Abstract—A new finite blocklength converse for the Slepian–Wolf coding problem, which significantly improves on the best-known converse due to Miyake and Kanaya, is presented. To obtain this converse, an extension of the linear programming (LP)-based framework for finite blocklength point-to-point coding problems is employed. However, a direct application of this framework demands a complicated analysis for the Slepian–Wolf problem. An analytically simpler approach is presented. To obtain this converse, an extension of the linear programming, duality.

Index Terms—Slepian-wolf coding, finite blocklength converses, linear programming, duality.

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 ([3]–[6]), 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.

Fig. 1. Two-user joint source-channel coding 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 following optimization problem,

\[
\min_{f_1,f_2,g} P\{(S_1, S_2) \neq (\hat{S}_1, \hat{S}_2)\}
\]

\[
\text{s.t. } \begin{align*}
X_1 &= f_1(S_1), \\
X_2 &= f_2(S_2), \\
(\hat{S}_1, \hat{S}_2) &= g(Y_1, Y_2),
\end{align*}
\]

where \(S_1, S_2, X_1, X_2, Y_1, Y_2, \hat{S}_1, \hat{S}_2\) are discrete random variables taking values in fixed, finite spaces \(S_1, S_2, X_1, X_2, Y_1, Y_2, \hat{S}_1, \hat{S}_2\), 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 separately encoded by functions \(f_1 : S_1 \rightarrow X_1\) and \(f_2 : S_2 \rightarrow X_2\) 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 ‘.’ is true and is zero otherwise. \((Y_1, Y_2)\) is then jointly decoded by \(g : Y_1 \times Y_2 \rightarrow \hat{S}_1 \times \hat{S}_2\) to obtain the output signal \((\hat{S}_1, \hat{S}_2) = g(Y_1, Y_2)\).

For the finite blocklength Slepian-Wolf coding problem, we note that spaces \(S_1 = \hat{S}_1, S_2 = \hat{S}_2, X_1 = \hat{Y}_1 = \{1, \ldots, M_1\}\) and \(X_2 = \hat{Y}_2 = \{1, \ldots, M_2\}\, M_1, M_2 \in \mathbb{N}\). An error in transmission occurs when \((S_1, S_2) \neq (\hat{S}_1, \hat{S}_2)\). 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 et al. [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 extend this theme towards the Slepian-Wolf coding problem. 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 clean finite blocklength converses in the point-to-point setting for lossy source-coding problems with side-information at the decoder; without side-information these converses are equivalent to the hypothesis testing-based converse of Kostina and Verdú [5, Th. 8], stronger than our earlier converses in [3], and imply the tilted information based converse of Kostina and Verdú [5, Th. 7] and the converse of Palzer and Timo [11, Th. 1], and are, to the best of our knowledge, the strongest known. Moreover, for the lossless case, they exactly equal the finite blocklength probability of error. 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 asymptotically 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 LP-based converse results, which improves on the converse of Miyake and Kanaya.

Multiterminal problems have various types of challenges associated with them. One fundamental challenge is that of extending results and techniques from point-to-point information theory to multiterminal settings in a systematic manner. This includes, amongst other problems, discovering the right quantities and inequalities to describe optimal rate regions. Another challenge, is that existing rate regions of several problems involve auxiliary random variables that are not easy to eliminate. The systematic synthesis of converses from point-to-point converses obtained via the LP approach, illustrated in this paper for the Slepian-Wolf problem is evidence that the LP approach helps tackle the first challenge. Moreover, the approach naturally suggests improvements to converses, which can come in handy in second-order analysis. And the lessons learnt from applying the approach to a “solved” problem such as Slepian-Wolf coding, could be applied to other unsolved multiterminal problems. This is part of our ongoing effort.

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, $S, T$ respectively; lower case letters $a, b$ represent the specific values these random variables take. $\mathbb{I}[\cdot]$ denotes the indicator function which is equal to one when its argument is true and is zero otherwise. $\mathbb{P}[\cdot]$ denotes the set of all probability distributions on $\cdot$. $Q$ represents a specific distribution. If $Q$ is a joint probability distribution, let $Q_{\cdot|\cdot}$ denote the marginal distribution of $\cdot$. For example, $Q_{X|S}$ represents the vector with $Q_{X|S}(x|s)$ for $x \in X, s \in S$ 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 $\text{OPT}(P)$ represents its optimal value and $\text{FEA}(P)$ represents its feasible region. LHS stands for Left Hand Side and RHS stands for Right Hand Side. The notation $a \perp b$ 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.

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,

\[
\text{SC:} \quad \min_{f, g} \mathbb{E}[\mathbb{I}[d(S, \hat{S}) > \Delta]] \quad \text{s.t.} \quad X = f(S), \quad \hat{S} = g(Y).
\]

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

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

\[
\text{SC:} \quad \min_{Q_{X|S}, Q_{\hat{S}|Y}} \sum_{x, \hat{s}} \mathbb{I}[d(s, \hat{s}) > \Delta] \sum_{x, y} Q(s, x, y, \hat{s})
\]

\[
Q(\tilde{z}) = P_S Q_{X|S} Q_{Y|X} Q_{\hat{S}|Y}(\tilde{z}),
\]

\[
\text{s.t.} \quad \sum_{x} Q_{X|S}(x|s) = 1 \quad \forall s \in S,
\]

\[
\sum_{\hat{s}} Q_{\hat{S}|Y}(\hat{y}|s) = 1 \quad \forall y \in Y,
\]

\[
Q_{X|S}(x|s) \geq 0 \quad \forall s \in S, x \in X,
\]

\[
Q_{\hat{S}|Y}(\hat{y}|s) \geq 0 \quad \forall \hat{s} \in \hat{S}, y \in Y.
\]

where $\tilde{z} := (x, s, y, \hat{s}) \in \tilde{Z}$, $\hat{z} := S \times X \times Y \times \hat{S}$ and $P_S Q_{X|S} Q_{Y|X} Q_{\hat{S}|Y}(\tilde{z}) = P_S(s) Q_{X|S}(x|s) Q_{Y|X}(y|x) Q_{\hat{S}|Y}(\hat{y}|s).$ Here, $Q_{X|S}$ represents a randomized encoder and $Q_{\hat{S}|Y}$ 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 $W(s, x, y, \hat{s}) = Q_{X|S}(x|s) Q_{\hat{S}|Y}(\hat{y}|s)$ and obtain valid constraints involving $W$ through the constraints of the problem. Specifically, multiply both sides of the constraint $\sum_{x} Q_{X|S}(x|s) = 1$ by $Q_{\hat{S}|Y}(\hat{y}|s)$ for all $s, y, \hat{s}$ and multiply both sides of $\sum_{\hat{s}} Q_{\hat{S}|Y}(\hat{y}|s) = 1$ 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 $Q_{X|S}, Q_{\hat{S}|Y}, W$, which together with $Q_{X|S} \in \mathcal{P}(X|S)$ and $Q_{\hat{S}|Y} \in \mathcal{P}(\hat{S}|Y)$ give the following LP relaxation.

\[
\text{LP:} \quad \min_{Q_{X|S}, Q_{\hat{S}|Y}, W} \sum_{\tilde{z}} \mathbb{I}[d(s, \hat{s}) > \Delta] \sum_{s} Q_{X|S}(x|s) Q_{Y|X}(y|x) W(\tilde{z})
\]

\[
\sum_{x} Q_{X|S}(x|s) = 1 : \gamma^a(s) \forall s
\]

\[
\sum_{y} Q_{\hat{S}|Y}(\hat{y}|x) = 1 : \gamma^b(y) \forall y
\]

\[
\sum_{s} W(\tilde{z}) - Q_{\hat{S}|Y}(\hat{y}|x) = 0 : \lambda_{s}(s, \hat{s}, y) \forall s, y, \hat{s}
\]

\[
\sum_{x} W(\tilde{z}) - Q_{X|S}(x|s) = 0 : \lambda_{x}(x, s, y) \forall x, s, y
\]

\[
Q_{X|S}, Q_{\hat{S}|Y}, W \geq 0.
\]

Above $\gamma^a, \gamma^b, \lambda_s$ and $\lambda_x$ 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 $s \in S$ and $y \in Y$, $W(s, x, y, \hat{s})$ is a coupling on $X \times \hat{S}$ between the marginals $Q_{X|S}(x|s)$ and $Q_{\hat{S}|Y}(\hat{y}|s)$; let the set of such $W$ be denoted by $\Xi(Q_{X|S}, Q_{\hat{S}|Y})$. The LP relaxation of SC is a nested minimization – the inner minimization is over all couplings $W \in \Xi(Q_{X|S}, Q_{\hat{S}|Y})$ and the outer minimization is over all randomized codes $(Q_{X|S}, Q_{\hat{S}|Y})$:

\[
\min_{Q_{X|S}, Q_{\hat{S}|Y}} \min_{W \in \Xi(Q_{X|S}, Q_{\hat{S}|Y})} \sum_{\tilde{z}} \mathbb{I}[d(s, \hat{s}) > \Delta] P_S Q_{Y|X} W(\tilde{z}).
\]
The original problem SC has the outer minimization over codes, but in place of the inner minimization over $W$ it employs the product $Q_{X|S}(\cdot)Q_{\bar{S}|Y}(\cdot) \in \mathcal{Z}(Q_{X|S}, Q_{\bar{S}|Y})$ to obtain the distribution. Thus the LP relaxation is arrived at by considering the term $Q_{X|S}(\cdot)Q_{\bar{S}|Y}(\cdot)$ in SC as an element of $\mathcal{Z}(Q_{X|S}, Q_{\bar{S}|Y})$ and minimizing the resulting cost over all elements of $\mathcal{Z}(Q_{X|S}, Q_{\bar{S}|Y})$. 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-A.1.

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

$$\text{DP} \quad \max_{\gamma^a, \gamma^b, \lambda_c, \lambda_s} \sum_s \gamma^a(s) + \sum_y \gamma^b(y)$$

$$\gamma^a(s) - \sum_y \lambda_c(x, s, y) \leq 0 \quad \forall x, s \quad (P1)$$

$$\gamma^b(y) - \sum_s \lambda_c(s, \bar{s}, y) \leq 0 \quad \forall \bar{s}, y \quad (P2)$$

$$\lambda_c(s, \bar{s}, y) + \lambda_s(s, \bar{s}, y) \leq \Sigma(\bar{s}) \forall \bar{s} \quad (P3)$$

where $\Sigma(\bar{s}) = I(d(s, \bar{s}) > \Delta)P_S(\bar{s}|y = x)$ for all $\bar{s}$, since $P_{Y|X}(y|x) = \mathbb{I}[y = x]$.

Since the objective of DP is to maximize summation over $s \in \mathcal{S}$ of $\gamma^a(s)$ and summation over $y \in \mathcal{Y}$ of $\gamma^b(y)$, it is optimal to choose $\gamma^a(s)$ and $\gamma^b(y)$ such that (P1) and (P2) hold with equality, i.e., $\gamma^a(s) = \min_s \sum_y \lambda_c(x, s, y)$ and $\gamma^b(y) = \min_y \sum_s \lambda_c(s, \bar{s}, y)$. Then the optimal value of DP with $\Sigma(\bar{s})$ as the RHS of (P3) evaluates to,

$$\text{OPT(DP, } \Sigma(\bar{s})\text{)} = \max_{\lambda_c, \lambda_s} \left\{ \sum_s \min_x \sum_y \lambda_c(x, s, y) + \sum_y \min_z \sum_s \lambda_s(s, \bar{s}, y) \right\}$$

$$\text{s.t. } \lambda_c(x, s, y) + \lambda_s(s, \bar{s}, y) \leq \Sigma(\bar{s}) \forall \bar{s}$$

(1)

It follows that if we construct functions $\lambda_c : \mathcal{S} \times \mathcal{S} \times \mathcal{Y} \rightarrow \mathbb{R}$ and $\lambda_s : \mathcal{S} \times \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ satisfying (1), then linear programming duality implies the following lower bound on OPT(SC),

$$\text{OPT(SC)} \geq \text{OPT(LP)} = \text{OPT(DP)}$$

$$\sum_s \min_x \sum_y \lambda_c(x, s, y) + \sum_y \min_z \sum_s \lambda_s(s, \bar{s}, y).$$

(2)

Notice that $\lambda_s$ and $\lambda_c$ are functions on subspaces of $\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{S}$. $\lambda_c$ 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, $\lambda_s$ is a function of the source signal $s$, the decoder input $y$ and decoder output $\bar{s}$. 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”, $\Sigma(\bar{s}) = P_S(\bar{s}|y = x)\mathbb{I}[d(s, \bar{s}) > \Delta]$, 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 LP-based converse for lossy source coding, which implies our improvement on the Kostina-Verdú converse and is equivalent to the hypothesis testing based converse in [5, Th. 8].

**Theorem 1 (LP-Based Converse for Lossy Source Coding):** Consider the problem SC. For any code, the following bound holds,

$$\mathbb{E}[\mathbb{I}(d(s, \bar{s}) > \Delta)] \geq \text{OPT(SC)} \geq \text{OPT(LP)} = \text{OPT(DP)}$$

$$\geq \sup_{0 \leq \phi(s) \leq P_S(s)} \left\{ \sum_s \phi(s) - M \max_s \sum_s \phi(s)\mathbb{I}[d(s, \bar{s}) \leq \Delta] \right\},$$

(3)

$$\geq \sup_{z(s) \geq 0} \left\{ \sum_s \min_s \{P_S(s), z(s)\} \right\} - M \max_s \sum_s \min_s \{P_S(s), z(s)\}\mathbb{I}[d(s, \bar{s}) \leq \Delta].$$

(4)

where the supremum in (3) is over all functions $\phi : \mathcal{S} \rightarrow [0, 1]$ such that $0 \leq \phi(s) \leq P_S(s)$ for all $s \in \mathcal{S}$ and in (4) is over all functions $z : \mathcal{S} \rightarrow [0, \infty]$.

**Proof:** To obtain the converse in (3), consider the following values of source and channel flows of DP,

$$\lambda_c(x, s, y) \equiv \mathbb{I}[y = x]\phi(s),$$

$$\lambda_s(s, \bar{s}, y) \equiv -\phi(s)\mathbb{I}[d(s, \bar{s}) \leq \Delta].$$

(5)

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

**Case 1:** $\mathbb{I}(d(s, \bar{s}) > \Delta) = 1$.

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

**Case 2:** $\mathbb{I}(d(s, \bar{s}) > \Delta) = 0$.

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 lower bound in (3) follows from (2) by taking supremum over $\phi$ such that $0 \leq \phi(s) \leq P_S(s) \forall s \in \mathcal{S}$.

To see the equivalence between (3) and (4), note that choosing $\phi(s) = \min_s \{P_S(s), z(s)\}$ in (3) where $z : \mathcal{S} \rightarrow [0, \infty]$, and taking the supremum over such $z$ lower bounds (3) to get (4). On the other hand, lower bounding (4) by choosing $z(s) = \phi^*(s)$, where $\phi^*(s)$ maximizes (3), gets to the lower bound in (3).

**Remark 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,

$$\Sigma(\bar{s}) \geq \phi(s)\mathbb{I}[y = x]\mathbb{I}[d(s, \bar{s}) > \Delta]$$

$$= \phi(s)\mathbb{I}[y = x] - \phi(s)\mathbb{I}[y = x]\mathbb{I}[d(s, \bar{s}) \leq \Delta].$$


An obvious choice of the flows would thus be \( \lambda_{\phi}(x, s, y) = \phi(s) \|\{y = x\} \) and \( \lambda_{\phi}(s, \hat{x}, y) \leq -\phi(s) \|\{y = x\} \|d(s, \hat{x}) \leq \Delta \), which results in our converse in (3).

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

**Corollary 2 (LP-Based Converse Recovers Improvement on Kostina-Verdú Converse From [3, Corollary 5.9]):** Consider the problem SC. Then, for any code,

\[
\mathbb{E}[\|d(S, \hat{S}) > \Delta\|] \geq \text{OPT}(SC) \geq \text{OPT}(LP) = \text{OPT}(DP)
\]

\[
\geq \sup_{\gamma} \left\{ \mathbb{P}[j_S(S, \Delta) \geq \gamma + \log M] + \frac{1}{M} \right\}
\times \sum SP(s) \exp(j_S(s, \Delta) - \gamma)\|j_S(s, \Delta) < \log M + \gamma\|
\left( - \max_{\gamma} \exp(-\gamma) \sum SP(s) \exp(j_S(s, \Delta))\|d(s, \hat{S}) \leq \Delta\| \right).
\]

**Proof:** To see this, take \( z(s) = \sum P(s) J_S(s, \Delta) - \gamma \) for any scalar \( \gamma \) and lower bound \(-M \sum \min_{\gamma} \{ P(s), z(s) \} \|d(s, \hat{S}) \leq \Delta\| \) in (4) with \(-M \sup_{\gamma} \sum z(s) \) \{d(s, \hat{S}) \leq \Delta\}. Subsequently, take supremum over \( \gamma \) to get the required bound.

Consider a binary hypothesis testing problem between distributions \( P_S \) and \( \hat{Q}_S \). Let \( a_0(P_S, Q_S) \) represent the minimum type-I error, \( \sum P_S(s) T(s) \) over all tests \( T \) such that the type-II error, \( \sum Q_S(s)(1 - T(s)) \) is at most \( \theta \). The following corollary then establishes the equivalence between the LP-based converse in Theorem 1 and the the hypothesis testing based converse of Kostina and Verdú [5, Th. 8].

**Corollary 3 (LP-Based Converse is Equivalent to Kostina-Verdú Hypothesis Testing Based Converse From [5, Th. 8]):** Consider the problem SC. The RHS of (4) equals

\[
\sup_{Q_S \in \mathbb{P}(S)} \sup_{\beta \geq 0} \left\{ \sum \min\{P_S(s), \beta Q_S(s)\} - \beta M^* \right\}
\]

(6)

\[
= \sup_{Q_S \in \mathbb{P}(S)} a_{M^*}(P_S, Q_S)
\]

(7)

Here, \( M^* = M \max_{\gamma} \sum d(s, \hat{S}) \leq \Delta \) and \( \mathbb{Q} \) is the measure on \( S \) induced by \( Q_S \).

**Proof:** The proof is included in Appendix A.

**Corollary 4 (LP-Based Converse Recovers Palzer-Timo Converse [11, Th. 1]):** Consider the problem SC. Then, for any code, the LP-based converse (4) satisfies

\[
\mathbb{E}[\|d(S, \hat{S}) > \Delta\|] \geq \text{OPT}(SC) \geq \text{OPT}(LP) = \text{OPT}(DP)
\]

\[
\geq \sup_{\beta \in \mathbb{R}} \left\{ \mathbb{P}[j_S(S, \Delta) > \beta]\right\}

- M \max \mathbb{P}[j_S(S, \Delta) > \beta, d(S, \hat{S}) \leq \Delta]\}
\]

**Proof:** To obtain this converse, take \( \zeta(s) = P_S(s)\|\{j_S(s, \Delta) > \beta\} \) in (4) where \( \beta \geq 0 \) and take supremum over \( \beta \geq 0 \). Notice that in this case, \( \min\{\zeta(s), P_S(s)\} = \zeta(s) \) since \( \zeta(s) = P_S(s)\|\{j_S(s, \Delta) > \beta\} \leq P_S(s) \).

The finite blocklength lossless data compression problem results from SC by setting \( d(S, \hat{S}) = \|S \neq \hat{S}\| \) and \( \Delta = 0 \).

The following corollary particularizes the LP-based converse in (3) to lossless data compression. In this case, the LP-based converse takes a particularly simple form.

**Corollary 5 (LP-Based Converse for Lossless Source Coding):** For lossless data compression, consider the problem SC with \( d(S, \hat{S}) = \|S \neq \hat{S}\| \) and \( \Delta = 0 \). Consequently, for any code, the following converse follows from Theorem 1,

\[
\mathbb{E}[\|S \neq \hat{S}\|] \geq \text{OPT}(SC) \geq \text{OPT}(LP)
\]

\[
= \sup_{0 \leq \gamma \leq \beta} \left\{ \|\phi_1 - M\|\phi_\infty \right\}
\]

(8)

**Proof:** To show the equivalence of (8) to OPT(SC), we first obtain a lower bound on (8). Towards this, consider space \( S \) with the source sequences arranged in order of decreasing probability, i.e., \( S = \{s_1, s_2, \ldots, s_M, s_{M+1}, \ldots, s_{|S|}\} \)

such that

\[
P_S(s_1) \geq \ldots \geq P_S(s_M) \geq \ldots \geq P_S(s_{|S|}).
\]

Consequently, choose

\[
\phi(s_i) \begin{cases} P_S(s_i) & \text{if } i \in \{M + 1, \ldots, |S|\}, \\
P_S(s_{M+1}) & \text{if } i \in \{1, \ldots, M\}.
\end{cases}
\]

The above choice of \( \phi \) in (8) subsequently yields the following lower bound,

\[
\text{OPT}(SC) \geq \sum_{i=|S|}^{M+1} P_S(s_i).
\]

Moreover, the above bound is in fact achievable by the following code for lossless source coding,

\[
Q_{S|x}(x|s_i) = \begin{cases} \|x = i\| & \text{if } i = 1, \ldots, M, \\
\|x = M\| & \text{if } i = M + 1, \ldots, |S|,
\end{cases}
\]

\[
Q_{S|y}(y|j) = \|y = s_j\|, \quad j = 1, \ldots, M,
\]

which yields that \( \sum_{i=M+1}^{|S|} P_S(s_i) \geq \text{OPT}(SC) \). Together with the lower bound obtained, we thus have \( \text{OPT}(SC) = \sup_{0 \leq \phi \leq \beta} \left\{ \|\phi_1 - M\|\phi_\infty \right\} \).

Note that in (8), we have viewed \( \phi \) as a vector in \( \mathbb{R}^{|S|} \) that is nonnegative and dominated by \( P_S \in \mathbb{R}^{|S|} \). The maximization in (8) is a tradeoff between increasing the \( \ell_1 \) norm of \( \phi \) on the one hand, and decreasing the \( \ell_\infty \) norm of \( \phi \) on the other. One plausible strategy for this tradeoff is to take \( \phi(s) = P_S(s) \) for those \( s \in S \) for which \( P_S(s) \) is not too large, and zero otherwise. Specifically, one may take \( \phi(s) = P_S(s) \|P_S(s) \leq \exp(-\gamma)\| \) for some \( \gamma \geq 0 \). Then the RHS of (8) is lower bounded by

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

where \( h(S) = -\log P_S(S) \). The above converse is derived in [7] and is same as the converse in [5, Th. 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 $f : S_1 \times S_2 \rightarrow X_1 \times X_2$ 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 $g : Y_1 \times Y_2 \rightarrow \hat{S}_1 \times \hat{S}_2$.

The objective, as for SW problem, is to losslessly recover $(\hat{S}_1, \hat{S}_2)$ 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 $S := (S_1, S_2), Y := (Y_1, Y_2), \bar{S} := (\hat{S}_1, \hat{S}_2)$ and $d(S, \bar{S}) = \mathbb{I}(S \neq \bar{S})$ with $\Delta = 0$.

Consequently, to obtain finite blocklength converses for the joint encoding problem of correlated sources, we resort to the following generalized version of DPJE for lossless source coding problem,

$$
\begin{align*}
\text{DPJE} \quad \max & \quad \sum \tilde{\gamma}^a(s_1, s_2) + \sum \tilde{\gamma}^b(y_1, y_2) \\
\text{s.t.} & \quad \tilde{\gamma}^a(s_1, s_2) - \sum_{y_1, y_2} \tilde{\lambda}_c(s_1, s_2, x_1, y_1, y_2) \leq 0 \\
& \quad \tilde{\gamma}^b(y_1, y_2) - \sum_{s_1, s_2} \tilde{\lambda}_b(s_1, s_2, y_1, y_2) \leq 0 \\
& \quad \tilde{\lambda}_b(s_1, s_2, \tilde{x}_1, \tilde{x}_2, y_1, y_2) + \tilde{\lambda}_c(s_1, s_2, x_1, y_1, y_2) \leq \Upsilon(\bar{z})
\end{align*}
$$

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

As in the case of DP, taking $\tilde{\gamma}^a$ and $\tilde{\gamma}^b$ such that (A1) and (A2) hold with equality, OPT(DPJE) can be written in terms of the channel flow $\tilde{\lambda}_c(s_1, s_2, x_1, x_2, y_1, y_2)$ and source flow $\tilde{\lambda}_b(s_1, s_2, \tilde{x}_1, \tilde{x}_2, y_1, y_2)$. The LP-based converse for lossless source coding problem in Corollary 5 then readily implies the following corollary.

**Corollary 6 (LP-Based Converse for Jointly Encoded Sources):** Consider problem SC with $S := (S_1, S_2), X := (X_1, X_2), Y := (Y_1, Y_2), \bar{S} := (\hat{S}_1, \hat{S}_2)$ and $d(S, \bar{S}) = \mathbb{I}(S \neq \bar{S})$ with $\Delta = 0$. Consequently for any code, we have from Corollary 5,

$$
\mathbb{E}[\mathbb{I}(S_1, S_2) \neq (\hat{S}_1, \hat{S}_2)] \leq \text{OPT}(\text{DPJE})
$$

where the supremum is over $\tilde{\phi} : S_1 \times S_2 \rightarrow [0, 1]$ such that $0 \leq \tilde{\phi}(s_1, s_2) \leq P_{S_1, S_2}(s_1, s_2)$ for all $s_1 \in S_1, s_2 \in S_2$.

**Proof:** To obtain the required converse, we consider the following choice for the flows in DPJE, which generalizes the ones adopted in (5),

$$
\begin{align*}
\tilde{\lambda}_c(x_1, x_2, s_1, s_2, y_1, y_2) & \equiv \mathbb{I}(y_1, y_2) = (x_1, x_2)\tilde{\phi}(s_1, s_2) \\
\tilde{\lambda}_b(s_1, s_2, \tilde{x}_1, \tilde{x}_2, y_1, y_2) & \equiv -\tilde{\phi}(s_1, s_2)\mathbb{I}(s_1, s_2) = (\hat{s}_1, \hat{s}_2).
\end{align*}
$$

The feasibility of these flows with respect to (A3) can be verified in the proof of Theorem 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 $f_1 : S_1 \rightarrow X_1$ 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 $g : S_2 \times Y_1 \rightarrow \hat{S}_1$ to get $\hat{S}_1$.

The finite blocklength source coding of $S_1$ given $S_2$ as the side information can be then posed as the following optimization problem,

$$
\begin{align*}
\text{SID}_{S_1 | S_2} & \quad \min \mathbb{E}[\mathbb{I}(S_1 \neq \hat{S}_1)] \\
\text{s.t.} & \quad X_1 = f_1(S_1), \quad \hat{S}_1 = g(S_2, Y_1).
\end{align*}
$$

Thus, $\text{SID}_{S_1 | S_2}$ seeks to obtain a code $(f_1, g)$ which minimizes $\mathbb{E}[\mathbb{I}(S_1 \neq \hat{S}_1)] = P[S_1 \neq \hat{S}_1]$, 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 $\text{SID}_{S_1 | S_2}$.

$$
\begin{align*}
\text{LPSI}_{S_1 | S_2} & \quad \min \sum_{s_1, s_2, \tilde{z}} \Psi(\tilde{z}) W(\tilde{z}) \\
\text{s.t.} & \quad \sum_{s_1} Q_{S_1 Y_1 | S_2} 1 : \tilde{\gamma}^a(s_1) \\
& \quad \sum_{s_1} Q_{S_1 Y_1 | S_2} 1 : \tilde{\gamma}^b(s_1, y_1) \\
& \quad \sum_{s_1} W(\tilde{z}) - \sum_{s_1} Q_{S_1 Y_1 | S_2} 1 : \tilde{\gamma}^c(s_1, y_1, y_2) \geq 0
\end{align*}
$$

where $s := (s_1, s_2), \tilde{z} := (x_1, s_1, s_2, \tilde{x}_1, \tilde{x}_2)$ and $\Psi(\tilde{z}) = P_{S_1, S_2}(s_1, s_2)\mathbb{I}[S_1 \neq \hat{S}_1]\mathbb{I}[Y_1 = X_1]$.

Employing the Lagrange multipliers $\tilde{\gamma}^a, \tilde{\gamma}^b, \tilde{\gamma}^c$, corresponding to the constraints of $\text{LPSI}_{S_1 | S_2}$, we obtain the following dual of $\text{LPSI}_{S_1 | S_2}$.

$$
\begin{align*}
\text{DPSI}_{S_1 | S_2} & \quad \max \sum_{s_1} \tilde{\gamma}^a(s_1) + \sum_{s_1} \tilde{\gamma}^b(s_1, y_1) \\
\text{s.t.} & \quad \tilde{\gamma}^a(s_1) - \sum_{s_1, y_1} \tilde{\gamma}^c(s_1, s_2, x_1, y_1, y_2) \leq 0 \\
& \quad \tilde{\gamma}^b(s_1, y_1) - \sum_{s_1, s_2, y_1} \tilde{\lambda}_b(s_1, s_2, \tilde{x}_1, \tilde{x}_2, y_1, y_2) \leq 0 \\
& \quad \tilde{\lambda}_b(s_1, s_2, \tilde{x}_1, \tilde{x}_2, y_1, y_2) \leq \Psi(\tilde{z}) \\
& \quad \tilde{\lambda}_c(s_1, s_2, \tilde{x}_1, \tilde{x}_2, y_1, y_2) \leq \Psi(\tilde{z}) \\
& \quad \tilde{\lambda}_c(s_1, s_2, \tilde{x}_1, \tilde{x}_2, y_1, y_2) \leq \Psi(\tilde{z}) \\
\end{align*}
$$

Fig. 3. Lossless source coding with side information available only at the decoder.
Choosing $\hat{\gamma}^a(s_1)$ and $\hat{\gamma}^b(s_2, y_1)$ such that (B1) and (B2) hold with equality, OPT(DPS1) can be written in terms of $\lambda^c(s_1, s_2, x_1, y_1)$ and $\lambda^b(s_1, s_2, \hat{s}_1, y_1)$. Notice that $\lambda^c(s_1, s_2, x_1, y_1)$ is a function of $x_1, s_1, s_2, y_1$. Thus, for each $s_2 \in S_2$, it is akin to a channel flow of the point-to-point source coding problem with $S_1$ as the source. Likewise, for each $s_2 \in S_2$, $\lambda^b(s_1, s_2, \hat{s}_1, y_1)$ 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 SID1.

**Theorem 7:** Consider the problem SID1. For any code, the following lower bound holds,

$$\mathbb{E}[\|S_1 \neq \hat{S}_1]\| \geq \text{OPT}(\text{SID1}) \geq \text{OPT}(\text{LPS1}) \geq \sup_{0 \leq \gamma \leq \gamma^1} \left\{ \sum_{s_1, s_2} \gamma^1(s_1, s_2) - M_1 \sum_{s_2} \max_{\hat{s}_1} \phi^1(\hat{s}_1, s_2) \right\},$$

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

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

$$\tilde{\gamma}^c(s_1, s_2, x_1, y_1) = \mathbb{I}[x_1 = y_1] \gamma^1(s_1, s_2),$$

$$\tilde{\gamma}^b(s_1, s_2, \hat{s}_1, y_1) = -\phi^1(s_1, s_2) \mathbb{I}[s_1 = \hat{s}_1].$$

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

Note that as in the case with (3), the converse in (11) is equivalent to the following bound,

$$\mathbb{E}[\|S_1 \neq \hat{S}_1]\| \geq \text{OPT}(\text{SID1}) \geq \text{OPT}(\text{LPS1}) \geq \sup_{\eta_1 \geq 0} \left\{ \sum_{s_1, s_2} \min \{P_{S_1, S_2}(s_1, s_2), \eta_1(s_1, s_2)\} - M_1 \sum_{s_2} \min_{\hat{s}_1} \{P_{S_1, S_2}(s_1, s_2), \eta_2(s_1, s_2)\} \right\},$$

where $\eta_1 : S_1 \times S_2 \to [0, +\infty)$. When particularized to $\eta_2(s_1, s_2) = P_{S_2}(s_2) \exp(-\beta)$ 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 $\beta$, the converse in (13) becomes,

$$\mathbb{E}[\|S_1 \neq \hat{S}_1]\| \geq \text{OPT}(\text{LPS1}) \geq \text{OPT}(\text{DPS1}) \geq \sup_{\beta \geq 0} \left\{ \text{OPT}(\text{SID1}) \geq \text{OPT}(\text{LPS1}) \geq \text{OPT}(\text{DPS1}) \geq \sup_{0 \leq \beta \leq \beta^1} \left\{ \text{OPT}(\text{SID1}) \geq \text{OPT}(\text{LPS1}) \geq \text{OPT}(\text{DPS1}) \right\} \right\},$$

where $\beta^1(\tilde{s}_1, \tilde{s}_2, y_2, \tilde{y}_2) = \text{OPT}(\text{SID1}) \geq \text{OPT}(\text{LPS1}) \geq \text{OPT}(\text{DPS1}) \geq \sup_{\beta \geq 0} \left\{ \text{OPT}(\text{SID1}) \geq \text{OPT}(\text{LPS1}) \geq \text{OPT}(\text{DPS1}) \right\}$.

Here $h_{A|B}(a|b) = -\log P_{A|B}(a|b)$ is the conditional entropy density. The inequality in (a) follows from the definition 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 (14) by zero and upper bounding $\min\{P_{S_1, S_2}(s_1, s_2), P_{S_2}(s_2) \exp(-\beta)\}$ with $P_{S_2}(s_2) \exp(-\beta)$. Notice that the converse in (15) is the well-known converse for lossless source-coding problem with side-information at the decoder. The converse in (14) provides new improvement on the standard converse in (15) due to the presence of the additional non-negative term corresponding to when $h_{S_1|S_2}(s_1, s_2) < \log M_1 + \beta$ and that $-M_1 \sum_{s_2} \sup_{\hat{s}_1} \min \{P_{S_1, S_2}(s_1, s_2), P_{S_2}(s_2) \exp(-\beta)\}$.

**Theorem 8:** Consider the problem SID2. For any code, the following lower bound holds,

$$\mathbb{E}[\|S_2 \neq \hat{S}_2]\| \geq \text{OPT}(\text{SID2}) \geq \text{OPT}(\text{LPS2}) \geq \sup_{\beta \geq 0} \left\{ \text{OPT}(\text{SID2}) \geq \text{OPT}(\text{LPS2}) \geq \text{OPT}(\text{DPS2}) \right\} \geq \sup_{0 \leq \beta \leq \beta^1} \left\{ \text{OPT}(\text{SID2}) \geq \text{OPT}(\text{LPS2}) \geq \text{OPT}(\text{DPS2}) \right\},$$

where $\beta^1(\tilde{s}_1, \tilde{s}_2, y_2, \tilde{y}_2) = \text{OPT}(\text{SID2}) \geq \text{OPT}(\text{LPS2}) \geq \text{OPT}(\text{DPS2}) \geq \sup_{\beta \geq 0} \left\{ \text{OPT}(\text{SID2}) \geq \text{OPT}(\text{LPS2}) \geq \text{OPT}(\text{DPS2}) \right\}$.

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 \( Q : S_1 \times S_2 \times X_1 \times X_2 \times Y_1 \times Y_2 = 0, 1 \) which can be factored as,

\[
Q(z) = P_{S_1,S_2}Q_{X_1|S_1}Q_{X_2|S_2}P_{Y_1|X_1X_2}Q_{S_1|S_2}Y_1Y_2(z),
\]

where \( Z := S_1 \times S_2 \times X_1 \times X_2 \times Y_1 \times Y_2 \times \hat{S}_1 \times \hat{S}_2 \) and \( z := (s_1, s_2, x_1, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2) \in Z \). Employing \( Q \), we obtain the following optimization problem over joint probability distributions,

\[
\begin{align*}
\text{min} & \quad \sum_{z} I((s_1, s_2) \neq (\hat{s}_1, \hat{s}_2))Q(z) \\
\text{s.t.} & \quad P_{S_1,S_2}Q_{X_1|S_1}Q_{X_2|S_2}P_{Y_1|X_1X_2}Q_{S_1|S_2}Y_1Y_2(z) \equiv Q(z),
\end{align*}
\]

where, \( P(A|B) := \frac{Q(A|B)}{\sum_a Q(A|B)} = 1 \), \( \forall a, b \). Here, \( Q_{X_1|S_1} \) and \( Q_{X_2|S_2} \) represent the two randomized encoders, and \( Q_{S_1, S_2}|Y_1 Y_2 \) 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 (two randomized encoders and the randomized decoder), 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 LP, 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,

\[
W(z) \equiv Q_{X_1|S_1}Q_{X_2|S_2}Q_{S_1, S_2|Y_1Y_2}(z),
\]

\[
U(z_1) \equiv Q_{X_1|S_1}(x_1|s_1)Q_{S_1, S_2|Y_1Y_2}(s_1, s_2, y_1, y_2),
\]

\[
V(z_2) \equiv Q_{X_2|S_2}(x_2|s_2)Q_{S_1, S_2|Y_1Y_2}(s_1, s_2, s_1, s_2),
\]

\[
T(z_3) \equiv Q_{X_1|S_1}(x_1|s_1)Q_{X_2|S_2}(x_2|s_2),
\]

where recall that \( z := (s_1, s_2, x_1, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2) \), \( z_1 := (s_1, x_1, y_1, y_2, \hat{s}_1, \hat{s}_2) \), \( z_2 := (s_2, x_2, \hat{s}_1, \hat{s}_2, y_1, y_2) \) and \( z_3 := (s_1, s_2, x_1, x_2) \). These new 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 \( s_1 \in S_1 \), we multiply both sides of the constraint \( \sum_{s_2} Q_{X_1|S_1}(x_1|s_1) = 1 \) by \( Q_{X_2|S_2}(x_2|s_2) \) for all \( x_2, s_2 \), by \( Q_{S_1, S_2|Y_1Y_2}(s_1, s_2, y_1, y_2) \) for all \( s_1, s_2, y_1, y_2 \) and by \( Q_{X_2|S_2}(x_2|s_2)Q_{S_1, S_2|Y_1Y_2}(s_1, s_2, y_1, y_2) \) for all \( x_2, s_2, \hat{s}_1, \hat{s}_2, y_1, y_2 \), to obtain three new sets of multilinear equality constraints. Replace the resulting multilinear product terms by the newly defined variables in (18)-(21). This results in the following new valid linear constraints in the lifted space,

\[
\begin{align*}
\sum_{s_1} T(x_1, x_2, s_1, s_2) & \equiv Q_{X_1|S_1}(x_1|s_1)Q_{X_2|S_2}(x_2|s_2) \\
\sum_{s_1} U(x_1, s_1, \hat{s}_1, s_2, y_1, y_2) & \equiv Q_{S_1, S_2|Y_1Y_2}(s_1, s_2, y_1, y_2) \\
\sum_{s_1} W(z) & \equiv V(x_2, s_2, \hat{s}_1, \hat{s}_2, y_1, y_2).
\end{align*}
\]

Similar set of linear constraints can be obtained corresponding to \( \sum_{s_2} Q_{X_1|S_1}(x_1|s_2) = 1 \) for all \( s_2 \) and \( \sum_{s_1} Q_{S_1, S_2|Y_1Y_2}(s_1, s_2, y_1, y_2) = 1 \) for all \( y_1, y_2 \). Subsequently, 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 (18)-(21) and we have the LP relaxation of SW, LPSW as as shown at the top of the next page. Notice that the constraints of LPSW are implied by the constraints of SW where LPSW is a relaxation of SW.

Here, \( \gamma^x(s_1), \gamma^y(s_2), \gamma^c(s_1, s_2) \), \( \mu_{(1)}(s_1, s_2, s_1, s_2) \), \( \mu_{(2)}(s_1, s_2, s_2, s_1) \), \( \mu_{(3)}(s_1, s_2, y_1, y_2) \), \( \mu_{(4)}(s_1, s_2, y_1, y_2) \), \( \mu_{(5)}(s_1, s_2, x_1, y_1, y_2, s_1, s_2) \), \( \mu_{(6)}(s_1, s_2, x_1, y_1, y_2, s_1, s_2) \) and \( \mu_{(7)}(s_1, s_2, x_1, y_1, y_2, s_1, s_2) \) 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 \( s_1 \in S_1, s_2 \in S_2, y_1 \in Y_1, y_2 \in Y_2 \), \( W(s_1, s_2, y_1, y_2) \) is a coupling on \( X_1 \times X_2 \times (\hat{S}_1 \times \hat{S}_2) \) between all marginal \( Q_{X_1|S_1}(\cdot|x_1), Q_{X_2|S_2}(\cdot|x_2) \) and \( Q_{S_1, S_2|Y_1Y_2}(\cdot, \cdot|y_1, y_2) \); denote the set of all such W by \( \Xi' \). However, note that \( W \in \Xi' \) does not automatically imply that, for instance,

\[
\sum_{s_1} W(z) \in \Xi(Q_{X_1|S_1}, Q_{S_1, S_2|Y_1Y_2}),
\]

and likewise, that

\[
\sum_{s_2} W(z) \in \Xi(Q_{X_1|S_1}, Q_{S_1, S_2|Y_1Y_2}),
\]

\[
\sum_{s_1} W(z) \in \Xi(Q_{X_1|S_1}, Q_{X_2|S_2}).
\]

LPSW is obtained by imposing not only that \( W \in \Xi' \), but also (22)-(24). 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 \( \Xi' \) that is constrained by (22)-(24), and then minimizing over all codes.

We note that the variables \( U, V, T \) in (19)-(21) are introduced in LPSW only to express the constraints (22)-(24) 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 \( (Q_{X_1|S_1}, Q_{X_2|S_2}, Q_{S_1, S_2|Y_1Y_2}) \).
JOSE AND KULKARNI: IMPROVED FINITE BLOCKLENGTH CONVERSES FOR SLEPIAN–WOLF CODING VIA LP

B. Duality and Converse

Employing the Lagrange multipliers corresponding to the constraints of LPSW, we now obtain the dual of LPSW, denoted DPSW, as shown at the top of this page. Here, \( \Pi \) is defined as:

\[
\Pi \triangleq \{ (s_1, s_2) \neq (\tilde{s}_1, \tilde{s}_2) \} \times P_{S_1, S_2}(s_1, s_2) \mathbb{I}( (y_1, y_2) = (x_1, x_2) )
\]

Let \( \Theta \) be defined as:

\[
\Theta \triangleq \{ \lambda_{s_1}^{(1)}, \lambda_{s_2}^{(1)}, \lambda_c^{(1)}, \mu_{s_1}^{(1)}, \mu_{s_2}^{(1)}, \mu_{c_1}^{(1)}, \mu_{c_2}^{(1)}, \mu_{c_3}^{(1)}, \mu_{c_4}^{(1)} \}
\]

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 1:** Any collection of functions \( \Theta \) satisfying constraints (D4)-(D7) yields the following lower bound on the optimal value of SW, i.e.,

\[
\text{OPT(SW)} \geq \text{OPT(LPSW)} \equiv \text{OPT(DPSW)}
\]

The inequality in (a) follows since LPSW is a relaxation of SW and (b) results from the duality of linear programming. The last inequality follows from \( \Theta \) 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,
\[
\begin{align*}
\lambda^{(12)}_S &: S_1 \times S_2 \times X_2 \times Y_1 \times Y_2 \times \bar{S}_1 \times \bar{S}_2 \rightarrow \mathbb{R}, \\
\lambda^{(21)}_S &: S_1 \times S_2 \times X_1 \times Y_1 \times Y_2 \times \bar{S}_1 \times \bar{S}_2 \rightarrow \mathbb{R}, \\
\lambda_c &: S_1 \times S_2 \times X_2 \times Y_1 \times Y_2 \rightarrow \mathbb{R}, \\
\mu^{(1)}_S &: S_1 \times \bar{S}_1 \times \bar{S}_2 \times Y_1 \times Y_2 \rightarrow \mathbb{R}, \\
\mu^{(2)}_S &: S_2 \times \bar{S}_1 \times \bar{S}_2 \times Y_1 \times Y_2 \rightarrow \mathbb{R}, \\
\mu^{(2)}_c &: S_2 \times X_1 \times \bar{S}_1 \times \bar{S}_2 \rightarrow \mathbb{R}, \\
\mu^{(1)}_c &: S_2 \times X_1 \times \bar{S}_1 \times \bar{S}_2 \rightarrow \mathbb{R}, \\
\mu^{(2)}_c &: X_1 \times S_1 \times S_2 \rightarrow \mathbb{R}, \\
\end{align*}
\]
(26)
such that the point-wise inequalities in (D4)-(D7) are satisfied.

We call the above collection of functions \( \Theta \), 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, DPSI\(_{12} \) and DPSI\(_{21} \), 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 SW coding in a formal relationship between the feasible regions of problems DPJE and DPSI\(_{12} \) and DPSI\(_{21} \) 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 \( \lambda^{(12)}_S \), \( \lambda^{(21)}_S \), \( \lambda_c \) and \( \mu^{(1)}_S \), \( \mu^{(2)}_S \), \( \mu^{(1)}_c \), \( \mu^{(2)}_c \). We will refer to these variables (recall that these are functions, as stated in (26)) also as flows. Our approach for interpreting and classifying these flows is based on relating these flows to flows of problems DPJE, DPSI\(_{12} \) and DPSI\(_{21} \). We remark that there may be other approaches that would yield a more refined understanding.

We begin with the \( \lambda \)'s. Consider the flow \( \lambda^{(12)}_S \). Observe that \( \lambda^{(12)}_S \) is a function of \( s_1, s_2, x_2, y_1, y_2, \bar{s}_1, \bar{s}_2 \) but is independent of \( x_1 \). Hence, for each fixed \( s_2, x_2, y_2 \) and \( \bar{s}_2 \), \( \lambda^{(12)}_S \) may be likened to a source flow from \( \bar{s}_1 \) to \( \bar{s}_1 \) (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 \( \lambda^{(12)}_S \) on \( s_2, x_2, y_2, \bar{s}_2 \) hints at (coded or uncoded) side-information about \( S_2 \) through the path \( S_2 \rightarrow X_2 \rightarrow Y_2 \rightarrow \bar{S}_2 \). This leads one to surmise that the flow \( \lambda^{(12)}_S \) would have a close relation to the source flow of the problem DPSI\(_{12} \). Thus we refer to \( \lambda^{(12)}_S \) as a source flow for DPSW. Note though, that this is not the only heuristic one can apply. If \( \lambda^{(12)}_S \) is assumed to be also independent of \( x_2 \), then \( \lambda^{(12)}_S (s_1, s_2, x_1, y_1, y_2, \bar{s}_1, \bar{s}_2) \) can be also interpreted to be the source flow in the problem DPJE where \( \bar{s}_1, \bar{s}_2 \) are jointly encoded. Thus one surmises that a value for \( \lambda^{(12)}_S \) could probably be arrived at by a combination of the source flows of DPJE and DPSI\(_{12} \). A similar heuristic can be applied to surmise that \( \lambda^{(21)}_S \) could be arrived at through the source flows of DPJE and DPSI\(_{21} \). The flow \( \lambda^{(21)}_S (s_1, s_2, x_1, x_2, y_1, y_2) \) which depends on correlated sources and the channel inputs and outputs, appears to be related to the channel flows of all three problems DPJE, DPSI\(_{12} \) and DPSI\(_{21} \), 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 \( \mu \)'s. Notice that the flows of problem DPSW fall into a hierarchy wherein the \( \lambda \)'s are constrained by the error density bottleneck (constraint (D4)), whereas the \( \mu \)'s are constrained by a bottleneck determined by the \( \lambda \)'s. Arguing as in the case of the \( \lambda \)'s we see that \( \mu^{(1)}_S \) and \( \mu^{(2)}_C \) are akin to source and channel flows of a coding problem along the path \( S_1 \rightarrow X_1 \rightarrow Y_1 \rightarrow \bar{S}_1 \). Note though the objective of the problem (source coding, or something else) would depend on \( \lambda^{(12)}_S \), the RHS of constraint (D6). Likewise \( \mu^{(2)}_S \) and \( \mu^{(2)}_C \) resemble source channel flows for a coding problem along \( S_2 \rightarrow X_2 \rightarrow Y_2 \rightarrow \bar{S}_2 \) whose objective is determined by \( \lambda^{(21)}_S \). The final set of dual variables \( \mu^{(1)}_S \) and \( \mu^{(2)}_C \) 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, DPSI\(_{12} \) and DPSI\(_{21} \).

Proposition 9: Let \( \hat{\Theta}_{12} := \{ \hat{\gamma}^a, \hat{\gamma}^b, \hat{\bar{\lambda}}^{(12)}, \hat{\lambda}^{(12)} \} \in \text{FEA}(\text{DPSI}_{12}) \) with its corresponding objective value, \( \text{OBJ}(\text{DPSI}_{12}) = \sum_{s_1} \gamma^a(s_1) + \sum_{x_1, y_1} \gamma^b(x_1, y_1) \). Then the following choice of values for the variables of DPSW is feasible.

\[
\begin{align*}
\lambda^{(12)}_S (s_1, s_2, x_2, y_1, y_2, \bar{s}_1, \bar{s}_2) &\equiv \hat{\lambda}^{(12)}_S (s_1, s_2, \bar{s}_1, y_1) \mathbb{I}[x_2 = y_2], \\
\lambda^{(21)}_S (s_1, s_2, x_1, x_2, y_1, y_2) &\equiv \hat{\lambda}^{(21)}_S (s_1, x_2, \bar{s}_1, y_1) \mathbb{I}[x_2 = y_2], \\
\mu^{(2)}_S (x_2, y_1, y_2) &\equiv \hat{\mu}^b(x_2, y_1) \mathbb{I}[x_2 = y_2], \\
\mu^{(2)}_c (x_1, x_2, y_1, y_2) &\equiv \hat{\gamma}^b(x_1, y_2) \mathbb{I}[x_2 = y_2], \\
\mu^{(1)}_S (s_1, \bar{s}_1, \bar{s}_2, y_1, y_2) &\equiv 0, \\
\mu^{(1)}_c (s_1, \bar{s}_1, \bar{s}_2, y_1, y_2) &\equiv 0, \\
\gamma^a (s_1) &\equiv \sum_{y_1} \gamma^a (s_1, y_1). \\
\end{align*}
\]
(27)
Consequently, \( \text{OPT}(\text{DPSW}) \geq \sum_{s_1} \gamma^a(s_1) + \sum_{x_1,y_1} \gamma^b(x_1,y_1) + \sum_{y_1,y_2} \gamma^a(y_1,y_2) \) which is equal to the objective of DPSI\(_{12} \) under \( \hat{\Theta}_{12} \). In particular, considering \( \Theta_{12} \) as the optimal solution of DPSI\(_{12} \) gives \( \text{OPT}(\text{DPSW}) \geq \text{OPT}(\text{DPSI}_{12}) \).
**Proof:** The proof is included in Appendix B. □

It thus becomes clear that given any feasible point \((\tilde{\gamma}^a, \tilde{\gamma}^b, \tilde{\lambda}_b, \tilde{\lambda}_c)\) in FEA(DPSI12), one can construct a feasible point of DPSW as given in (27). Moreover, the objective value of the resulting feasible point gives a lower bound on OPT(DPSW). Similarly, it can be shown that given any feasible point of DPSI21, one can construct a feasible point of DPSW with the same cost, thereby implying

\[
\text{OPT(DPSW)} \geq \text{OPT(DPSI21)}.
\]

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

**Proposition 10:** The following relationship between the feasible region of DPJE and DPSW holds. Let \(\tilde{\Theta} := (\tilde{\gamma}^a, \tilde{\gamma}^b, \tilde{\lambda}_b, \tilde{\lambda}_c)\) in FEA(DPJE) with its corresponding objective value, \(\text{OBJ(DPJE)} = \sum_{s_1, s_2} \tilde{\gamma}^a(s_1, s_2) + \sum_{y_1, y_2} \tilde{\gamma}^b(y_1, y_2)\). Then the following choice of values of the variables of DPSW are feasible.

\[
\begin{align*}
\lambda_b^{(12)} & (s_1, s_2, x_1, y_1, y_2, \tilde{\gamma}, \tilde{\lambda}_c) \equiv \tilde{\lambda}_b(s_1, s_2, \tilde{\gamma}, \tilde{\lambda}_c), \\
\lambda_c^{(12)} & (s_1, s_2, x_1, y_1, y_2, \tilde{\gamma}, \tilde{\lambda}_c) \equiv \tilde{\lambda}_c(s_1, s_2, \tilde{\gamma}, \tilde{\lambda}_c), \\
\mu_c^{(2)} & (s_2, x_1, y_1, y_2) \equiv 0, \\
\mu_{c}^{(12)} & (s_1, s_2) \equiv 0, \\
\gamma^{c} (s_1, y_1, y_2) & \equiv \tilde{\gamma}^{b}(s_1, y_1, y_2), \\
\gamma^{c} (s_2, \tilde{\gamma}, \tilde{\lambda}_c) & \equiv \sum_{s_1} \tilde{\lambda}_c(s_1, s_2, \tilde{\gamma}, \tilde{\lambda}_c).
\end{align*}
\]

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

**Proof:** The proof is similar to the proof of Proposition 9 and we skip the proof here. □

The relationships between the feasible regions of DPSW with that of DPJE, DPSI12 and DPSI21 established through Propositions 9 and 10 help in establishing a formal interpretation for the roles of the dual variables of DPSW. From Proposition 9, we see that the dual variable \(\lambda_{b}^{(12)}\) that we considered as akin to a source flow for DPSI12 and also as a source flow for the DPJE, has a somewhat more complex interpretation. Specifically, while the latter interpretation holds thanks to Proposition 10, the former holds only along the diagonal \(x_2 = y_2\), as seen in (27). A similar caveat holds for the channel flow \(\lambda_{c}\) and the other source flow \(\lambda_{c}^{(12)}\).

We remark here that the choices of the \(\mu^{c}\)’s in Propositions 10 and 9 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 \(\sum_{s_1} \lambda_{b}^{(12)}\) is considered as in constraint (D5) (recall here that \(\lambda_{b}^{(12)}\) resembles a source flow from \(S_1\) to \(\tilde{S}_1\) given the side-information of \(S_2\) available through the path \(S_2 \rightarrow X_2 \rightarrow Y_2 \rightarrow \tilde{S}_2\), the source flow from \(S_1\) to destination \(\tilde{S}_1\) is averaged out, and what is left, is the influence of side-information of \(S_2\). Thus, the RHS of constraint (D5) accounts for the side-information of \(S_2\) while the LHS is determined by \(\mu_{c}^{(12)}\) and \(\mu_{c}^{(2)}\). From Proposition 10 it can then be seen that \(\mu_{c}^{(2)}\) being a function of \(s_2, y_1, y_2, \tilde{\gamma}, \tilde{\lambda}_c\) and not of \(s_2\), accounts for the point-to-point like source flow through the side-information path from \(S_2\) to the destination node \((\tilde{S}_1, \tilde{S}_2)\). Further, (27) implies that \(\mu_{c}^{(12)}\) accounts for the channel flow through this path.

Similarly, \(\sum_{s_2} \lambda_{b}^{(12)}\) in the RHS of constraint (D6) averages out the source flow from \(S_2\) to \(\tilde{S}_2\), leaving behind the influence of side-information of \(S_1\). Consequently, as before, \(\mu_{b}^{(12)}\) represents the point-to-point like source flow through the side-information path from \(S_1\) to \((\tilde{S}_1, \tilde{S}_2)\) and \(\mu_{c}^{(1)}\) represents the corresponding channel flow for this path. Finally, thanks to the relation \(\mu_{c}^{(12)} = \sum_{s_1} \gamma_{c}^{(12)}\) in (27) (recall that \(\lambda_{c}^{(12)}\) represents the point-to-point channel flow from source \(S_1\) given information about \(S_2\)), we can interpret that \(\mu_{c}^{(12)}\) represents an average channel flow from \(S_1\) to \((\tilde{S}_1, \tilde{S}_2)\) given the side-information about \(S_2\). Similarly, \(\mu_{c}^{(1)}\) represents an average channel flow from \(S_2\) given information about \(S_1\).

A. Synthesizing Converges for SW From Point-to-Point Duals

As an immediate consequence of Proposition 9 and Proposition 10, we get that the point-to-point LP-based converges in (11), (17) and (9) are all lower bounds on OPT(DPSW). Consequently, the following is a straightforward lower bound on OPT(DPSW),

\[
\text{OPT(DPSW)} \geq \max \left\{ \sup_{0 \leq \phi(s_1, s_2) \in P_{s_1, s_2}} \left| \sum_{s_1, s_2} \tilde{\phi}(s_1, s_2) \right| - M_1 M_2 \max_{s_1, s_2} \tilde{\phi}(s_1, s_2), \right. \\
\left. \left| \sum_{s_1, s_2} \phi^{(12)}(s_1, s_2) \right| - M_1 \max_{s_1} \phi^{(12)}(s_1, s_2), \right. \\
\left. \left| \sum_{s_1, s_2} \phi^{(2)}(s_1, s_2) \right| - M_2 \max_{s_1} \phi^{(2)}(s_1, s_2) \right\} \right). \tag{28}
\]

Convex analytically speaking, the above bound considers a convex combination of feasible points of DPSW obtained via Propositions 9 and 10. 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, DPSI12 and DPSI21. We will subsequently apply specific LP-based converges from Corollary 6, Theorem 7 and Theorem 8 to get our new LP-based converges.

**Theorem 11:** Let \((\tilde{\gamma}^a, \tilde{\gamma}^b, \tilde{\lambda}_b, \tilde{\lambda}_c) \in \text{FEA(DPSI12)}, (\gamma^a, \gamma^b, \gamma^{(2)}(1), \gamma^{(2)}(1)) \in \text{FEA(DPSI21)}\) and \((\gamma^a, \gamma^b, \tilde{\lambda}_b, \tilde{\lambda}_c) \in \text{FEA(DPJE)}\). Then, any choice of values for the variables
of DPSW satisfying the following equations is feasible for DPSW.

\[
\lambda^{(12)}_b(s_1, s_2, x_2, y_1, y_2, \tilde{s}_1, \tilde{s}_2) = \left[ \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, x_1, y_1) \right] \times \\
\|x_2 = y_2\| + a \lambda_b(s_1, s_2, \tilde{s}_1, \tilde{s}_2, y_1, y_2) \\
\|\{s_1, s_2\} = (\tilde{s}_1, \tilde{s}_2)\},
\]

\[
\lambda^{(21)}_b(s_1, s_2, x_1, y_1, y_2, \tilde{s}_1, \tilde{s}_2) = \left[ \lambda^{(21)}_b(s_1, s_2, \tilde{s}_2, y_2) \right] \\
\|x_1 = y_1\| + (1 - a) \lambda_b(s_1, s_2, \tilde{s}_1, \tilde{s}_2, y_1, y_2) \times \\
\|\{s_1, s_2\} = (\tilde{s}_1, \tilde{s}_2)\},
\]

\[
\lambda_c(s_1, s_2, x_1, y_1, y_2) = \min \left\{ P(s_1, s_2) \| (y_1, y_2) = (x_1, x_2) \right\},
\]

\[
\tilde{\gamma}_c(s_1, s_2, x_1, y_1, y_2) = \left[ \tilde{\gamma}^{(12)}_c(s_1, s_2, \tilde{s}_1, x_1, y_1) \right] \times \\
\|x_1 = y_1\| + \left( 1 - a \right) \tilde{\gamma}_b(s_1, s_2, \tilde{s}_1, \tilde{s}_2, y_1, y_2) \times \\
\|\{s_1, s_2\} = (\tilde{s}_1, \tilde{s}_2)\},
\]

\[
\mu_c^{(2)}(s_2, x_2, y_2) \leq \left[ \gamma^b(y_1, s_2) - \sum_{s_1 \neq \tilde{s}_1} \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_2) \right] \times \\
\|x_2 = y_2\| \tilde{s}_2, 
\]

\[
\mu_c^{(2)}(s_2, \tilde{s}_1, \tilde{s}_2, y_1, y_2) = \alpha \lambda_b(s_1, s_2, \tilde{s}_1, \tilde{s}_2, y_1, y_2) \times \\
\|\{s_1, s_2\} = (\tilde{s}_1, \tilde{s}_2)\},
\]

\[
\mu_c^{(1)}(s_1, x_1, y_1, y_2) \leq \left[ \gamma^b(s_1, y_2) - \sum_{s_2 \neq \tilde{s}_2} \gamma^{(21)}_b(s_2, s_2, \tilde{s}_2, y_2) \right] \times \\
\|\{x_1 = y_1\} \tilde{s}_1, 
\]

\[
\mu_c^{(1)}(s_1, \tilde{s}_1, \tilde{s}_2, y_1, y_2) = (1 - a) \tilde{\gamma}_b(s_1, \tilde{s}_2, \tilde{s}_1, \tilde{s}_2, y_1, y_2) \times \\
\|\{s_1 = \tilde{s}_1\},
\]

\[
\mu_c^{(21)}(x_2, s_1, s_2) + \mu_c^{(12)}(x_1, s_1, s_2) \leq \sum_{y_1, y_2} \lambda_c(s_1, s_2, x_1, x_2, y_1, y_2),
\]

where \( a \in (0, 1) \) 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 11 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 \( \lambda^{(12)}_b \) is taken as a superset of the source flow for DPSI\(_{12} \) (which acts only when \( x_2 = y_2 \)) and a fraction of the source flow for DPJE. Similarly, \( \lambda^{(21)}_b \) is a superset of the source flow for DPSI\(_{21} \) (which acts only when \( x_1 = y_1 \)) with the remaining fraction of the source flow for DPJE. While \( \lambda^{(12)}_b \) and \( \lambda^{(21)}_b \) are linear combinations of the point-to-point source flows, the channel flow \( \lambda_c \) is quite different. Instead of a linear combination of the point-to-point channel flows, \( \lambda_c \) considers a minimum of the combination of the channel flows of DPSI\(_{12} \), DPSI\(_{21} \) and DPJE, and \( P_{S_1, S_2}(s_1, s_2) \| (y_1, y_2) = (x_1, x_2) \), such that the bottleneck in constraint (D4) is satisfied. This in turn, makes \( \lambda_c \) a non-linear combination of these flows. Since \( \lambda_c \) dictates the choice of \( \mu^{(12)}_c \) and \( \mu^{(21)}_c \), the nonlinearity is also inherited in the relation of \( \mu \)'s.

Moreover, the nonlinearity of \( \lambda_c \) resulting from taking the minimum between \( P_{S_1, S_2}(s_1, s_2) \| (y_1, y_2) = (x_1, x_2) \) and the point-to-point flows is one of the reasons for the improvement on the classical converse of Miyake and Kanaya, as can be seen later in (32). Precisely, as can be seen later from (53), the choice of \( \lambda_c \) is of the form \( \min\{a, b\} \) which evaluates to \( a1[a \leq b] + b1[a > b] \), and thus results in the non-negative term in the LP-based converse of (30). In contrast, the classical converse of Miyake-Kanaya accounts for only \( a1[a \leq b] \) term as can be seen later in (46). Thus, by also accounting for the term \( b1[a > b] \), the choice of \( \lambda_c \) presents one avenue for improvement over the classical Miyake-Kanaya converse. 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 11, 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 DPSI\(_{12} \), DPSI\(_{21} \) and DPJE which yield the LP-based converses in (11), (17) and (9) for the corresponding point-to-point sub-problems and subsequently employing Theorem 11, we obtain the following new finite blocklength converse for SW.

**Theorem 12 (LP-Based Converse for Slepian-Wolf Coding):** Consider the problem SW. Consequently, for any code, the following bound holds:

\[
\mathbb{E}[\|S_1, S_2\| \neq (\tilde{S}_1, \tilde{S}_2)] \geq \text{OPT}(SW) \geq \text{OPT}(DPSW)
\]

\[
\geq \sup_{\phi, \phi^{(12)}_b, \phi^{(21)}_b} \sum_{s_1, s_2} \min\{P_{S_1, S_2}(s_1, s_2), \phi^{(12)}_b(s_1, s_2), \phi^{(21)}_b(s_1, s_2)\} - M_1 M_2 \max_{s_1, s_2} \phi^{(12)}_b(s_1, s_2)
\]

\[
- M_2 \sum_{s_1} \max_{s_2} \phi^{(21)}_b(s_1, s_2) - M_1 \sum_{s_2} \max_{s_1} \phi^{(12)}_b(s_1, s_2)
\]

where the supremum is over \( \tilde{\phi}, \phi^{(12)}_b, \phi^{(21)}_b : S_1 \times S_2 \rightarrow [0, 1] \) such that \( 0 \leq \tilde{\phi}(s_1, s_2), \phi^{(12)}_b(s_1, s_2), \phi^{(21)}_b(s_1, s_2) \leq P_{S_1, S_2}(s_1, s_2) \) for all \( s_1 \in S_1, s_2 \in S_2 \).

**Proof:** The proof is included in Appendix B.

Moreover, the converse in (30) is in fact equivalent to the following converse:

\[
\mathbb{E}[\|S_1, S_2\| \neq (\tilde{S}_1, \tilde{S}_2)] \geq \text{OPT}(SW) \geq \text{OPT}(DPSW)
\]

\[
\geq \sup_{\eta_1, \eta_2, \eta_3 \geq 0} \left\{ \sum_{s_1, s_2} \min\{P_{S_1, S_2}(s_1, s_2), \eta_1(s_1, s_2) + \eta_2(s_1, s_2) + \eta_3(s_1, s_2)\} \right\}
\]

\[
- M_1 \sum_{s_1} \max_{s_2} \min\{P_{S_1, S_2}(s_1, s_2), \eta_3(s_1, s_2)\}
\]

\[
- M_2 \sum_{s_2} \max_{s_1} \min\{P_{S_1, S_2}(s_1, s_2), \eta_2(s_1, s_2)\}
\]

(31)
where $\eta_1, \eta_2, \eta_3 : S_1 \times S_2 \to [0, \infty)$. To see the equivalence, note that (31) follows from (30) by considering
\[ \tilde{\phi}(s_1, s_2) = \min\{P_{S_1|S_2}(s_1, s_2), \eta_1(s_1, s_2)\} \]
\[ \phi^{(1)}(s_1, s_2) = \min\{P_{S_1|S_2}(s_1, s_2), \eta_2(s_1, s_2)\} \]
\[ \phi^{(2)}(s_1, s_2) = \min\{P_{S_1|S_2}(s_1, s_2), \eta_3(s_1, s_2)\}, \]
and subsequently taking supremum over $\eta_1, \eta_2, \eta_3$. Further, lower bounding (31) by choosing $\eta_1(s_1, s_2) = \tilde{\phi}^*(s_1, s_2)$, $\eta_2(s_1, s_2) = \phi^{(1)}(s_1, s_2)$, $\eta_3(s_1, s_2) = \phi^{(2)}(s_1, s_2)$, where $0 \leq \phi^*, \phi^{(1)}(s_1, s_2) \leq P_{S_1|S_2}(s_1, s_2)$ maximizes (30), yields the converse in (30).

Further, the new converse in (31) improves on the information spectrum based converse of Miyake and Kanaya [2] as shown in the following corollary.

**Corollary 13 (Improvements on Miyake-Kanaya Converse):** Consider problem SW. Then, for any code, the LP-based converse in (31) recovers the following improvement on the converse of Miyake and Kanaya,
\[ \mathbb{E}\mathbb{I}\{[(S_1, S_2) \neq (\hat{S}_1, \hat{S}_2)] \} \geq \text{OPT(SW)} \geq \text{OPT(DPSW)} \]
\[ \geq \sup_{\beta > 0} \left\{ \mathbb{P} \left[ h_{S_1, S_2}(S_1, S_2) \geq \log M_1 M_2 + \beta \text{ or } h_{S_1|S_2}(S_1|S_2) \right] \right\} \]
\[ + \sum_{s_1, s_2} \max \left\{ \frac{\exp(-\beta)}{M_1 M_2}, \frac{\exp(-\beta)}{M_1} P_{S_2|S_1}(s_2), \frac{\exp(-\beta)}{M_2} \right\} \times \mathbb{I}\left\{ P_{S_1|S_2}(s_1|s_2) > \frac{\exp(-\beta)}{M_1}, P_{S_2|S_1}(s_2|s_1) > \frac{\exp(-\beta)}{M_2} \right\} \]
\[ \times \mathbb{I}\left\{ P_{S_1|S_2}(s_1, s_2) > \frac{\exp(-\beta)}{M_1}, \frac{\exp(-\beta)}{M_2} \right\} - 3 \exp(-\beta) \right\}, \]  \( (32) \)
where $h_{A|B}(a|b) = -\log P_{A|B}(a|b)$ is the conditional entropy density and $h_{A,B}(a, b) = -\log P_{A,B}(a, b)$ is the joint entropy density.

**Proof:** To obtain the above converse, weaken (31) by choosing $\eta_1(s_1, s_2) = \exp(-\beta) M_1$, $\eta_2(s_1, s_2) = P_{S_2}(s_2)$ 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 $\min\{P(s_1, s_2), \eta_2(s_1, s_2), \eta_3(s_1, s_2)\}$. Further, bound $\min\{P_{S_1|S_2}(s_1|s_2), \eta_1(s_1, s_2)\}$ by $\min\{P_{S_1|S_2}(s_1|s_2), \eta_2(s_1, s_2), \eta_3(s_1, s_2)\}$ 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 2 (Recovering the Converse of Miyake and Kanaya):** Lower bounding the non-negative term in (32) corresponding to $\mathbb{I}\left\{ P_{S_1|S_2}(s_1|s_2) > \frac{\exp(-\beta)}{M_1}, P_{S_2|S_1}(s_2|s_1) > \frac{\exp(-\beta)}{M_2} \right\}$ with zero, we recover the converse of Miyake and Kanaya given as,
\[ \mathbb{E}\mathbb{I}\{[(S_1, S_2) \neq (\hat{S}_1, \hat{S}_2)] \} \geq \sup_{\beta > 0} \left\{ \mathbb{P} \left[ h_{S_1, S_2}(S_1, S_2) \geq \log M_1 M_2 + \beta \right] \right\} \]
\[ \text{or } h_{S_1|S_2}(S_1|S_2) \geq \log M_1 + \beta \]
\[ \text{or } h_{S_2|S_1}(S_2|S_1) \geq \log M_2 + \beta \right\} - 3 \exp(-\beta) \right\}, \]  \( (33) \)

Before we conclude, we note that the relevance of (28), particularly in the analysis of second-order asymptotics is limited. As pointed out in [12, Sec. 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 (30) or the union bound in (32) are more relevant. In fact, with the flexibility of choosing $\eta_1, \eta_2, \eta_3$ which are functions of $(s_1, s_2)$, the converse in (30) may even yield refined third order terms in the asymptotic analysis.

**B. Illustrative Example**

In this section, we consider the following example of a discrete memoryless correlated source and illustrate the performance of our new LP-based converse in comparison to the converse of Miyake and Kanaya. Let $S_1 = S_2 = \{0, 1\}^n$, the source distribution is stationary and memoryless with
\[ P_{S_1, S_2}(s_1, s_2) = \prod_{i=1}^n P_{S_{1i}, S_{2i}}(s_{1i}, s_{2i}), \]
where
\[ P_{S_{1i}, S_{2i}}(s_{1i}, s_{2i}) = \begin{cases} (1 - 3q) & \text{if } s_{1i} = s_{2i} = 0 \\ q & \text{else} \end{cases} \]
\[ 0 \leq q \leq \frac{1}{4} \]
is the source parameter. Consequently, $P_{S_1, S_2}$ can be equivalently written as
\[ P_{S_1, S_2}(s_1, s_2) = (1 - 3q)^{s_{1i} s_{2i}} q^{n - s_{1i} - s_{2i}}, \]
where $s_{1i}, s_{2i} = \sum_{i=1}^n s_{1i} = s_{2i} = 0$. Moreover,
\[ P_{S_i}(s_i) = (2q)^{w(s_i)} (1 - 2q)^{n - w(s_i)}, \quad i = 1, 2, \]  \( (34) \)
where $w(s_i)$ represents the Hamming weight of $s_i$. The converse in (31) can be particularized to the above correlated source by choosing
\[ \eta_1(s_1, s_2) = P_{S_2}(s_2) \frac{2^\beta}{M_1}, \quad \eta_3(s_1, s_2) = P_{S_1}(s_1) \frac{2^\beta}{M_2} \]
\[ \eta_1(s_1, s_2) = \frac{2^\beta}{M_1}, \quad \eta_3(s_1, s_2) = \frac{2^\beta}{M_2} \]  \( (35) \)
which results in the following converse,
\[ \mathbb{E}\mathbb{I}\{[(S_1, S_2) \neq (\hat{S}_1, \hat{S}_2)] \} \geq \text{OPT(DPSW)} \]
\[ \geq \sup_{\beta > 0} \left\{ \sum_{u=0}^{n} \sum_{z=0}^{n-u} \sum_{v=0}^{n-z} C_u \left( 1 - 3q \right)^{n-u} \frac{1}{M_1 M_2} + \frac{(1 - 2q)^{n-u} (2q)^v}{M_1} \right\} - \frac{M_1 M_2}{M_2} \times \]
where the first positive summation term results from (31) since for any \( s_1 \in S_1 \) with \( w(s_1) = u \), there exists \( n-u C \) number of \( s_2 \in S_2 \) such that \( w(s_2) = v \) and \( z_{s_1,s_2} = z \).

In comparison, the Miyake-Kanaya converse (33) particularizes to

\[
\mathbb{E}[I(S_1, S_2) \neq (\tilde{S}_1, \tilde{S}_2)] \geq \text{OPT(DPSW)}
\]

\[
\geq \sup_{\beta > 0} \left\{ \sum_{u=0}^{n} C_u \sum_{n-u}^{n} C_z \sum_{n-u}^{n} C \right\} - M_1 M_2 \left\{ \left( 1 - 3q \right)^{q^{n-\tau}} \left( 1 - 3q \right)^{q^{n-\tau}} \leq 2^{-\beta} \max \left\{ \frac{1}{M_1 M_2} \right\} - 32^{-\beta} \right\}, \tag{37}
\]

We now study the maximum of the point-to-point converses (28) in relation to the LP based converse in (36). Towards this, we particularize (28) to the above example by choosing

\[
\hat{\phi}(s_1, s_2) = \min \left\{ (1 - 3q)^{q^{n-\tau}}, \eta_1(s_1, s_2) \right\},
\]

\[
\phi^{(1,2)}(s_1, s_2) = \min \left\{ (1 - 3q)^{q^{n-\tau}}, \eta_2(s_1, s_2) \right\},
\]

\[
\phi^{(2,1)}(s_1, s_2) = \min \left\{ (1 - 3q)^{q^{n-\tau}}, \eta_3(s_1, s_2) \right\},
\]

where \( \eta_1, \eta_2, \eta_3 \) are defined as in (35). This results in the following converse,

\[
\text{OPT(DPSW)} \geq \max \left\{ \sup_{\beta > 0} \left\{ \sum_{u=0}^{n} C_u \sum_{n-u}^{n} C_z \right\} - M_1 M_2 \sup_{\beta > 0} \left\{ \left( 1 - 3q \right)^{q^{n-\tau}} \leq 2^{-\beta} \max \left\{ \frac{1}{M_1 M_2} \right\} - 32^{-\beta} \right\} \right\}
\]

\[
\sup_{\beta > 0} \left\{ \sum_{u=0}^{n} C_u \sum_{n-u}^{n} C_z \right\} - M_1 M_2 \sup_{\beta > 0} \left\{ \left( 1 - 3q \right)^{q^{n-\tau}} \leq 2^{-\beta} \max \left\{ \frac{1}{M_1 M_2} \right\} - 32^{-\beta} \right\}, \tag{38}
\]

Figure 4 and Figure 5 compare the performance of converses in (36), (37) and the maximum converse in (38). Figure 4 shows that the LP-based converse outperforms both the Miyake-Kanaya converse and the converse in (38) for small blocklengths when \( R_1, R_2 \notin R_{SW} \), where \( R_{SW} = \{(R_1, R_2) | R_1 \geq H(S_1|S_2), R_2 \geq H(S_2|S_1), R_1 + R_2 \geq H(S_1, S_2) \} \). Figure 5 compares the converses when \( (R_1, R_2) \in R_{SW} \) and shows that our LP-based converse outperforms the other two.

V. DISCUSSION

The strongest finite blocklength converse on OPT(SW) derivable from employing the LP-based framework is OPT(DPSW), the exact evaluation of which is difficult. However, a hierarchy of lower bounds on it can be derived through a series of optimization problems. This also helps us conceptually situate the LP-based converse in (30) and the Miyake-Kanaya converse in the hierarchy, as discussed below.
Recall that OPT(DPSW) evaluates to the following optimization problem,

\[
\max_\theta \left\{ \sum_{s_1} \min_{y_1,y_2} \left[ \sum_{y_1,y_2} \mu_c^{(1)}(s_1,x_1,y_1,y_2) + \sum_{s_2} \mu_c^{(1)}(x_1,s_1,s_2) \right] + \sum_{s_2} \min_{s_2} \left[ \sum_{y_1,y_2} \mu_c^{(2)}(s_2,x_2,y_1,y_2) + \sum_{s_1} \mu_c^{(1)}(x_1,s_1,s_2) \right] + \sum_{s_1,s_2} \min_{y_1,y_2} \left[ \sum_{y_1,y_2} \mu_s^{(2)}(s_2,\tilde{s}_1,\tilde{s}_2,y_1,y_2) + \sum_{s_1} \mu_s^{(1)}(s_1,\tilde{s}_1,\tilde{s}_2,y_1,y_2) \right] \right\},
\]

s.t. (D4)-(D7) hold, \hspace{1cm} (39)

which can be equivalently posed as

\[
\text{OPT(DPSW)} = \max_\theta \left\{ \sum_{s_1} \min_{y_1,y_2} \left[ \sum_{y_1,y_2} \mu_c^{(1)}(s_1,x_1,y_1,y_2) + \sum_{s_2} \mu_c^{(1)}(x_1,s_1,s_2) \right] + \sum_{s_2} \min_{s_2} \left[ \sum_{y_1,y_2} \mu_c^{(2)}(s_2,x_2,y_1,y_2) + \sum_{s_1} \mu_c^{(1)}(x_1,s_1,s_2) \right] + \sum_{s_1,s_2} \min_{y_1,y_2} \left[ \sum_{y_1,y_2} \mu_s^{(2)}(s_2,\tilde{s}_1,\tilde{s}_2,y_1,y_2) + \sum_{s_1} \mu_s^{(1)}(s_1,\tilde{s}_1,\tilde{s}_2,y_1,y_2) \right] \right\},
\]

s.t. (D4)-(D7) hold. \hspace{1cm} (40)

To see the equivalence, note that (39) ≥ (40) which follows from the following set of relations,

\[
\sum_{s_1} \min_{y_1,y_2} \left[ \sum_{y_1,y_2} \mu_c^{(1)} + \sum_{s_2} \mu_c^{(1)} \right] \geq \sum_{s_1} \min_{y_1,y_2} \left[ \sum_{y_1,y_2} \mu_c^{(1)} + \sum_{s_2} \mu_c^{(1)} \right] + \sum_{s_2} \min_{s_2} \left[ \sum_{y_1,y_2} \mu_c^{(2)} + \sum_{s_1} \mu_c^{(1)} \right]
\]
\[
\sum_{s_2} \min_{s_2} \left[ \sum_{y_1,y_2} \mu_c^{(2)} + \sum_{s_1} \mu_c^{(1)} \right] \geq \sum_{s_2} \min_{s_2} \left[ \sum_{y_1,y_2} \mu_c^{(2)} + \sum_{s_1} \mu_c^{(1)} \right] + \sum_{s_1,s_2} \min_{y_1,y_2} \left[ \sum_{y_1,y_2} \mu_s^{(2)} + \sum_{s_1} \mu_s^{(1)} \right]
\]
\[
\sum_{y_1,y_2} \min_{s_2} \left[ \sum_{y_1,y_2} \mu_s^{(2)} + \sum_{s_1} \mu_s^{(1)} \right] \geq \sum_{y_1,y_2} \min_{s_2} \left[ \sum_{y_1,y_2} \mu_s^{(2)} + \sum_{s_1} \mu_s^{(1)} \right] + \sum_{s_1,s_2} \min_{s_1,s_2} \left[ \sum_{s_1,s_2} \mu_s^{(1)} \right]
\]

We now show that (40) ≥ (39). Towards this, lower bound (40) by considering the following choice of dual variables,

\[
\lambda_c(s_1,s_2,x_1,x_2,y_1,y_2) = \lambda_c^{(1)}(s_1,s_2,x_1,x_2,y_1,y_2) + \lambda_c^{(2)}(s_1,s_2,x_1,x_2,y_1,y_2)
\]
\[
\mu_c(s_1,x_1,y_1,y_2) = \mu_c^{(1)}(s_1,x_1,y_1,y_2) + \mu_c^{(2)}(s_1,x_1,y_1,y_2)
\]

and \(\hat{s}_1,\hat{s}_2\) is a feasible solution of (40). Consequently, the resulting objective value results in the optimal value of (39).
Note that (43) now has an outer optimization over $\lambda_{S}^{(1/2)}$, $\lambda_{S}^{(2/1)}$, $\lambda_{C}$ satisfying the error density bottleneck (D4) and three inner optimization problems over each of the pairs, $(\mu_{S}^{(1/2)}, \mu_{C}^{(1/2)})$, $(\mu_{S}^{(2/1)}, \mu_{C}^{(2/1)})$, and $(\mu_{S}^{(2/1)}, \mu_{C}^{(2/1)})$, with bottlenecks imposed by constraints (D5), (D5), and (D7) respectively.

We further lower bound (43) by restricting the choice of $\lambda_{S}^{(1/2)}$ and $\lambda_{S}^{(2/1)}$ such that $\sum_{s_1} \lambda_{S}^{(1/2)}$ is independent of $\hat{s}_1$ and $\sum_{s_2} \lambda_{S}^{(2/1)}$ is independent of $\hat{s}_2$. Under this assumption, constraints (D5) and (D6) imply that $\mu_{S}^{(1)}$ and $\mu_{S}^{(2)}$ are independent of $\hat{s}_1$ and $\hat{s}_2$, respectively. Hence, for each $y_1 \in \mathcal{Y}_1$, $\sum_{s_2} \min_{s_2} \sum_{s_2} \mu_{C}^{(2)}(s_2, x_2, y_1, y_2) + \frac{1}{\mathcal{Y}_2} \sum_{s_2} \mu_{C}^{(2)}(s_2, \hat{s}_1, \hat{s}_2, y_1, y_2)$ represents the objective corresponding to the packing of source flow $\mu_{S}^{(2)}$ and channel flow $\mu_{C}^{(2)}$ through the path $S_2 \rightarrow X_2 \rightarrow Y_2 \rightarrow \hat{S}_2$ satisfying the bottleneck, $\mu_{S}^{(2)}(s_2, \hat{s}_1, \hat{s}_2, y_1, y_2) + \mu_{C}^{(2)}(s_2, x_2, y_1, y_2) \leq \sum_{s_1} \lambda_{S}^{(1/2)}(s_1, s_2, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2)$ for all $s_2, x_2, y_2, \hat{s}_2$. Taking the maximum over $\mu_{S}^{(2)}$, $\mu_{C}^{(2)}$ in (43) inside the summation over $y_1$, we can express the optimal packing of these flows as $\text{OPT}(\text{DP}, \sum_{s_1} \lambda_{S}^{(1/2)}(s_1, s_2, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2))$, defined as in (1). Note that here, the RHS of the bottleneck is not necessarily an error density, but a function of $(s_2, x_2, y_1, y_2, \hat{s}_2)$ and $\text{OPT}(\text{DP}, \sum_{s_1} \lambda_{S}^{(1/2)}(s_1, s_2, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2))$ is not necessarily the relaxation of a source coding problem.

Similarly, for each $y_2 \in \mathcal{Y}_2$, $\sum_{s_2} \min_{s_2} \sum_{s_2} \mu_{C}^{(1)}(s_1, x_1, y_1, y_2) + \frac{1}{\mathcal{Y}_2} \sum_{s_2} \mu_{C}^{(1)}(s_1, \hat{s}_1, \hat{s}_2, y_1, y_2)$ represents the objective corresponding to the packing of source flow $\mu_{S}^{(1)}$ and the channel flow $\mu_{C}^{(1)}$ through the path $S_1 \rightarrow X_1 \rightarrow Y_1 \rightarrow \hat{S}_1$ satisfying the bottleneck imposed by (D6). The resultant optimal packing can be expressed as $\text{OPT}(\text{DP}, \sum_{s_2} \lambda_{S}^{(2/1)}(s_1, s_2, x_1, y_1, y_2, \hat{s}_1, \hat{s}_2))$. Employing these yields the following lower bound on (43),

$$
\begin{align*}
\max \left\{ \sum_{s_1} \text{OPT}(\text{DP}, \sum_{s_1} \lambda_{S}^{(1/2)}(s_1, s_2, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2)) \right\} \\
\text{s.t. (D4) holds} \\
\sum_{s_1} \lambda_{S}^{(1/2)} \leq \hat{s}_1, \\
\sum_{s_2} \lambda_{S}^{(2/1)} \leq \hat{s}_2, \\
\sum_{s_2} \lambda_{S}^{(2/1)} \lambda_{S}^{(1/2)} \\
\max \left\{ \sum_{s_1} \min_{s_2} \sum_{s_2} \mu_{C}^{(1)}(x_1, s_1, s_2) \right\} \\
\text{s.t. (D7) holds} \\
\sum_{s_2} \min_{s_1} \sum_{s_1} \mu_{C}^{(1)}(x_2, s_1, s_2) \\
\right\}.
\end{align*}
$$

(44)

Note that for a given choice of $\lambda_{S}^{(1/2)}$, $\lambda_{S}^{(2/1)}$, the bound in (44) comprises of optimal value of the dual of point-to-point problems, $\text{OPT}(\text{DP}, \sum_{s_1} \lambda_{S}^{(1/2)})$ and $\text{OPT}(\text{DP}, \sum_{s_2} \lambda_{S}^{(2/1)})$. However, the objective of these problems is not necessarily source coding since RHS of (D5) or (D6)) is not the source coding error density.

Thus, the bounds in (43) and (44) 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 12 with $\lambda_{S}^{(1/2)}(s_1, s_2, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2)$

$$
\begin{align*}
\lambda_{S}^{(2/1)}(s_1, s_2, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2) &= -\left[ \phi^{(1/2)}(s_1, s_2) \| \{ s_1 = \hat{s}_1 \} \right] \\
&\times \| \{ y_2 = y_2 \} + a \varphi(s_1, s_2) \| \{ s_1 = \hat{s}_1, s_2 = \hat{s}_2 \} \\
\lambda_{S}^{(2/1)}(s_1, s_2, x_1, y_1, y_2, \hat{s}_1, \hat{s}_2) &= -\left[ \phi^{(2/1)}(s_1, s_2) \| \{ s_2 = \hat{s}_2 \} \right] \\
&\times \| \{ x_1 = y_1 \} + (1 - a) \varphi(s_1, s_2) \| \{ s_1 = \hat{s}_1, s_2 = \hat{s}_2 \},
\end{align*}
$$

for $a \in (0, 1)$, it is easy to verify that our LP-based converse in (30) follows via the construction of a feasible solution to the optimization problem in (43). Note that here, $\sum_{s_1} \lambda_{S}^{(1/2)}$ in general depends on $\hat{s}_1$ and $\sum_{s_2} \lambda_{S}^{(2/1)}$ depends on $\hat{s}_2$, whereby this construction is not feasible for (44). On the other hand, we find that the derivation of the Miyake and Kanaya converse from (32) corresponds to the following choice of variables,

$$
\begin{align*}
\lambda_{S}^{(1/2)}(s_1, s_2, x_2, y_1, y_2, \hat{s}_1, \hat{s}_2) &= -\| \{ s_1 = \hat{s}_1, s_2 = \hat{s}_2 \} \right| \| \{ y_2 = y_2 \} \right| P_{S_1, S_2}(s_1, s_2) \right| e^{-A} \\
&\times \| \{ y_1 = x_1 \} \right| P_{S_1, S_2}(s_1, s_2) \right| e^{-A} + \| \{ y_1 = x_1 \} \right| P_{S_1, S_2}(s_1, s_2) \right| e^{-A}
\end{align*}
$$

(45)

for $\beta > 0$. In this case, $\sum_{s_1} \lambda_{S}^{(1/2)}$ is independent of $\hat{s}_1$ and $\sum_{s_2} \lambda_{S}^{(2/1)}$ is independent of $\hat{s}_2$. Hence, with the following choice for the remaining dual variables satisfying (D4)–(D7),

$$
\begin{align*}
\lambda_{C}(s_1, s_2, x_1, x_2, y_1, y_2) &= \| \{ y_1 = y_2 \} \right| \| \{ x_1, x_2 \} \right| \\
\lambda_{C}(s_1, x_1, y_1, y_2) &= -\| \{ y_1 = y_2 \} \right| \| \{ y_1 = x_1 \} \right| \\
\lambda_{C}(s_1, \hat{s}_1, \hat{s}_2, y_1, y_2) &= -\| \{ y_1 = y_2 \} \right| \| \{ y_1 = x_1 \} \right|
\end{align*}
$$

(46)

and $\mu_{S}^{(2)}, \mu_{C}^{(1/2)} \equiv 0$, it becomes clear that the resulting Miyake-Kanaya converse follows as a lower bound on the lower level optimization problem in (44). This also implies that the converse of Miyake and Kanaya can be thought to be obtained by the $\lambda$’s in DPSW as inducing source-coding like
function is defined as

\[ \inf \{ \sum \sigma_k \lambda_k^{(1)}(\hat{S}_i) + \sum \sigma_k \lambda_k^{(2)}(\hat{S}_i) + \tau \geq \sum \sigma_k \lambda_k^{(1)}(z) + \sum \sigma_k \lambda_k^{(2)}(z) \} \]

as in (45).

Converse (43)

\[ \sum \sigma_k \lambda_k^{(1)}(\hat{S}_i) + \sum \sigma_k \lambda_k^{(2)}(\hat{S}_i) = \sum \sigma_k \lambda_k^{(1)}(z) + \sum \sigma_k \lambda_k^{(2)}(z) \]

Miyake-Kanaya Converse (33)

\[ \lambda_k^{(1)}(\hat{S}_i) \geq \lambda_k^{(1)}(z) \]

\[ \lambda_k^{(2)}(\hat{S}_i) \geq \lambda_k^{(2)}(z) \]

Fig. 6. Hierarchy of lower bounds derived. An arrow from \( A \rightarrow B \) implies that \( A \geq B \), the heading above the arrow indicate how \( B \) is obtained from \( A \). \( \min(\sum s') \geq \sum(\min s') \) as in (42).

problems in the DP's in (44). On the other hand our LP-based converse in (30) follows from a more complicated bound.

In summary, our LP-based converse in (30) and the Miyake-Kanaya converse in (33) 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 (40).

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 LP-based converses for lossy source coding and lossless source coding with side information problems, and recover the 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 LP-based converses for these point-to-point problems, a new LP-based converse for Slepian-Wolf coding was derived.

APPENDIX A

HYPOTHESIS TESTING BASED CONVERSE FOR LOSSY SOURCE CODING

For a source \( S \) with distribution \( P_S \), distortion function \( d : S \times \hat{S} \rightarrow [0, +\infty] \) and distortion level \( \Delta \), the rate-distortion function is defined as

\[ R_S(\Delta) = \inf_{P_{\hat{S}|S}} \{ I(S; \hat{S}) \} \]

where the infimum is over \( P_{\hat{S}|S} \in \mathcal{P}(\hat{S}|S) \). Assume that the infimum in (47) is achieved by a unique \( P_{\hat{S}|S} \) and \( \Delta_{\min} \)

\[ \inf \{ \Delta : R_S(\Delta) < \infty \} \]. The hypothesis testing based converse of Kostina and Verdú [5, Th. 8] is then obtained as below.

**Converse 1 (KV-hypothesis testing):** Consider problem SC with \( X = Y = \{1, \ldots, M\} \). Any code \((f, g)\) such that \( E[I(d(S, \hat{S}) > \Delta)] \leq \epsilon \) and \( \Delta > \Delta_{\min} \) must satisfy,

\[ M \geq \inf_{\hat{S} \in \mathcal{S}} \beta_{1-\epsilon}(P_S, Q) \]

(48)

for all \( Q \in \mathcal{P}(S) \) where \( \beta_{1-\epsilon}(P_S, Q) \) is the minimum type-II error \( \sum \{ Q(s)T(s) \} \) over all tests \( T \) such that the type-I error, \( \sum \{ P(s)(1 - T(s)) \} \leq \epsilon \). Moreover, the converse in (48) is equivalent to the following lower bound on the probability of error (see [13, eq. (72)]),

\[ \epsilon \geq \sup_{Q_S \in \mathcal{P}(S)} \alpha_{M^{*}}(P_S, Q_S), \]

(49)

with \( M^{*} = M \max_{\hat{S}} \sup_{Q_S} [I(d(S, \hat{S}) \leq \Delta)] \).

Further, it can be easily verified that the following holds,

\[ \alpha_{M^{*}}(P_S, Q_S) = \beta_{1-M^{*}}(Q_S, P_S). \]

(50)

**Corollary 14:** For any \( Q_S \in \mathcal{P}(S) \), the following relationship holds,

\[ \alpha_{M^{*}}(P_S, Q_S) \leq \sup_{\lambda > 0} \left\{ \sum \min \{ P_S(s), \lambda Q_S(s) \} - \lambda M^{*} \right\}. \]

(51)

Crucial to establishing the relationship in (51) is the following variational lemma of Elkayam and Feder [14, Lemma 1].

**Lemma 2 (Variational Lemma):** For \( 0 \leq \delta \leq 1, \)

\[ \beta_{\delta}(P_S, Q_S) = \sup_{\lambda > 0} \left\{ \sum \min \{ Q_S(s), \lambda P_S(s) \} - \lambda (1 - \delta) \right\}. \]

(52)

We now prove Corollary 14.

**Proof:** Note that for any \( Q_S \in \mathcal{P}(S), \)

\[ \alpha_{M^{*}}(P_S, Q_S) \]

\[ \leq \sup_{\lambda > 0} \left\{ \sum \min \{ P_S(s), \lambda Q_S(s) \} - \lambda M^{*} \right\}, \]

where the equality in (a) follows from (50) and equality in (b) follows from (52).

We now prove Corollary 3.

**Proof of Corollary 3:** The proof of the relation in (7) follows from Corollary 14. To prove the equivalence between the bounds in (4) and (6), we first show that (4) implies (6). To see this, take \( z(s) = \beta Q_S(s) \) in (4) where \( \beta \geq 0 \) and lower bound \( -M \sup_{s} \sum \min \{ P_S(s), z(s) \} \) with \( -M \sup_{s} \sum \{ z(s) \} \leq \Delta = -\beta M^{*} \). Subsequently, take the supremum over \( \beta \geq 0 \) and \( Q_S \in \mathcal{P}(S) \) to get the hypothesis testing bound (6).
We now show that (6) ≥ (4). To see this, we have the following sequences of observations,

\[
(6) = \sup_{Q_x \in \mathcal{P}(\mathcal{S})} \sup_{\beta \geq 0} \left\{ \sum_s \min \{ P_S(s), \beta Q_S(s) \} - \beta M^a \right\}
\]

\[\equiv \sup_{z_s \geq 0} \left\{ \sum_s \min \{ P_S(s), z_s \} - \max_s \sum_s z_s \mathbb{I}[d(s, \tilde{s}) \leq \Delta] \right\}
\]

\[\geq \sup_{\nu_s \geq 0} \left\{ \sum_s \min \{ P_S(s), \nu_s \} \right\} - \max_s \sum_s \min \{ P_S(s), \nu_s \} \mathbb{I}[d(s, \tilde{s}) \leq \Delta] = (4).
\]

Here, (a) results from taking \(z_s = \beta Q_S(s)\) such that \(z_s \geq 0\). The inequality in (b) follows by choosing \(z_s = \min \{ P_S(s), \nu_s \}\), where \(\nu_s \geq 0\).

**Appendix B**

**Proofs of Theorems in Section IV**

**Proof of Proposition 9:** Let \((\bar{y}^a, \bar{y}^b, \bar{\lambda}^{(12)}_b, \bar{\lambda}^{(1)}_c)\) ∈ FEA(DPS1\(_2\)). We now show that the choice of dual variables in (27) 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,

\[
\sum_{y_1, y_2} \mu^{(12)}_c(s_1, x_1, y_1, y_2) + \sum_{x_2} \mu^{(12)}_c(x_1, s_1, s_2)
\]

\[= \sum_{y_1, y_2} \bar{\lambda}^{(12)}_b(x_1, s_1, s_2, y_1) \geq \bar{y}^a(s_1) = \mu^{(1)}_a(s_1),
\]

thereby satisfying (D1). The inequality in (c) follows from the constraint (B1) of DPS1\(_2\). For checking feasibility with respect to constraint (D2), we get that

\[
\sum_{y_1, y_2} \mu^{(2)}_c(x_2, s_1, s_2, y_1, y_2) + \sum_{x_1} \mu^{(2)}_c(x_1, s_1, s_2)
\]

\[= \sum_{y_1, y_2} \bar{\lambda}^b(s_2, y_1) \mathbb{I}[x_2 = y_2] = \sum_{y_1} \bar{\lambda}^b(s_2, y_1) = \mu^b(s_2),
\]

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

\[
\lambda^{(12)}_b(s_1, s_2, x_2, y_1, 1, y_2, s_2, y_2) + \lambda_c(s_1, s_2, x_1, x_2, y_1, y_2),
\]

\[= \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_1) \mathbb{I}[x_2 = y_2]
\]

\[+ \lambda^{(12)}_c(s_1, s_2, x_1, 1) \mathbb{I}[x_2 = y_2]
\]

\[\leq \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_1) \mathbb{I}[x_1 = y_2] + \lambda^{(12)}_c(s_1, s_2, x_1, 1) \mathbb{I}[x_2 = y_2]
\]

\[\leq \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_1) + \lambda^{(12)}_c(s_1, x_1, 1) \mathbb{I}[x_2 = y_2]
\]

\[= \lambda^{(12)}_b(s_1, x_2, y_1, 1, y_2, s_2, y_2),
\]

which is the RHS, thereby satisfying (D4). Here, the inequality in (a) follows from the constraint (B3) of DPS1\(_2\).

To verify feasibility with respect to (D5), we have

\[\mu^{(2)}_c(s_2, s_1, \tilde{s}_1, y_1, y_2) + \mu^{(2)}_c(x_2, s_1, y_1, y_2) = \]

\[\bar{y}^b(s_2, y_1) \mathbb{I}[x_2 = y_2] \leq \sum_{s_1} \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_1, y_2, \tilde{s}_2),
\]

which is the RHS of (D5), thereby satisfying it. The inequality in (b) follows from constraint (B2) of DPS1\(_2\). Since \(\lambda^{(2)}_b(s_1, s_2, x_1, x_2, y_1, y_2) = 0\), the constraint (D6) is trivially satisfied. To verify feasibility with respect to (D7), we have

\[\mu^{(2)}_c(s_2, s_1, y_2) + \mu^{(2)}_c(x_1, s_1, s_2) = \sum_{y_1} \lambda^{(12)}_c(x_1, s_1, s_2, y_1) \mathbb{I}[x_2 = y_2] = \sum_{y_1, y_2} \lambda^{(1)}_c(s_1, s_2, x_1, x_2, y_1, y_2) = \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, s_2, s_2, \tilde{s}_2),
\]

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

**Proof of Theorem 11:** It is enough to show that the choice of dual variables in (29) 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:** \(\mathbb{I}(s_1, s_2) \neq (\tilde{s}_1, \tilde{s}_2)\) holds. In this case, \(\lambda^{(12)}_b(s_1, s_2, x_1, x_2, y_1, y_2, \tilde{s}_1, \tilde{s}_2) = 0\) and \(\lambda^{(12)}_c(s_1, s_2, x_1, x_2, y_1, y_2, \tilde{s}_1, \tilde{s}_2) = 0\). The LHS of (D4) becomes

\[\lambda^{(12)}_c(s_1, s_2, x_1, x_2, y_1, y_2, \tilde{s}_1, \tilde{s}_2) = \mathbb{I}(x_1, s_1, s_2, y_1, y_2) \mathbb{I}[x_2 = y_2] = \mathbb{I}(s_1, s_2) \mathbb{I}[x_2 = y_2],
\]

which is the RHS of (D4) thereby satisfying the constraint.

**Case 2:** \(\mathbb{I}(s_1, s_2) = (\tilde{s}_1, \tilde{s}_2)\) holds. In this case, \(\lambda^{(12)}_c(s_1, s_2, x_1, x_2, y_1, y_2, \tilde{s}_1, \tilde{s}_2) = 0\). The LHS of (D4) becomes

\[\lambda^{(12)}_c(s_1, s_2, x_1, x_2, y_1, y_2, \tilde{s}_1, \tilde{s}_2) = \mathbb{I}(x_1, s_1, s_2, y_1, y_2) \mathbb{I}[x_2 = y_2] = \mathbb{I}(s_1, s_2),
\]

which is non-positive, thereby satisfying the constraint (D4). The non-positivity follows since \(\lambda^{(12)}_b(s_1, s_2, x_1, x_2, y_1, y_2)\) and \(\lambda^{(12)}_c(s_1, s_2, x_1, x_2, y_1, y_2)\) satisfy the constraints (B3), (C3) and (A3) (corresponding to the case when \(s_1, s_2 = (\tilde{s}_1, \tilde{s}_2)\)) of dual programs DPS1\(_2\), DPS1\(_{12}\) and DPJE respectively.

To verify feasibility with respect to (D5), \(\lambda^{(12)}_b(s_1, s_2, x_1, x_2, y_1, y_2, \tilde{s}_1, \tilde{s}_2)\) evaluates to

\[\lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_1) \mathbb{I}[x_2 = y_2] + \lambda^{(12)}_c(s_1, s_2, \tilde{s}_1, y_1, y_2) \mathbb{I}[x_2 = y_2] \]

\[= \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_1) \mathbb{I}[x_2 = y_2] + \lambda^{(12)}_c(s_1, s_2, \tilde{s}_1, y_1, y_2) \mathbb{I}[x_2 = y_2] \]

\[\leq \lambda^{(12)}_b(s_1, s_2, \tilde{s}_1, y_1) + \lambda^{(12)}_c(s_1, s_2, x_1, y_1) \mathbb{I}[x_2 = y_2] \]

\[\leq \mathbb{I}(s_1, s_2) \mathbb{I}[x_2 = y_2],
\]

which is the RHS, thereby satisfying (D5). Here, the inequality in (a) follows from the constraint (B3) of DPS1\(_2\).
\[ \geq \mu_c^{(2)}(x_1, s_1, s_2, y_1, y_2) + \mu_c^{(2)}(s_1, \tilde{s}_1, \tilde{s}_2, y_2, y_2), \]

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

To verify the feasibility with respect to (D6), \( \sum_{s_2} \lambda_s^{(2)}(s_1, s_2, \tilde{s}_1, \tilde{s}_2, y_2) \) evaluates to
\[
\left[ \lambda_s^{(2)}(s_1, \tilde{s}_1, \tilde{s}_2, y_2) \right] I(x_1 = y_1) \\
+ (1 - a) \lambda_s^{(2)}(s_1, \tilde{s}_1, \tilde{s}_2, y_2) \left[ I(x_1 = y_1) \\
+ (1 - a) \lambda_s^{(2)}(s_1, \tilde{s}_1, \tilde{s}_2, y_2) \right] I(x_1 = y_1, s_1 = \tilde{s}_1) \\
\geq \mu_c^{(1)}(x_1, s_1, y_1, y_2) + \mu_c^{(1)}(s_1, \tilde{s}_1, \tilde{s}_2, y_1, y_2),
\]

thereby satisfying (D6). The inequality in (b) follows from the constraint (C2) of DPSI, (21). The feasibility with respect to (D7) is trivially satisfied. Hence, the considered choice of dual variables are feasible for DPSW.

**Proof of Theorem 12**: To get to the required converse, take \( \lambda_s^{(1,2)}, \lambda_s^{(1,3)} \) and \( \gamma^{a}, \gamma^{b} \) as in (12) (\( \lambda_s^{(1,2)}, \lambda_s^{(1,3)}, \gamma^{a}, \gamma^{b} \)) as in (16) and \( \lambda_s^{(2)}, \lambda_s^{(3)} \) as in (10) and substitute in (29) to get the values of the variables \( \gamma^{(1,2)}, \lambda_s^{(1,2)}, \mu_c^{(1,2)} \), and \( \mu_s^{(1,2)} \) of DPSW. For the remaining variables, choose the following values of dual variables,
\[
\lambda_c(s_1, s_2, x_1, x_2, y_1, y_2) = \min \{ P_{s_1, s_2}(s_1, s_2), \phi(\tilde{s}_1, s_2), \phi(\tilde{s}_2, s_1) \} \\
\mu_c^{(2)}(s_2, x_2, y_1, y_2) = -\max \{ \phi^{(1)}(\tilde{s}_2, s_1) \} , \\
\mu_c^{(1)}(x_1, s_1, y_1, y_2) = -\max \{ \phi^{(2)}(s_1, \tilde{s}_2) \} , \\
\mu_c^{(21)}(x_1, s_1, s_2) \equiv \min \{ P_{s_1, s_2}(s_1, s_2), \phi(\tilde{s}_1, s_2) + \phi^{(1)}(s_1, s_2), \phi(\tilde{s}_2, s_1) \} \\
\gamma^{(c)}(y_1, y_2) = -\max \{ \phi^{(1)}(s_1, \tilde{s}_2) \} , \\
\gamma^{(b)}(s_2) = -M_1 \max \{ \phi^{(1)}(\tilde{s}_1, s_2) \} \\
+ \sum \mu_c^{(21)}(x_1, s_1, s_2), \\
\gamma^{(d)}(s_1) = -M_2 \max \{ \phi^{(2)}(s_1, \tilde{s}_2) \}, \\
\mu_c^{(1,2)}(x_1, s_1, s_2) \equiv 0.
\]

The above choice of variables can be easily verified to satisfy the constraints in (29).

**Acknowledgments**

The authors would like to sincerely thank the reviewers and the Associate Editor for their comments on an earlier version of this paper.

**References**

[1] S. T. Jose and A. A. Kulkarni, “Linear programming based finite blocklength converses for some network-like problems,” in *Proc. IEEE Inf. Theory Workshop*, Nov. 2017, pp. 544–548.

[2] S. Miyake and F. Kanaya, “Coding theorems on correlated general sources,” *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, 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 Trans. Inf. Theory*, vol. 63, no. 11, pp. 7066–7094, Nov. 2017.

[4] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[5] V. Kostina and S. Verdú, “Fixed-length lossy compression in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3309–3338, Jun. 2012.

[6] V. Kostina and S. Verdú, “Lossy joint source-channel coding in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2545–2575, May 2013.

[7] T. Han, “Information-spectrum methods in information theory,” in *Stochastic Modelling and Applied Probability*, vol. 1. Berlin, Germany: Springer, no. 6, 2003, pp. 1–3.

[8] A. A. Kulkarni and T. P. Coleman, “An optimizer’s approach to stochastic control problems with nonclassical information structures,” *IEEE Trans. Autom. Control*, vol. 60, no. 4, pp. 937–949, Apr. 2015.

[9] M. Conforti, G. Cornuejols, and G. Zambelli, *Integer Programming*, vol. 271. Cham, Switzerland: Springer, 2014.

[10] C. Villani, *Optimal Transport: Old and New*, vol. 338. Berlin, Germany: Springer, 2008.

[11] L. Balzer and R. Timo, “A converse for lossy source coding in the finite blocklength regime,” in *Proc. 24th Int. Zurich Seminar Commun.*, 2016, pp. 15–19.

[12] V. Y. F. Tan, “Asymptotic estimates in information theory with vanishing error probabilities,” *Found. Trends Commun. Inf. Theory*, vol. 11, nos. 1–2, pp. 1–184, Sep. 2014.

[13] G. Vázquez-Vilar, A. T. Campo, A. G. Fábregas, and A. Martínez, “Bayesian M-Ary hypothesis testing: The meta-converse and Verdú-Han bounds are tight,” *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2324–2333, May 2016.

[14] N. Elkayam and M. Feder, “On the calculation of the minimax-converse of the channel coding problem,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)* Jun. 2017, pp. 1247–1251.

**Sharu Theresa Jose** received her B.Tech. in Electrical and Electronics Engineering from TKM College of Engineering, Kerala, India in 2011, M.Tech. in Systems and Control from the Indian Institute of Technology Roorkee in 2013 and Ph.D. from the Systems and Control Engineering group of Indian Institute of Technology Bombay in 2018. Her research interests include information theory, stochastic control and optimization. She is a recipient of the best paper award at the National Conference on Communications, Chennai, 2017, runner up best paper award at the International Conference on Signal Processing and Communication, Bengaluru, 2018 and Dr. Jaykrisna Medal for Outstanding M.Tech. Graduate, 2013 at IIT Roorkee.

**Ankur A. Kulkarni** is an Associate Professor with the Systems and Control Engineering group at Indian Institute of Technology Bombay (IITB). He received his B.Tech. in Aerospace Engineering from IITB in 2006, followed by M.S. in 2008 and Ph.D. in 2010, both from the University of Illinois at Urbana-Champaign (UIUC). From 2010-2012 he was a post-doctoral researcher at the Coordinated Science Laboratory at UIUC. His research interests include information theory, the role of information in stochastic control, game theory, combinatorial coding theory problems, optimization and variational inequalities, and operations research. He is an Associate (from 2015-2018) of the Indian Academy of Sciences, Bangalore, a recipient of the INSPIRE Faculty Award of the Department of Science and Technology, Government of India, 2013, Best paper awards at the National Conference on Communications, Chennai, 2017, and Indian Control Conference 2018, runner up best paper award at the International Conference on Signal Processing and Communication, Bengaluru, 2018, Excellence in Teaching Award 2018 at IITB and the William A. Chittenden Award, 2008 at UIUC. He is a consultant to the Securities and Exchange Board of India on regulation of high frequency trading.