theorem 4.3 generates a new feasible point for DPSW using an appropriate and nonlinear combination of the feasible points of the point-to-point source coding problems. Notice that the source flow $\lambda_{\mathsf{s}}^{(1|2)}$ is taken as a superposition of the source flow for $\mathrm{DPSI}_{1|2}$ (which acts only when $x_2 = y_2$) and a fraction of the source flow for DPJE. Similarly, $\lambda_{\mathsf{s}}^{(2|1)}$ is a superposition of the source flow for $\mathrm{DPSI}_{2|1}$ (which acts only when $x_1 = y_1$) with the remaining fraction of the source flow for DPJE. While $\lambda_{\mathsf{s}}^{(1|2)}$ and $\lambda_{\mathsf{s}}^{(2|1)}$ are linear combinations of the point-to-point source flows, the channel flow λ_{c} considers a nonlinear combination of the channel flows of $\mathrm{DPSI}_{1|2}$, $\mathrm{DPSI}_{2|1}$ and DPJE, so as to satisfy the bottleneck in constraint (D4). Since λ_{c} dictates the choice of $\mu_{\mathsf{c}}^{(1|2)}$ and $\mu_{\mathsf{c}}^{(2|1)}$, the nonlinearity is also inherited in the relation of μ's.

Moreover, the nonlinear relation between λ_{c} and the channel flows of the point-to-point problems is one of the reasons for the improvement on the classical converse of Miyake and Kanaya, as can be seen later in (31). This improvement is

otherwise hard to deduce from a feasible point of DPSW resulting from a convex combination of point-to-point feasible points.

Thanks to Theorem 4.3, to obtain finite blocklength converses for Slepian-Wolf coding, it suffices to consider the simpler point-to-point source-coding problems and construct good feasible points for them. In particular, considering those feasible points of $\mathrm{DPSI}_{1|2}$, $\mathrm{DPSI}_{2|1}$ and DPJE which yield the metaconverses in (10), (16) and (8) for the corresponding point-to-point sub-problems and subsequently employing Theorem 4.3, we obtain the following new finite blocklength converse for SW.

Theorem 4.4 (Metaconverse for Slepian-Wolf Coding): Consider the problem SW. Consequently, for any code, the following bound holds:

$$\mathbb{E}[\mathbb{I}\{(S_1,S_2)\neq(\widehat{S}_1,\widehat{S}_2)\}]\geq \mathrm{OPT(SW)}\geq \mathrm{OPT(DPSW)}\geq \sup_{\widehat{\phi},\phi^{(1|2)},\phi^{(2|1)}}\Bigg\{\sum_{s_1,s_2}\min\{P_{S_1,S_2}(s_1,s_2),\widehat{\phi}(s_1,s_2)+\phi^{(1|2)}(s_1,s_2) + \phi^{(2|1)}(s_1,s_2)\} - M_1M_2\max_{\widehat{s}_1,\widehat{s}_2}\widehat{\phi}(\widehat{s}_1,\widehat{s}_2) - M_2\sum_{s_1}\max_{\widehat{s}_2}\phi^{(2|1)}(s_1,\widehat{s}_2) - M_1\sum_{s_2}\max_{\widehat{s}_1}\phi^{(1|2)}(\widehat{s}_1,s_2)\Bigg\}, \tag{29}$$

where the supremum is over $\widehat{\phi},\phi^{(1|2)},\phi^{(2|1)}:\mathcal{S}_1\times\mathcal{S}_2\to[0,1]$ such that $0\leq\widehat{\phi}(s_1,s_2),\phi^{(1|2)}(s_1,s_2),\phi^{(2|1)}(s_1,s_2)\leq P_{S_1,S_2}(s_1,s_2)$ for all $s_1\in\mathcal{S}_1,s_2\in\mathcal{S}_2$.

Proof : The proof is included in Appendix B. ■

In particular, choosing

$$\widehat{\phi}(s_1,s_2)=\min\{P_{S_1,S_2}(s_1,s_2),\eta_1(s_1,s_2)\}$$
$$\phi^{(1|2)}(s_1,s_2)=\min\{P_{S_1,S_2}(s_1,s_2),\eta_2(s_1,s_2)\}$$
$$\phi^{(2|1)}(s_1,s_2)=\min\{P_{S_1,S_2}(s_1,s_2),\eta_3(s_1,s_2)\}$$

in (29), where $\eta_1,\eta_2,\eta_3:\mathcal{S}_1\times\mathcal{S}_2\to[0,\infty)$, we get the following bound.

$$\mathbb{E}[\mathbb{I}\{(S_1,S_2)\neq(\widehat{S}_1,\widehat{S}_2)\}]\geq \mathrm{OPT(SW)}\geq \mathrm{OPT(DPSW)}\geq \sup_{\eta_1,\eta_2,\eta_3\geq 0}\Bigg\{\sum_{s_1,s_2}\min\{P_{S_1,S_2}(s_1,s_2),\eta_1(s_1,s_2)+\eta_2(s_1,s_2)+ \eta_3(s_1,s_2)\} - M_1M_2\max_{\widehat{s}_1,\widehat{s}_2}\min\{P_{S_1,S_2}(\widehat{s}_1,\widehat{s}_2),\eta_1(\widehat{s}_1,\widehat{s}_2)\} - M_2\sum_{s_1}\max_{\widehat{s}_2}\min\{P_{S_1,S_2}(s_1,\widehat{s}_2),\eta_3(s_1,\widehat{s}_2)\} - M_1\sum_{s_2}\max_{\widehat{s}_1}\min\{P_{S_1,S_2}(\widehat{s}_1,s_2),\eta_2(\widehat{s}_1,s_2)\}\Bigg\} \tag{30}$$

Further, the new converse in (30) improves on the information spectrum based converse of Miyake and Kanaya [2] as shown in the following corollary.

Corollary 4.5 (Improvements on Miyake-Kanaya Converse): The converse in (30) implies the following improvement on the converse of Miyake and Kanaya,

$$\mathbb{E}[\mathbb{I}\{(S_1,S_2)\neq(\widehat{S}_1,\widehat{S}_2)\}]\geq \mathrm{OPT(SC)}\geq \mathrm{OPT(DP)}\geq \sup_{\beta>0}\Bigg\{\mathbb{P}\Bigg[h_{S_1,S_2}(S_1,S_2)\geq\log M_1M_2+\beta \text{ or } h_{S_1|S_2}(S_1|S_2) \geq\log M_1+\beta \text{ or } h_{S_2|S_1}(S_2|S_1)\geq\log M_2+\beta\Bigg]+ \sum_{s_1,s_2}\max\Bigg\{\frac{\exp(-\beta)}{M_1M_2},\frac{\exp(-\beta)}{M_1}P_{S_2}(s_2),\frac{\exp(-\beta)}{M_2}P_{S_1}(s_1)\Bigg\} \times\mathbb{I}\Bigg\{P_{S_1|S_2}(s_1|s_2)>\frac{\exp(-\beta)}{M_1},P_{S_2|S_1}(s_2|s_1)>\frac{\exp(-\beta)}{M_2}, P_{S_1,S_2}(s_1,s_2)>\frac{\exp(-\beta)}{M_1M_2}\Bigg\}-3\exp(-\beta)\Bigg\}, \tag{31}$$

where $h_{A|B}(a|b)\equiv-\log P_{A|B}(a|b)$ is the conditional entropy density and $h_{A,B}(a,b)\equiv-\log P_{A,B}(a,b)$ is the joint entropy density.

Proof : To obtain the above converse, weaken (30) by choosing $\eta_1(s_1,s_2)\equiv\frac{\exp(-\beta)}{M_1M_2}$, $\eta_2(s_1,s_2)\equiv P_{S_2}(s_2)\frac{\exp(-\beta)}{M_1}$, $\eta_3(s_1,s_2)\equiv P_{S_1}(s_1)\frac{\exp(-\beta)}{M_2}$ and bound $\min\{P(s_1,s_2),\eta_1(s_1,s_2)+\eta_2(s_1,s_2)+\eta_3(s_1,s_2)\}$ by $\min\{P(s_1,s_2),\max\{\eta_1(s_1,s_2),\eta_2(s_1,s_2),\eta_3(s_1,s_2)\}\}$. Further, bound $\min\{P_{S_1,S_2}(\widehat{s}_1,\widehat{s}_2),\eta_1(\widehat{s}_1,\widehat{s}_2)\}$ by $\eta_1(s_1,s_2)$, $\min\{P_{S_1,S_2}(\widehat{s}_1,\widehat{s}_2),\eta_2(\widehat{s}_1,\widehat{s}_2)\}$ by $\eta_2(s_1,s_2)$ and $\min\{P_{S_1,S_2}(\widehat{s}_1,\widehat{s}_2),\eta_3(\widehat{s}_1,\widehat{s}_2)\}$ by $\eta_3(s_1,s_2)$. Subsequently, employing the definition of $h_{A|B}(a|b)$, $h_{A,B}(a,b)$ and taking supremum over $\beta>0$, we get the required converse. ■

Remark IV.2. (Recovering the Converse of Miyake and Kanaya) Lower bounding the non-negative term in (31) corresponding to $\mathbb{I}\Bigg\{P_{S_1|S_2}(s_1|s_2)>\frac{\exp(-\beta)}{M_1},P_{S_2|S_1}(s_2|s_1)>\frac{\exp(-\beta)}{M_2},P_{S_1,S_2}(s_1,s_2)>\frac{\exp(-\beta)}{M_1M_2}\Bigg\}$ with zero, we recover the converse of Miyake and Kanaya given as,

$$\mathbb{E}[\mathbb{I}\{(S_1,S_2)\neq(\widehat{S}_1,\widehat{S}_2)\}]\geq\sup_{\beta>0}\Bigg\{\mathbb{P}\Bigg[h_{S_1,S_2}(S_1,S_2) \geq\log M_1M_2+\beta \text{ or } h_{S_1|S_2}(S_1|S_2)\geq\log M_1+\beta \text{ or } h_{S_2|S_1}(S_2|S_1)\geq\log M_2+\beta\Bigg]-3\exp(-\beta)\Bigg\}. \tag{32}$$

□

Before we conclude, we note that the relevance of (27), particularly in the analysis of second-order asymptotics is limited. As pointed out in [12, Section 6.2], the second-order analysis centered at a corner point of the first order rate region of Slepian-Wolf problem, is determined by the multivariate Gaussian CDF with respect to jointly encoded and side-information problems together. Consequently, the lower bound in (29) or the union bound in (31) are more relevant. In fact, with the flexibility of choosing η_1,η_2,η_3 which are functions of (s_1,s_2), the converse in (29) may even yield refined third order terms in the asymptotic analysis.

B. Illustrative Example:Doubly Symmetric Binary Sources

In this section, we consider the example of a Doubly Symmetric Binary Source (DSBS) with $\mathcal{S}_1 = \mathcal{S}_2 = \{0,1\}^n$ and the joint probability distribution given as,

$$P_{S_1,S_2}(s_1,s_2) \equiv \frac{1}{2^n} p^{d(s_1,s_2)}(1-p)^{n-d(s_1,s_2)}, \tag{33}$$

where $p < 0.5$ and $d(s_1,s_2)$ represents the Hamming distance between $s_1 \in \mathcal{S}_1$ and $s_2 \in \mathcal{S}_2$. Further, $M_1 = 2^{nR_1}$, $M_2 = 2^{nR_2}$, $P_{S_1|S_2}(s_1|s_2) = P_{S_2|S_1}(s_2|s_1) \equiv p^{d(s_1,s_2)}(1-p)^{n-d(s_1,s_2)}$ and $P_{S_1}(s_1) = P_{S_2}(s_2) = \frac{1}{2^n}$. The optimal rate region for this DSBS is given by [13],

$$\mathcal{R}_{SW} = \{(R_1,R_2) \mid R_1, R_2 \geq H(p), R_1 + R_2 \geq 1 + H(p)\}.$$

Particularizing the converse in (30) to the case of DSBS by choosing $\eta_2(s_1,s_2) \equiv P_{S_2}(s_2)\frac{2^{-\beta}}{M_1}$, $\eta_3(s_1,s_2) \equiv P_{S_1}(s_1)\frac{2^{-\beta}}{M_2}$ and $\eta_1(s_1,s_2) \equiv \frac{2^{-\beta}}{M_1 M_2}$ results in the following converse.

$$\begin{aligned}
&\mathbb{E}[\mathbb{I}\{(S_1,S_2) \neq (\widehat{S}_1,\widehat{S}_2)\}] \geq \mathrm{OPT(DPSW)} \\
&\geq \sup_{\beta>0}\Bigg\{\sum_{k=0}^{n} {}^nC_k \min\Bigg\{p^k(1-p)^{n-k}, \frac{2^{-\beta}}{M_1} + \frac{2^{-\beta}}{M_2} + \frac{2^{-\beta+n}}{M_2 M_1}\Bigg\} \\
&\quad - M_1 2^n \max_{k\in\{0,\ldots n\}} \min\Bigg\{\frac{p^k(1-p)^{n-k}}{2^n}, \frac{2^{-\beta}}{2^n M_1}\Bigg\} \\
&\quad - M_2 2^n \max_{k\in\{0,\ldots n\}} \min\Bigg\{\frac{p^k(1-p)^{n-k}}{2^n}, \frac{2^{-\beta}}{2^n M_2}\Bigg\} \\
&\quad - M_1 M_2 \max_{k\in\{0,\ldots n\}} \min\Bigg\{\frac{p^k(1-p)^{n-k}}{2^n}, \frac{2^{-\beta}}{M_1 M_2}\Bigg\}\Bigg\}.
\end{aligned} \tag{34}$$

The above bound follows from (29) since for any $s_1 \in \mathcal{S}_1$, the number of s_2's at a Hamming distance of $k \in \{0,\ldots,n\}$ is given by ${}^nC_k$.

Figure 4 and Figure 5 compare our improved converse (34) with the Miyake-Kanaya converse in (32). It is seen see that the improved converse in (34) shows a nontrivial improvement on the Miyake-Kanaya converse. Note the differences in scale in both Figure 4 and Figure 5.

Remark IV.3. For the case of DSBS whose joint distribution depends only on the Hamming distance, choosing η_1, η_2, η_3 in (30) to be independent of (s_1,s_2) results in (34) performing weaker than (27). To see this, lower bound (27) with the converse from jointly encoded sources i.e., (8), and choose

$$\begin{aligned}
\widehat{\phi}(s_1,s_2) = \min\Bigg\{&p^{d(s_1,s_2)}(1-p)^{n-d(s_1,s_2)}\frac{1}{2^n}, \\
&\min\Bigg\{p^{d(s_1,s_2)}(1-p)^{n-d(s_1,s_2)}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_1}\Bigg\} \\
&+ \min\Bigg\{p^{d(s_1,s_2)}(1-p)^{n-d(s_1,s_2)}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_2}\Bigg\} \\
&+ \min\Bigg\{p^{d(s_1,s_2)}(1-p)^{n-d(s_1,s_2)}\frac{1}{2^n}, \frac{2^{-\beta}}{M_1 M_2}\Bigg\}\Bigg\}
\end{aligned} \tag{35}$$

and subsequently, upper bound $\max_{\widehat{s}_1,\widehat{s}_2} \widehat{\phi}(\widehat{s}_1,\widehat{s}_2)$ with

$$\begin{aligned}
\max_{\widehat{s}_1,\widehat{s}_2}\Bigg\{&\min\{p^{d(\widehat{s}_1,\widehat{s}_2)}(1-p)^{n-d(\widehat{s}_1,\widehat{s}_2)}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_1}\} + \\
&\min\{p^{d(\widehat{s}_1,\widehat{s}_2)}(1-p)^{n-d(\widehat{s}_1,\widehat{s}_2)}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_2}\} \\
&+ \min\{p^{d(\widehat{s}_1,\widehat{s}_2)}(1-p)^{n-d(\widehat{s}_1,\widehat{s}_2)}\frac{1}{2^n}, \frac{2^{-\beta}}{M_2 M_1}\}\Bigg\}.
\end{aligned}$$

Further, noting that $\widehat{\phi}(s_1,s_2)$ is equivalent to, $\min\{p^{d(s_1,s_2)}(1-p)^{n-d(s_1,s_2)}\frac{1}{2^n}, \frac{2^{-\beta}}{M_1 2^n} + \frac{2^{-\beta}}{2^n M_2} + \frac{2^{-\beta}}{M_2 M_1}\}$ and $\sum_{s_1,s_2}\widehat{\phi}(s_1,s_2) = \sum_{k=0}^{n} {}^nC_k \min\Bigg\{p^k(1-p)^{n-k}, \frac{2^{-\beta}}{M_1} + \frac{2^{-\beta}}{M_2} + \frac{2^{-\beta+n}}{M_2 M_1}\}$ the following lower bound follows from (27),

$$\begin{aligned}
&\sum_{k=0}^{n} {}^nC_k \min\Bigg\{p^k(1-p)^{n-k}, \frac{2^{-\beta}}{M_1} + \frac{2^{-\beta}}{M_2} + \frac{2^{-\beta+n}}{M_2 M_1}\Bigg\} \\
&- M_1 M_2 \max_{k\in\{0,\ldots,n\}}\Bigg\{\min\{p^k(1-p)^{n-k}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_1}\} + \\
&\min\{p^k(1-p)^{n-k}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_2}\} + \min\{p^k(1-p)^{n-k}\frac{1}{2^n}, \\
&\quad \frac{2^{-\beta}}{M_2 M_1}\}\Bigg\} \\
&\geq \sum_{k=0}^{n} {}^nC_k \min\Bigg\{p^k(1-p)^{n-k}, \frac{2^{-\beta}}{M_1} + \frac{2^{-\beta}}{M_2} + \frac{2^{-\beta+n}}{M_2 M_1}\Bigg\} \\
&- M_1 M_2 \max_{k\in\{0,\ldots,n\}} \min\{p^k(1-p)^{n-k}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_1}\} \\
&- M_1 M_2 \max_{k\in\{0,\ldots,n\}} \min\{p^k(1-p)^{n-k}\frac{1}{2^n}, \frac{2^{-\beta}}{2^n M_2}\} \\
&- M_1 M_2 \max_{k\in\{0,\ldots,n\}} \min\{p^k(1-p)^{n-k}\frac{1}{2^n}, \frac{2^{-\beta}}{M_2 M_1}\}\Bigg\}.
\end{aligned} \tag{36}$$

It is now easy to see that when $M_1, M_2 \leq 2^n$, (36) outperforms (34). Note, however, that this outperformance relies on a particular choice of the η_1, η_2, η_3 in Theorem 4.4 which leads to (34) and on the structure of the DSBS. In particular, it *does not* imply that Theorem 4.4 is weaker than (27). □

V. Discussion

The tightest finite blocklength converse derivable for the SW problem employing the LP-based framework is OPT(DPSW), the exact evaluation of which is difficult. However, since the framework poses the tightest converse as an optimization problem, a hierarchy of lower bounds on it can be derived through a series of optimization problems. This also helps us conceptually situate the metaconverse in (29) and the Miyake-Kanaya converse in the hierarchy, as discussed below.

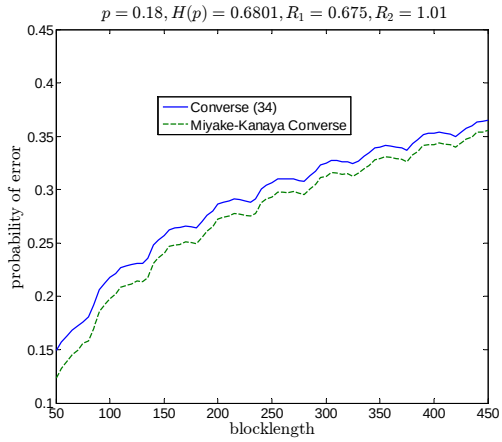

Fig. 4: DSBS with $(R_1, R_2) \notin \mathcal{R}_{SW}$.

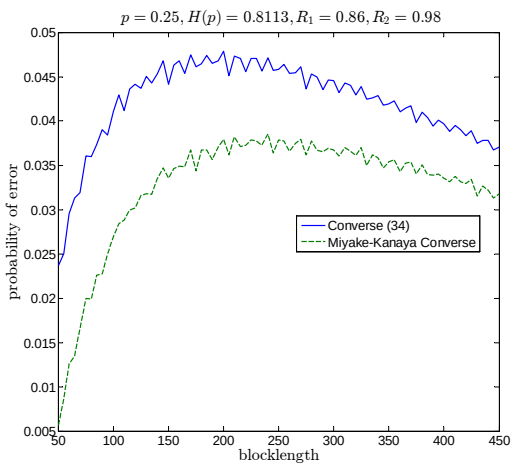

Fig. 5: DSBS with $(R_1, R_2) \in \mathcal{R}_{SW}$.

Recall that OPT(DPSW) evaluates to the following optimization problem,

$$\max_{\bar{\Theta}} \Bigg\{ \sum_{s_1} \min_{x_1} \Bigg\{ \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2) + \sum_{s_2} \mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2) \Bigg\} + \sum_{s_2} \min_{x_2} \Bigg\{ \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2) + \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2) \Bigg\} + \sum_{y_1,y_2} \min_{\widehat{s}_1,\widehat{s}_2} \Bigg\{ \sum_{s_2} \mu_{\mathsf{s}}^{(2)}(s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) + \sum_{s_1} \mu_{\mathsf{s}}^{(1)}(s_1, \widehat{s}_1, \widehat{s}_2, y) \Bigg\} \Bigg\},$$

s.t. (D4)-(D7) hold.

Note that in the above equation, the source flow $\mu_{\mathsf{s}}^{(1)}$ and channel flow $\mu_{\mathsf{c}}^{(1)}$, though along the same path are under two different minimum's. So are the pairs $(\mu_{\mathsf{s}}^{(2)}, \mu_{\mathsf{c}}^{(2)})$ and $(\mu_{\mathsf{c}}^{(1|2)}, \mu_{\mathsf{c}}^{(2|1)})$. Consequently, we first try to bring the terms in each of the above pairs together. Towards this, we first separate out the terms inside the minimum's in OPT(DPSW)

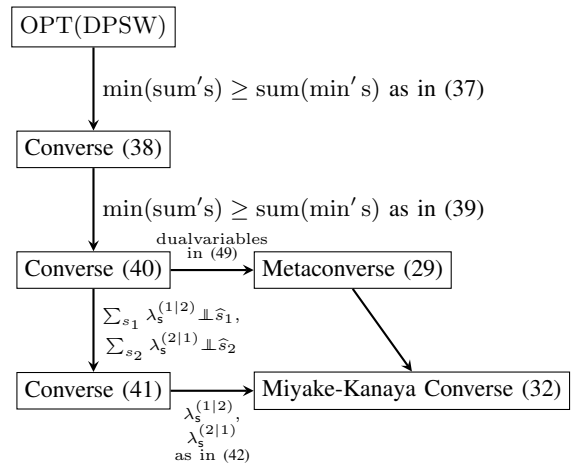

Fig. 6: Hierarchy of lower bounds derived. An arrow from $A \to B$ implies that $A \geq B$, the heading above the arrow indicate how B is obtained from A. $\min(\text{sum}'\text{s}) \geq \text{sum}(\min'\text{s})$ represent that minimum of sums is atleast equal to the sum of minimums.

by employing that

$$\sum_{s_1} \min_{x_1} \Bigg\{ \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(1)} + \sum_{s_2} \mu_{\mathsf{c}}^{(1|2)} \Bigg\} \geq \sum_{s_1} \Bigg[\min_{x_1} \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(1)} + \min_{x_1} \sum_{s_2} \mu_{\mathsf{c}}^{(1|2)} \Bigg],$$

$$\sum_{s_2} \min_{x_2} \Bigg\{ \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(2)} + \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)} \Bigg\} \geq \sum_{s_2} \Bigg[\min_{x_2} \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(2)} + \min_{x_2} \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)} \Bigg],$$

$$\sum_{y_1,y_2} \min_{\widehat{s}_1,\widehat{s}_2} \Bigg\{ \sum_{s_2} \mu_{\mathsf{s}}^{(2)} + \sum_{s_1} \mu_{\mathsf{s}}^{(1)} \Bigg\} \geq \sum_{y_1,y_2} \Bigg[\min_{\widehat{s}_1,\widehat{s}_2} \sum_{s_2} \mu_{\mathsf{s}}^{(2)} + \min_{\widehat{s}_1,\widehat{s}_2} \sum_{s_1} \mu_{\mathsf{s}}^{(1)} \Bigg]. \quad (37)$$

This results in the following optimization problem, the optimal value of which yields a lower bound on OPT(DPSW),

$$\text{OPT(DPSW)} \geq \max_{\bar{\Theta}} \Bigg\{ \sum_{s_1} \min_{x_1} \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2) + \sum_{s_1} \min_{x_1} \sum_{s_2} \mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2) + \sum_{s_2} \min_{x_2} \sum_{y_1,y_2} \mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2) + \sum_{s_2} \min_{x_2} \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2) + \sum_{y_1,y_2} \min_{\widehat{s}_1,\widehat{s}_2} \sum_{s_2} \mu_{\mathsf{s}}^{(2)}(s_2, \widehat{s}_1, \widehat{s}_2, y_1, y_2) + \sum_{y_1,y_2} \min_{\widehat{s}_1,\widehat{s}_2} \sum_{s_1} \mu_{\mathsf{s}}^{(1)}(s_1, \widehat{s}_1, \widehat{s}_2, y) \Bigg\}$$

s.t. (D4)-(D7) hold. (38)

We now further lower bound (38) by using,

$$\min_{x_1}\sum_{y_1,y_2}\mu_{\mathsf{c}}^{(1)} \geq \sum_{y_2}\min_{x_1}\sum_{y_1}\mu_{\mathsf{c}}^{(1)}$$
$$\min_{x_2}\sum_{y_1,y_2}\mu_{\mathsf{c}}^{(2)} \geq \sum_{y_1}\min_{x_2}\sum_{y_2}\mu_{\mathsf{c}}^{(2)}, \tag{39}$$

to get the following optimization problem, whose optimal value is a lower bound on (38),

$$\begin{aligned}
&\max_{\substack{\lambda_{\mathsf{s}}^{(2|1)},\lambda_{\mathsf{s}}^{(1|2)},\lambda_{\mathsf{c}}\\ \text{s.t (D4) holds}}}\Bigg\{\max_{\substack{\mu_{\mathsf{s}}^{(2)},\mu_{\mathsf{c}}^{(2)}\\ \text{s.t (D5) holds}}}\Bigg\{\sum_{y_1}\Bigg[\sum_{s_2}\min_{x_2}\sum_{y_2}\mu_{\mathsf{c}}^{(2)}(s_2,x_2,y_1,y_2)\\
&\quad+\sum_{y_2}\min_{\widehat{s}_1,\widehat{s}_2}\sum_{s_2}\mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)\Bigg]\Bigg\}+ \qquad (40)\\
&\max_{\substack{\mu_{\mathsf{s}}^{(1)},\mu_{\mathsf{c}}^{(1)}\\ \text{s.t (D6) holds}}}\Bigg\{\sum_{y_2}\Bigg[\sum_{s_1}\min_{x_1}\sum_{y_1}\mu_{\mathsf{c}}^{(1)}(s_1,x_1,y_1,y_2)\\
&\quad+\sum_{y_1}\min_{\widehat{s}_1,\widehat{s}_2}\sum_{s_1}\mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2)\Bigg]\Bigg\}+\\
&\max_{\substack{\mu_{\mathsf{c}}^{(2|1)},\mu_{\mathsf{c}}^{(1|2)}\\ \text{s.t (D7) holds}}}\Bigg\{\sum_{s_1}\min_{x_1}\sum_{s_2}\mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2)\\
&\quad+\sum_{s_2}\min_{x_2}\sum_{s_1}\mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2)\Bigg\}\Bigg\}.
\end{aligned}$$

Note that (40) now has an outer optimization over $\lambda_{\mathsf{s}}^{(1|2)},\lambda_{\mathsf{s}}^{(2|1)},\lambda_{\mathsf{c}}$ satisfying the error density bottleneck (D4) and three inner optimization problems over each of the pairs, $(\mu_{\mathsf{s}}^{(1)},\mu_{\mathsf{c}}^{(1)})$, $(\mu_{\mathsf{s}}^{(2)},\mu_{\mathsf{c}}^{(2)})$ and $(\mu_{\mathsf{c}}^{(1|2)},\mu_{\mathsf{c}}^{(2|1)})$ with bottlenecks imposed by constraints (D6), (D5) and (D7) respectively.

We further lower bound (40) by restricting the choice of $\lambda_{\mathsf{s}}^{(1|2)}$ and $\lambda_{\mathsf{s}}^{(2|1)}$ such that $\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}$ is independent of $\widehat{s}_1$ and $\sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)}$ is independent of $\widehat{s}_2$. Under this assumption, constraints (D5) and (D6) imply that $\mu_{\mathsf{s}}^{(2)}$ and $\mu_{\mathsf{s}}^{(1)}$ are independent of $\widehat{s}_1$ and $\widehat{s}_2$, respectively. Hence, for each $y_1\in\mathcal{Y}_1$, $\sum_{s_2}\min_{x_2}\sum_{y_2}\mu_{\mathsf{c}}^{(2)}(s_2,x_2,y_1,y_2)+\sum_{y_2}\min_{\widehat{s}_1,\widehat{s}_2}\sum_{s_2}\mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)$ represents the objective corresponding to the packing of source flow $\mu_{\mathsf{s}}^{(2)}$ and channel flow $\mu_{\mathsf{c}}^{(2)}$ through the path $S_2\to X_2\to Y_2\to\widehat{S}_2$ satisfying the bottleneck, $\mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)+\mu_{\mathsf{c}}^{(2)}(s_2,x_2,y_1,y_2)\leq\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2)$ for all $s_2,x_2,y_2,\widehat{s}_2$. Taking the maximum over $\mu_{\mathsf{s}}^{(2)},\mu_{\mathsf{c}}^{(2)}$ inside the summation over y_1, we can express the optimal packing of these flows as $\mathrm{OPT}(\mathrm{DP},\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2))$, defined as in (1). Note that here, the RHS of the bottleneck is not necessarily an error density, but a function of $(s_2,x_2,y_1,y_2,\widehat{s}_2)$ and $\mathrm{OPT}(\mathrm{DP},\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2))$ is not necessarily the relaxation of a source coding problem.

Similarly, for each $y_2\in\mathcal{Y}_2$, $\sum_{s_1}\min_{x_1}\sum_{y_1}\mu_{\mathsf{c}}^{(1)}(s_1,x_1,y_1,y_2)+\sum_{y_1}\min_{\widehat{s}_1,\widehat{s}_2}\sum_{s_1}\mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2)$ represents the objective corresponding to the packing of source flow $\mu_{\mathsf{s}}^{(1)}$ and the channel flow $\mu_{\mathsf{c}}^{(1)}$ through the path $S_1\to X_1\to Y_1\to\widehat{\mathcal{S}}_1$ satisfying the bottleneck imposed by (D6). The resultant optimal packing can be expressed as $\mathrm{OPT}(\mathrm{DP},\sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2))$. Employing these yields the following lower bound on (40),

$$\begin{aligned}
&\max_{\substack{\lambda_{\mathsf{s}}^{(1|2)},\lambda_{\mathsf{s}}^{(2|1)},\lambda_{\mathsf{c}}\\ \text{s.t (D4) holds}\\ \sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}\perp\widehat{s}_1,\\ \sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)}\perp\widehat{s}_2}}\Bigg\{\sum_{y_1}\mathrm{OPT}(\mathrm{DP},\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2))\\
&+\sum_{y_2}\mathrm{OPT}(\mathrm{DP},\sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2))\\
&+\max_{\substack{\mu_{\mathsf{c}}^{(1|2)},\mu_{\mathsf{c}}^{(2|1)}\\ \text{s.t (D7) holds}}}\Bigg\{\sum_{s_1}\min_{x_1}\sum_{s_2}\mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2)\\
&\qquad+\sum_{s_2}\min_{x_2}\sum_{s_1}\mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2)\Bigg\}\Bigg\}. \qquad (41)
\end{aligned}$$

Note that for a given choice of $\lambda_{\mathsf{s}}^{(1|2)},\lambda_{\mathsf{s}}^{(2|1)}$ the bound in (41) comprises of optimal value of the duals of point-to-point problems, $\mathrm{OPT}(\mathrm{DP},\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)})$ and $\mathrm{OPT}(\mathrm{DP},\sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)})$. However, the objective of these problems is necessarily source coding since RHS of (D5) (or (D6)) is not the source coding error density.

Thus, the bounds in (38), (40) and (41) illustrate a hierarchy of lower bounds on the optimal value of DPSW. We now show that the Miyake and Kanaya converse falls lower in this hierarchy than our converse. Considering the choice of flows as in the proof of Theorem 4.4 with

$$\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2)=-\Big[\phi^{(1|2)}(s_1,s_2)\mathbb{I}\{s_1=\widehat{s}_1\}\times\mathbb{I}\{x_2=y_2\}+\alpha\widehat{\phi}(s_1,s_2)\Big]\mathbb{I}\{(s_1,s_2)=(\widehat{s}_1,\widehat{s}_2)\},$$
$$\lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2)=-\Big[\phi^{(2|1)}(s_1,s_2)\mathbb{I}\{s_2=\widehat{s}_2\}\times\mathbb{I}\{x_1=y_1\}+(1-\alpha)\widehat{\phi}(s_1,s_2)\Big]\mathbb{I}\{(s_1,s_2)=(\widehat{s}_1,\widehat{s}_2)\},$$

for $\alpha\in(0,1)$, it is easy to verify that our metaconverse in (29) follows via the construction of a feasible solution to the optimization problem in (40). Note that here, $\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}$ in general *depends* on $\widehat{s}_1$ and $\sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)}$ *depends* on $\widehat{s}_2$, whereby this construction is not feasible for (41). On the other hand, we find that the derivation of the Miyake and Kanaya converse from (31) corresponds to the following choice of variables,

$$\begin{aligned}
\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2)&=-\mathbb{I}\{(s_1,s_2)=(\widehat{s}_1,\widehat{s}_2)\}\\
&\Bigg[\mathbb{I}\{y_2=x_2\}P_{S_2}(s_2)\frac{\exp(-\beta)}{M_1}\Bigg], \qquad (42)\\
\lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2)&=-\mathbb{I}\{(s_1,s_2)=(\widehat{s}_1,\widehat{s}_2)\}\\
&\Bigg[\mathbb{I}\{y_1=x_1\}P_{S_1}(s_1)\frac{\exp(-\beta)}{M_2}+\frac{\exp(-\beta)}{M_1M_2}\Bigg],
\end{aligned}$$

for $\beta>0$. In this case, $\sum_{s_1}\lambda_{\mathsf{s}}^{(1|2)}$ is independent of $\widehat{s}_1$ and $\sum_{s_2}\lambda_{\mathsf{s}}^{(2|1)}$ is independent of $\widehat{s}_2$. Hence, with the following

choice for the remaining dual variables satsifying (D4)–(D7),

$$\lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2)=\mathbb{I}\{(y_1,y_2)=(x_1,x_2)\}P_{S_1,S_2}(s_1,s_2)\times \mathbb{I}\Bigg\{P_{S_1,S_2}(s_1,s_2)\leq \max\Bigg\{P_{S_2}(s_2)\frac{\exp(-\beta)}{M_1}, P_{S_1}(s_1)\frac{\exp(-\beta)}{M_2}, \frac{\exp(-\beta)}{M_1M_2}\Bigg\}\Bigg\},$$

$$\mu_{\mathsf{c}}^{(2)}(s_2,x_2,y_1,y_2)=-\frac{\exp(-\beta)}{M_1}P_{S_2}(s_2)\mathbb{I}\{y_2=x_2\},$$
$$\mu_{\mathsf{c}}^{(1)}(s_1,x_1,y_1,y_2)=-\frac{\exp(-\beta)}{M_2}P_{S_1}(s_1)\mathbb{I}\{y_1=x_1\},$$
$$\mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2)=-\frac{\exp(-\beta)}{M_1M_2}\mathbb{I}\{s_1=\widehat{s}_1\},$$
$$\mu_{\mathsf{c}}^{(2|1)}(s_1,s_2,x_2)=\sum_{y_1,y_2}\lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2),$$

and $\mu_{\mathsf{s}}^{(2)},\mu_{\mathsf{c}}^{(1|2)}\equiv 0$, it becomes clear that the resulting Miyake-Kanaya converse follows as a lower bound on the lower level optimization problem in (41). This also implies that the converse of Miyake and Kanaya can be thought to be obtained by the λ's in DPSW as inducing source-coding like problems in the DP's in (41). On the other hand our metaconverse in (29) follows from a more complicated bound.

In summary, our metaconverse in (29) and the Miyake-Kanaya converse in (32) can be placed in the hierarchy of lower bounds as illustrated in Fig 6. Moreover, this hierarchy also provides structured avenues for obtaining tighter bounds on the finite blocklength Slepian-Wolf coding problem – by appropriately bounding optimization problems lying higher in the hierarchy in (38).

VI. Conclusion

We presented a new finite blocklength converse for the Slepian-Wolf coding problem which improves on the converse of Miyake and Kanaya. The converse was derived by employing the linear programming based framework discussed in [3]. The proposed framework was shown to imply new metaconverses for lossy source coding and lossless source coding with side information problems, and recover the tightest hypothesis testing based converse of Kostina and Verdú [5]. For finite blocklength Slepian-Wolf coding, a systematic approach was developed to synthesize new LP-based converses from those of lossless source coding problems with side information. By appropriately combining the metaconverses for these point-to-point problems, our metaconverse for Slepian-Wolf coding was derived.

VII. Appendices

Appendix A
Hypothesis Testing Based Converse for Lossy Source Coding

For a source S with distribution P_S, distortion function $d:\mathcal{S}\times\widehat{\mathcal{S}}\to[0,+\infty]$ and distortion level $\mathbf{d}$, the rate-distortion function is defined as

$$R_S(\mathbf{d})=\inf_{P_{\widehat{S}|S}:\mathbb{E}[d(S,\widehat{S})]\leq\mathbf{d}} I(S;\widehat{S}), \tag{43}$$

where the infimum is over $P_{\widehat{S}|S}\in\mathcal{P}(\widehat{\mathcal{S}}|\mathcal{S})$. Assume that the infimum in (43) is achieved by a unique $P_{\widehat{S}^*|S}$ and $\mathbf{d}_{\min}=\inf\{\mathbf{d}:R_S(\mathbf{d})<\infty\}$. The hypothesis testing based converse of Kostina and Verdú [5, Theorem 8] is then obtained as below.

Converse A.1 (KV-hypothesis testing): Consider problem SC with $\mathcal{X}=\mathcal{Y}=\{1,\ldots,M\}$. Any code (f,g) such that $\mathbb{E}[\mathbb{I}\{d(S,\widehat{S})>\mathbf{d}\}]\leq\epsilon$ and $\mathbf{d}>\mathbf{d}_{\min}$ must satisfy,

$$M\geq\sup_{Q\in\mathcal{P}(\mathcal{S})}\inf_{\widehat{s}\in\widehat{\mathcal{S}}}\frac{\beta_{1-\epsilon}(P_S,Q)}{M\mathbb{Q}[d(S,\widehat{s})\leq\mathbf{d}]}, \tag{44}$$

where $\beta_{1-\epsilon}(P_S,Q)$ is the minimum type-II error $\sum_s Q(s)T(s)$ over all tests T such that the type-I error, $\sum_s P(s)(1-T(s))\leq\epsilon$. Moreover, the converse in (44) is equivalent to the following lower bound on the probability of error (see [14, Equation 72]),

$$\epsilon\geq\sup_{Q_S\in\mathcal{P}(\mathcal{S})}\alpha_{M^*}(P_S,Q_S), \tag{45}$$

with $M^*=M\max_{\widehat{s}}\mathbb{Q}[d(S,\widehat{s})\leq\mathbf{d}]$.

Corollary A.1: The following relationship holds,

$$\sup_{Q_S\in\mathcal{P}(\mathcal{S})}\Big\{\alpha_{M^*}(P_S,Q_S)\Big\} = \sup_{Q_S\in\mathcal{P}(\mathcal{S})}\sup_{\beta\geq0}\Bigg\{\sum_s\min\{P_S(s),\beta Q_S(s)\}-\beta M^*\Bigg\}. \tag{46}$$

Proof : To see the above equivalence, we consider the Neyman-Pearson (NP) optimal test for $\alpha_{M^*}(P_S,Q_S)$. The NP optimal test is given by $T^*(s)=\mathbb{I}\{\frac{P_S}{Q_S}(s)\leq\gamma^*\}$ such that

$$\alpha_{M^*}(P_S,Q_S)=\sum_s P_S(s)\mathbb{I}\Bigg\{\frac{P_S}{Q_S}(s)\leq\gamma^*\Bigg\}\quad\text{and} \tag{47}$$

$$\sum_s Q_S(s)\mathbb{I}\Bigg\{\frac{P_S}{Q_S}(s)\leq\gamma^*\Bigg\}=1-M^*. \tag{48}$$

Now, consider $\alpha_{M^*}(P_S,Q_S)-\gamma^*(1-M^*)$ which evaluates to

$$\begin{aligned}&\sum_s P_S(s)\mathbb{I}\Bigg\{\frac{P_S}{Q_S}(s)\leq\gamma^*\Bigg\}-\gamma^*\sum_s Q(s)\mathbb{I}\Bigg\{\frac{P_S}{Q_S}(s)\leq\gamma^*\Bigg\}\\&=\sum_s P_S(s)\mathbb{I}\Bigg\{\frac{P_S}{Q_S}(s)\leq\gamma^*\Bigg\}-\gamma^*\\&\quad+\gamma^*\sum_s Q_S(s)\mathbb{I}\Bigg\{\frac{P_S}{Q_S}(s)>\gamma^*\Bigg\}\\&=\sum_s\min\{P_S(s),\gamma^*Q(s)\}-\gamma^*,\end{aligned}$$

which implies that,

$$\alpha_{M^*}(P_S,Q_S)=\sum_s\min\{P_S(s),\gamma^*Q(s)\}-\gamma^*M^*.$$

Moreover, the RHS of the above equality can be equivalently written as,

$$\begin{aligned}&\sum_s\min\{P_S(s),\gamma^*Q(s)\}-\gamma^*M^*\\&=\sup_{\beta\geq0}\Bigg\{\sum_s\min\{P_S(s),\beta Q(s)\}-\beta M^*\Bigg\}.\end{aligned}$$

The proof for the last equality follows in the same line as in the proof of [15, Lemma 1] and we skip the proof here. Now, taking the supremum over $Q_S \in \mathcal{P}(\mathcal{S})$ yields the required result. ■

Appendix B
Proofs of Theorems in Section IV

Proof of Proposition 4.1: Let $(\bar{\gamma}^a, \bar{\gamma}^b, \bar{\lambda}_{\mathsf{s}}^{(1|2)}, \bar{\lambda}_{\mathsf{c}}^{(1|2)}) \in \mathrm{FEA}(\mathrm{DPSI}_{1|2})$. We now show that the choice of dual variables in (26) is feasible for DPSW. We first verify the feasibility of the choice of dual variables with respect to constraint (D1) of DPSW. We get that,

$$\sum_{y_1,y_2} \mu_{\mathsf{c}}^{(1)}(s_1,x_1,y_1,y_2) + \sum_{s_2} \mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2)$$
$$= \sum_{s_2}\sum_{y_1} \bar{\lambda}_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2,y_1) \overset{(c)}{\geq} \bar{\gamma}^a(s_1) = \gamma^a(s_1),$$

thereby satisfying (D1). The inequality in (c) follows from the constraint (B1) of $\mathrm{DPSI}_{1|2}$. For checking feasibility with respect to constraint (D2), we get that

$$\sum_{y_1,y_2} \mu_{\mathsf{c}}^{(2)}(x_2,s_2,y_1,y_2) + \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2)$$
$$= \sum_{y_1,y_2} \bar{\gamma}^b(s_2,y_1)\mathbb{I}\{x_2=y_2\} = \sum_{y_1} \bar{\gamma}^b(s_2,y_1) = \gamma^b(s_2),$$

thereby satisyfing (D2). The feasibility with respect to (D3) is trivially satisfied. For feasibility with respect to (D4), the LHS of (D4) becomes

$$\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2) + \lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2)$$
$$= \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1,s_2,\widehat{s}_1,y_1)\mathbb{I}\{x_2=y_2\}$$
$$+ \bar{\lambda}_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2,y_1)\mathbb{I}\{x_2=y_2\}$$
$$\overset{(a)}{\leq} P(s_1,s_2)\mathbb{I}\{y_1=x_1\}\mathbb{I}\{y_2=x_2\}\mathbb{I}\{s_1\neq\widehat{s}_1\}$$
$$\leq P(s_1,s_2)\mathbb{I}\{(y_1,y_2)=(x_1,x_2)\}\mathbb{I}\{(s_1,s_2)\neq(\widehat{s}_1,\widehat{s}_2)\},$$

which is the RHS, thereby satisfying (D4). Here, the inequality in (a) follows from the constriant (B3) of $\mathrm{DPSI}_{1|2}$. To verify feasibility with respect to (D5), we have, $\mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2) + \mu_{\mathsf{c}}^{(2)}(x_2,s_2,y_1,y_2) =$

$$\bar{\gamma}^b(s_2,y_1)\mathbb{I}\{x_2=y_2\} \overset{(b)}{\leq} \sum_{s_1} \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1,s_2,\widehat{s}_1,y_1)\mathbb{I}\{x_2=y_2\}$$
$$= \sum_{s_1} \lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2),$$

which is the RHS of (D5), thereby satisfying it. The inequality in (b) follows from constraint (B2) of $\mathrm{DPSI}_{1|2}$. Since $\lambda_{\mathsf{s}}^{(2|1)}, \mu_{\mathsf{s}}^{(1)}, \mu_{\mathsf{c}}^{(1)} \equiv 0$, the constraint (D6) is trivially satisfied. To verify feasibility with respect to (D7), we have, $\mu_{\mathsf{c}}^{(2|1)}(x_2,s_1,s_2) + \mu_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2) =$

$$\sum_{y_1} \bar{\lambda}_{\mathsf{c}}^{(1|2)}(x_1,s_1,s_2,y_1) = \sum_{y_1,y_2} \bar{\lambda}_{\mathsf{c}}^{(1)}(x_1,s_1,s_2,y_1)\mathbb{I}\{x_2=y_2\}$$
$$= \sum_{y_1,y_2} \lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2),$$

thereby satisfying (D7). Thus, the considered choice of dual variables is feasible for DPSW. ■

Proof of Theorem 4.3: It is enough to show that the above choice of dual variables are feasible with respect to the constraints (D4)-(D7) of DPSW. To verify the feasibility of dual variables with respect to (D4), consider the following two cases.

Case 1: $\mathbb{I}\{(s_1,s_2)\neq(\widehat{s}_1,\widehat{s}_2)\}=1$.

In this case, $\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2) = 0$ and $\lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2) = 0$. The LHS of (D4) becomes, $\lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2) \leq P(s_1,s_2)\mathbb{I}\{(y_1,y_2)=(x_1,x_2)\}$, which is the RHS of (D4) thereby satisfying the constraint.

Case 2: $\mathbb{I}\{(s_1,s_2)\neq(\widehat{s}_1,\widehat{s}_2)\}=0$.

In this case, $s_1=\widehat{s}_1$, $s_2=\widehat{s}_2$, the RHS of (D4) is zero and the LHS becomes,

$$\lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2) + \lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_1)$$
$$+ \lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2)$$
$$= \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1,s_2,\widehat{s}_1,y_1)\mathbb{I}\{x_2=y_2\} + \alpha\widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)$$
$$+ \tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1,s_2,\widehat{s}_2,y_2)\mathbb{I}\{x_1=y_1\} + \lambda_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2)$$
$$+ (1-\alpha)\widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)$$
$$\leq [\bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1,s_2,\widehat{s}_1,y_1) + \bar{\lambda}_{\mathsf{c}}^{(1|2)}(s_1,s_2,x_1,y_1)]\mathbb{I}\{x_2=y_2\}$$
$$+ [\tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1,s_2,\widehat{s}_2,y_2) + \tilde{\lambda}_{\mathsf{c}}^{(2|1)}(s_1,s_2,x_2,y_2)]\mathbb{I}\{x_1=y_1\}$$
$$+ \widehat{\lambda}_{\mathsf{c}}(s_1,s_2,x_1,x_2,y_1,y_2) + \widehat{\lambda}_{\mathsf{s}}(s_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)$$

which is non-positive, thereby satisfying the constraint (D4). The non-positivity follows since $(\bar{\lambda}_{\mathsf{s}}^{(1|2)}, \bar{\lambda}_{\mathsf{c}}^{(1|2)})$, $(\tilde{\lambda}_{\mathsf{s}}^{(2|1)}, \tilde{\lambda}_{\mathsf{c}}^{(2|1)})$ and $(\widehat{\lambda}_{\mathsf{s}}, \widehat{\lambda}_{\mathsf{c}})$ satisfy the constraints (B3), (C3) and (A3) (corresponding to the case when $(s_1,s_2)=(\widehat{s}_1,\widehat{s}_2)$) of dual programs $\mathrm{DPSI}_{1|2}$, $\mathrm{DPSI}_{2|1}$ and DPJE respectively.

To verify feasibility with respect to (D5), $\sum_{s_1} \lambda_{\mathsf{s}}^{(1|2)}(s_1,s_2,x_2,y_1,y_2,\widehat{s}_1,\widehat{s}_2)$ evaluates to

$$\left[\bar{\lambda}_{\mathsf{s}}^{(1|2)}(\widehat{s}_1,s_2,\widehat{s}_1,y_1)\mathbb{I}\{x_2=y_2\} + \alpha\widehat{\lambda}_{\mathsf{s}}(\widehat{s}_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)\right]$$
$$\times \mathbb{I}\{s_2=\widehat{s}_2\}$$
$$\overset{(a)}{\geq} [\bar{\gamma}^b(s_2,y_1) - \sum_{s_1\neq\widehat{s}_1} \bar{\lambda}_{\mathsf{s}}^{(1|2)}(s_1,s_2,\widehat{s}_1,y_1)]\mathbb{I}\{x_2=y_2,s_2=\widehat{s}_2\}$$
$$+ \alpha\widehat{\lambda}_{\mathsf{s}}(\widehat{s}_1,s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)\mathbb{I}\{s_2=\widehat{s}_2\}$$
$$\geq \mu_{\mathsf{c}}^{(2)}(x_2,s_2,y_1,y_2) + \mu_{\mathsf{s}}^{(2)}(s_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2),$$

thereby satisfying (D5). The inequality in (a) results from the constraint (A2).

To verify the feasibility with respect to (D6), $\sum_{s_2} \lambda_{\mathsf{s}}^{(2|1)}(s_1,s_2,x_1,y_1,y_2,\widehat{s}_1,\widehat{s}_2)$ evaluates to

$$\Bigg[\tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1,\widehat{s}_2,\widehat{s}_2,y_2)\mathbb{I}\{x_1=y_1\}$$
$$+ (1-\alpha)\widehat{\lambda}_{\mathsf{s}}(s_1,\widehat{s}_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)\Bigg]\mathbb{I}\{s_1=\widehat{s}_1\},$$
$$\overset{(b)}{\geq} \left[\tilde{\gamma}^b(s_1,y_2) - \sum_{s_2\neq\widehat{s}_2} \tilde{\lambda}_{\mathsf{s}}^{(2|1)}(s_1,s_2,\widehat{s}_2,y_2)\right]\mathbb{I}\{x_1=y_1,s_1=\widehat{s}_1\}$$
$$+ (1-\alpha)\widehat{\lambda}_{\mathsf{s}}(s_1,\widehat{s}_2,\widehat{s}_1,\widehat{s}_2,y_1,y_2)\mathbb{I}\{s_1=\widehat{s}_1\}$$
$$\geq \mu_{\mathsf{c}}^{(1)}(x_1,s_1,y_1,y_2) + \mu_{\mathsf{s}}^{(1)}(s_1,\widehat{s}_1,\widehat{s}_2,y_1,y_2),$$

thereby satisfying (D6). The inequality in (b) follows from the constraint (C2) of $\mathrm{DPSI}_{2|1}$. The feasibility with respect to (D7) is trivially satisfied. Hence, the considered choice of dual variables are feasible for DPSW. ■

Proof of Theorem 4.4: To get to the above converse, take $(\bar{\lambda}_{\mathsf{s}}^{(1|2)}, \bar{\lambda}_{\mathsf{c}}^{(1|2)}, \bar{\gamma}^a, \bar{\gamma}^b)$ as in (11), $(\tilde{\lambda}_{\mathsf{s}}^{(2|1)}, \tilde{\lambda}_{\mathsf{c}}^{(2|1)}, \tilde{\gamma}^a, \tilde{\gamma}^b)$ as in (15) and $(\widehat{\lambda}_{\mathsf{s}}, \widehat{\lambda}_{\mathsf{c}}, \widehat{\gamma}^a, \widehat{\gamma}^b)$ as in (9) and substitute in (28) to get the values of the variables $\lambda_{\mathsf{s}}^{(1|2)}, \lambda_{\mathsf{s}}^{(2|1)}, \mu_{\mathsf{s}}^{(2)}$, and $\mu_{\mathsf{s}}^{(1)}$ of DPSW. For the remaining variables, choose the following values of dual variables,

$$
\begin{aligned}
&\lambda_{\mathsf{c}}(s_1, s_2, x_1, x_2, y_1, y_2) = \mathbb{I}\{(y_1, y_2) = (x_1, x_2)\}\times \\
&\min\{P_{S_1,S_2}(s_1, s_2), \widehat{\phi}(s_1, s_2) + \phi^{(1|2)}(s_1, s_2) + \phi^{(2|1)}(s_1, s_2)\}, \\
&\mu_{\mathsf{c}}^{(2)}(s_2, x_2, y_1, y_2) = -\max_{\widehat{s}_1} \phi^{(1|2)}(\widehat{s}_1, s_2)\mathbb{I}\{x_2 = y_2\}, \\
&\mu_{\mathsf{c}}^{(1)}(s_1, x_1, y_1, y_2) = -\max_{\widehat{s}_2} \phi^{(2|1)}(s_1, \widehat{s}_2)\mathbb{I}\{x_1 = y_1\}, \\
&\mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2) \equiv \min\{P_{S_1,S_2}(s_1, s_2), \\
&\qquad \widehat{\phi}(s_1, s_2) + \phi^{(1|2)}(s_1, s_2) + \phi^{(1|2)}(s_1, s_2)\}, \\
&\gamma^c(y_1, y_2) = -\max_{\widehat{s}_1, \widehat{s}_2} \widehat{\phi}(\widehat{s}_1, \widehat{s}_2), \\
&\gamma^b(s_2) = -M_1 \max_{\widehat{s}_1} \phi^{(1|2)}(\widehat{s}_1, s_2) + \\
&\qquad \sum_{s_1} \mu_{\mathsf{c}}^{(2|1)}(x_2, s_1, s_2), \\
&\gamma^a(s_1) = -M_2 \max_{\widehat{s}_2} \phi^{(2|1)}(s_1, \widehat{s}_2),
\end{aligned}
\tag{49}
$$

$\mu_{\mathsf{c}}^{(1|2)}(x_1, s_1, s_2) \equiv 0$. The above choice of variables can be easily verified to satisfy the constraints in (28). ■

References

[1] S. T. Jose and A. A. Kulkarni, "Linear programming based converses for some network-like problems," in *to appear in the Proceedings of Information Theory Workshop*, 2017.

[2] S. Miyake and F. Kanaya, "Coding theorems on correlated general sources," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 78, no. 9, pp. 1063–1070, 1995.

[3] S. T. Jose and A. A. Kulkarni, "Linear programming based converses for finite blocklength lossy joint source-channel coding," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7066–7094, November, 2017.

[4] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[5] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime," *Information Theory, IEEE Transactions on*, vol. 58, no. 6, pp. 3309–3338, 2012.

[6] ——, "Lossy joint source-channel coding in the finite blocklength regime," *Information Theory, IEEE Transactions on*, vol. 59, no. 5, pp. 2545–2575, 2013.

[7] T. Han, "Information-spectrum methods in information theory [english translation]. series: Stochastic modelling and applied probability, vol. 50," *Springer*, vol. 1, no. 6, pp. 3–1, 2003.

[8] A. A. Kulkarni and T. P. Coleman, "An optimizer's approach to stochastic control problems with nonclassical information structure," *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 937–949, 2015.

[9] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer programming*. Springer, 2014, vol. 271.

[10] C. Villani, *Optimal transport: old and new*. Springer, 2008, vol. 338.

[11] L. Palzer and R. Timo, "A converse for lossy source coding in the finite blocklength regime," 2016.

[12] V. Y. Tan *et al.*, "Asymptotic estimates in information theory with non-vanishing error probabilities," *Foundations and Trends® in Communications and Information Theory*, vol. 11, no. 1-2, pp. 1–184, 2014.

[13] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.

[14] G. Vazquez-Vilar, A. T. Campo, A. G. i Fàbregas, and A. Martinez, "Bayesian m-ary hypothesis testing: The meta-converse and verdú-han bounds are tight," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2324–2333, 2016.

[15] N. Elkayam and M. Feder, "On the calculation of the meta-converse," *CoRR*, vol. abs/1512.09333, 2015. [Online]. Available: http://arxiv.org/abs/1512.09333