



HAL
open science

A Converse Bound on the Mismatched Distortion-Rate Function

Maël Le Treust, Tristan Tomala

► **To cite this version:**

Maël Le Treust, Tristan Tomala. A Converse Bound on the Mismatched Distortion-Rate Function. ISIT 2024 - IEEE International Symposium on Information Theory, Jul 2024, Athènes, Greece. pp.1-6. hal-04679207

HAL Id: hal-04679207

<https://hal.science/hal-04679207v1>

Submitted on 27 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A Converse Bound on the Mismatched Distortion-Rate Function

Maël Le Treust

Univ. Rennes, CNRS, Inria, IRISA UMR 6074
F-35000 Rennes, France
Email: mael.le-treust@cnrs.fr

Tristan Tomala

HEC Paris
78351 Jouy-en-Josas, France
Email: tomala@hec.fr

Abstract—We investigate the mismatched distortion-rate problem formulated by Lapidoth in 1997. The single-letter characterization of the distortion-rate tradeoff is an open problem. When the rate is large enough, the mismatched distortion-rate problem is deeply related to the mechanism design problem formulated by Jackson and Sonnenschein in 2007. Inspired by this solution, we formulate a new converse bound that matches the achievability bound in four special cases: 1) when the rate is large enough, 2) when the source is binary and uniform and the decoder output is binary, 3) when the encoder distortion measure does not depend on the source symbols, 4) when the sum of the two distortion measures is zero. An important feature of the model is the stochastic post-processing of the codewords. By carefully selecting the randomness of the decoding function, it is optimal for the encoder to correctly transmit the source.

I. INTRODUCTION

Inspired by decentralized decision problems, we investigate the mismatched distortion-rate problem formulated by Lapidoth [1], in which the encoder and the decoder optimize distinct distortion measures d_0 and d_1 . A single-letter characterization of the optimal distortion-rate trade-off is missing. This problem can be viewed as a Stackelberg game [2] in which the codebook and the decoding function are known by the encoder [3]. Recent achievability bounds have been obtained by Kanabar and Scarlett [4], by using superposition coding and expurgated parallel coding. In [5], the authors pointed out interesting analogies between the mismatched rate-distortion problem and the mismatched channel coding problem, for which converse bounds are provided in [6], [7].

The mismatched distortion-rate problem is deeply related to the mechanism design problem of Jackson and Sonnenschein [8], when the rate is large enough. Before the transmission starts, the decoder commits to a decoding function, in which the number of codewords is unrestricted. By knowing this decoding function, the encoder matches each source realization to the optimal codeword, evaluated with respect to its distortion measure d_0 .

Inspired by the solution of [8], we formulate a new converse bound that matches the achievability bound inspired by [1], in four special cases: 1) when the rate is large enough, 2) when the source is binary and uniform and the decoder output is binary, 3) when the encoder distortion measure does not depend on the source symbols, 4) when the sum of the two distortion measures is always zero.

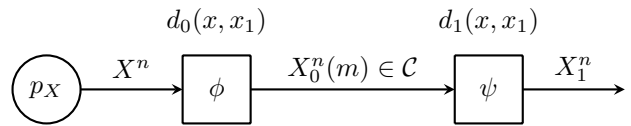


Fig. 1. The mismatched distortion-rate problem.

The problem under study is depicted in Fig. 1, where \mathcal{X} , \mathcal{X}_0 , \mathcal{X}_1 , denote the finite sets of symbols of source, codeword and decoder output. The encoder and decoder optimize two different distortion measures

$$d_0 : \mathcal{X} \times \mathcal{X}_1 \rightarrow \mathbb{R}, \quad d_1 : \mathcal{X} \times \mathcal{X}_1 \rightarrow \mathbb{R}, \quad (1)$$

that are evaluated with respect to the source symbol $x \in \mathcal{X}$ and to the decoder output symbol $x_1 \in \mathcal{X}_1$.

The main difference between our model and the models of Lapidoth [1] and of Scarlett et al. [5, Chap. 3] concerns the distortion measures. In [1], the distortion measure \bar{d}_0 is evaluated with respect to the codeword symbol $x_0 \in \mathcal{X}_0$, rather than to the decoder output symbol $x_1 \in \mathcal{X}_1$, that is

$$\bar{d}_0 : \mathcal{X} \times \mathcal{X}_0 \rightarrow \mathbb{R}, \quad d_1 : \mathcal{X} \times \mathcal{X}_1 \rightarrow \mathbb{R}. \quad (2)$$

In [5, Chap. 3], both distortion measures \bar{d}_0 and \bar{d}_1 are defined with respect to the source symbol $x \in \mathcal{X}$ and to the codeword symbol $x_0 \in \mathcal{X}_0$,

$$\bar{d}_0 : \mathcal{X} \times \mathcal{X}_0 \rightarrow \mathbb{R}, \quad \bar{d}_1 : \mathcal{X} \times \mathcal{X}_0 \rightarrow \mathbb{R}. \quad (3)$$

As discussed on [5, pp. 240], the trivial decoder does not allow post-processing of the codeword in order to generate the final estimate.

We think that the trivial decoder imposes a restriction on the model that makes the solution more complex to characterize. The solution of Jackson and Sonnenschein [8] requires post-processing of the codewords. In addition, this post-processing must be stochastic. By carefully selecting the randomness of the decoding function, it is optimal for the encoder to correctly transmit the source.

Notations. We denote by $\text{conv}(\mathcal{Y})$, the convex hull of \mathcal{Y} . The support of $p_X \in \mathcal{P}(\mathcal{X})$ is denoted by $\text{supp } p_X = \{x \in \mathcal{X}, p_X(x) > 0\}$. We denote by $\mathbb{1}_{\mathcal{X}_1|\mathcal{X}_0} \in \mathcal{P}(\mathcal{X}_1)^{|\mathcal{X}_0|}$ the conditional distribution where $\mathbb{1}_{\mathcal{X}_1|\mathcal{X}_0}(x_1|x_0) = 1$ if $x_1 = x_0$, and 0 otherwise.

II. SYSTEM MODEL

A memoryless source takes values in the finite set \mathcal{X} , according to the probability distribution $p_X \in \mathcal{P}(\mathcal{X})$. The encoder and decoder optimize two different distortion measures

$$d_0 : \mathcal{X} \times \mathcal{X}_1 \rightarrow \mathbb{R}, \quad d_1 : \mathcal{X} \times \mathcal{X}_1 \rightarrow \mathbb{R}, \quad (4)$$

where \mathcal{X}_1 denotes the finite set of decoder output symbols.

Definition 1 A codebook of rate $R \geq 0$ and blocklength $n \in \mathbb{N}^*$ over the finite set \mathcal{X}_0 s.t. $|\mathcal{X}_0| \geq |\mathcal{X}|$, is defined by

$$\mathcal{C} = \left\{ x_0^n(1), \dots, x_0^n(2^{nR}) \right\} \subseteq \mathcal{X}_0^n. \quad (5)$$

The stochastic decoding function is defined by

$$\psi : \mathcal{C} \rightarrow \mathcal{X}_1^n. \quad (6)$$

Given a codebook \mathcal{C} and a stochastic decoding function ψ , the stochastic mismatched encoding function,

$$\phi : \mathcal{X}^n \rightarrow \mathcal{C}, \quad (7)$$

satisfies for all $x^n \in \mathcal{X}^n$,

$$\text{supp } \phi(\cdot | x^n) \subseteq \underset{m \in \{1, \dots, 2^{nR}\}}{\text{argmin}} \sum_{x_1^n} \psi(x_1^n | m) \left(\frac{1}{n} \sum_{t=1}^n d_0(x_t, x_{1,t}) \right). \quad (8)$$

$\mathcal{A}(\mathcal{C}, \psi)$ denotes the set of mismatched encoding functions.

The codebook \mathcal{C} and the encoding and decoding functions ϕ, ψ induce the expected long-run distortion

$$\sum_{x^n, m, x_1^n} p_X^{\otimes n}(x^n) \phi(m | x^n) \psi(x_1^n | m) \left(\frac{1}{n} \sum_{t=1}^n d_0(X_t, X_{1,t}) \right).$$

Definition 2 The pair (R, D_1) is achievable if for all $\varepsilon > 0$, there exists $\bar{n} \in \mathbb{N}^*$, for all $n \geq \bar{n}$, there exists a codebook \mathcal{C} and a decoding function ψ , such that for all $\phi \in \mathcal{A}(\mathcal{C}, \psi)$,

$$\mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] \leq D_1 + \varepsilon. \quad (9)$$

The mismatched distortion-rate function $D_1(R)$ is the infimum of d_1 such that (R, d_1) is achievable.

In the above definition, we adopt the pessimistic tie-breaking rule. In [1] and [5], the definition of achievable pair (R, D_1) relies on the convergence in probability. Since \mathcal{X} and \mathcal{X}_1 are finite sets, this is equivalent to the convergence in expectation of (9), due to Markov's inequality.

Lemma 1 For $n \in \mathbb{N}^*$ and $R \geq 0$, we denote

$$D_1^n(R) = \inf_{\mathcal{C}, \psi} \max_{\phi \in \mathcal{A}(\mathcal{C}, \psi)} \mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right]. \quad (10)$$

We have

$$D_1(R) = \inf_{n \in \mathbb{N}^*} D_1^n(R) = \lim_{n \rightarrow +\infty} D_1^n(R). \quad (11)$$

The proof of Lemma 1 is stated in App. A.

We adapt the achievability result of [1], see also [5], to this model in which both distortion measures in (4) are evaluated with respect to the decoder output symbol $x_1 \in \mathcal{X}_1$.

Definition 3 (Achievability bound)

$$D_1^*(R) = \inf_{p_{X_0}, p_{X_1|X_0}} \max_{\substack{f_{X X_0} \\ \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, R)}} \mathbb{E}_{f_{X X_0} p_{X_1|X_0}} [d_1(X, X_1)], \quad (12)$$

where

$$\mathcal{F}(p_{X_0}, p_{X_1|X_0}, R) = \underset{f_{X X_0} \in \mathcal{D}(p_{X_0}, R)}{\text{argmin}} \mathbb{E}_{f_{X X_0} p_{X_1|X_0}} [d_0(X, X_1)], \quad (13)$$

$$\mathcal{D}(p_{X_0}, R) = \left\{ f_{X X_0} \in \mathcal{P}(\mathcal{X} \times \mathcal{X}_0), f_X = p_X, f_{X_0} = p_{X_0}, I(X; X_0) \leq R \right\}. \quad (14)$$

The infimum in (12) may not be achieved since the function

$$(p_{X_0}, p_{X_1|X_0}) \mapsto \max_{\substack{f_{X X_0} \\ \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, R)}} \mathbb{E}_{f_{X X_0} p_{X_1|X_0}} [d_1(X, X_1)], \quad (15)$$

may not be continuous.

Now, we provide a converse bound inspired by [8].

Definition 4 (Converse bound)

$$\tilde{D}_1(R) = \min_{p_{X_1|X} \in \mathcal{G}(R)} \mathbb{E}_{p_{X_1|X}} [d_1(X, X_1)], \quad (16)$$

where $\mathcal{G}(R)$ is the set of distributions $p_{X_1|X}$ that satisfy:

• *Information constraint:*

$$I(X; X_1) \leq R. \quad (17)$$

• *Mismatched encoding constraint:* For each reporting rule of the encoder $f_{X'|X} \in \mathcal{H}$ that preserve the source distribution

$$\begin{aligned} & \sum_{x, x', x_1} p_X(x) f_{X'|X}(x'|x) p_{X_1|X}(x_1|x') d_0(x, x_1) \\ & \geq \sum_{x, x_1} p_X(x) p_{X_1|X}(x_1|x) d_0(x, x_1), \end{aligned} \quad (18)$$

$$\mathcal{H} = \left\{ f_{X'|X} \in \mathcal{P}(\mathcal{X})^{|\mathcal{X}|}, \forall x' \in \mathcal{X}, \sum_x p_X(x) f_{X'|X}(x'|x) = p_X(x') \right\}. \quad (19)$$

• *Tie-breaking rule:* $\forall f_{X'|X} \in \mathcal{H}$ satisfying (18) with equality,

$$\begin{aligned} & \sum_{x, x', x_1} p_X(x) f_{X'|X}(x'|x) p_{X_1|X}(x_1|x') d_1(x, x_1) \\ & \leq \sum_{x, x_1} p_X(x) p_{X_1|X}(x_1|x) d_1(x, x_1). \end{aligned} \quad (20)$$

The set $\mathcal{G}(R)$ is compact convex since it is the intersection of compact convex sets, thus the infimum in (16) is a minimum.

Theorem 1 The mismatched distortion-rate function satisfies

$$\tilde{D}_1(R) \leq D_1(R) \leq D_1^*(R). \quad (21)$$

III. ACHIEVABILITY PROOF OF THEOREM 1

Fix $p_{X_1|X_0}$ and define the auxiliary distortion measures

$$\tilde{d}_0(x, x_0) = \sum_{x_1} p_{X_1|X_0}(x_1|x_0) d_0(x, x_1), \quad (22)$$

$$\tilde{d}_1(x, x_0) = \sum_{x_1} p_{X_1|X_0}(x_1|x_0) d_1(x, x_1). \quad (23)$$

This implies that the set of distributions $\tilde{\mathcal{F}}(p_{X_0}, \mathbb{R})$ defined in [5, (4.6)] coincides with $\mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})$ defined in (24).

$$\tilde{\mathcal{F}}(p_{X_0}, \mathbb{R}) = \underset{f_{X X_0} \in \mathcal{D}(p_{X_0}, \mathbb{R})}{\operatorname{argmin}} \mathbb{E}_{f_{X X_0}} [\tilde{d}_0(X, X_0)] \quad (24)$$

$$= \mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R}). \quad (25)$$

According to [5, Theorem 4.2] derived from [1],

$$D_1(\mathbb{R}) \leq \inf_{p_{X_0}} \max_{f_{X X_0} \in \tilde{\mathcal{F}}(p_{X_0}, \mathbb{R})} \mathbb{E}_{f_{X X_0}} [\tilde{d}_1(X, X_1)], \quad (26)$$

$$= \inf_{p_{X_0}} \max_{f_{X X_0} \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})} \mathbb{E}_{f_{X X_0} p_{X_1|X_0}} [d_1(X, X_1)]. \quad (27)$$

Equations (26) and (27) are valid for all conditional distribution $p_{X_1|X_0}$. Therefore $D_1(\mathbb{R}) \leq D_1^*(\mathbb{R})$.

IV. CONVERSE PROOF OF THEOREM 1

The main idea is to adapt the ‘‘revelation principle’’ of [8]: the decoder implements $p_{X_1|X'} \circ f_{X'|X}$ when a reporting rule $f_{X'|X} \in \mathcal{H}$ is optimal for the encoder.

We consider a blocklength $n \in \mathbb{N}^*$, a codebook \mathcal{C} , a decoding function ψ , and an encoding function $\phi \in \mathcal{A}(\mathcal{C}, \psi)$ that achieves the maximum in

$$\max_{\phi' \in \mathcal{A}(\mathcal{C}, \psi)} \mathbb{E}_{\mathcal{C}, \phi', \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right]. \quad (28)$$

We define the average distribution

$$q_{X X_1}(x, x_1) = \frac{1}{n} \sum_{t=1}^n p_{X_t X_{1,t}}(x, x_1), \quad \forall (x, x_1), \quad (29)$$

where $p_{X_t X_{1,t}}$ denotes the distribution of $(X_t, X_{1,t})$ induced by $p_X^{\otimes n}(x^n) \phi(m|x^n) \psi(x_1^n|m)$. Note that $q_X = p_{X_t} = p_X$ and

$$\mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] = \mathbb{E}_{q_{X_1|X}} [d_1(X, X_1)]. \quad (30)$$

Lemma 2 *The conditional distribution $q_{X_1|X} \in \mathcal{G}(\mathbb{R})$.*

Proof. [Lemma 2] *Information constraint.*

$$\begin{aligned} & I_{q_{X_1|X}}(X; X_1) \\ & \leq \frac{1}{n} \sum_{t=1}^n I_{p_{X_t X_{1,t}}}^t(X_t; X_{1,t}) \leq \frac{1}{n} \sum_{t=1}^n I(X_t; X_1^n, X^{t-1}) \end{aligned} \quad (31)$$

$$= \frac{1}{n} \sum_{t=1}^n I(X_t; X_1^n | X^{t-1}) = \frac{1}{n} I(X^n; X_1^n) \quad (32)$$

$$\leq \frac{1}{n} I(X^n; X_0^n) \leq \mathbb{R}. \quad (33)$$

Equation (31) comes from the convexity of $I(X; X_1)$ in $q_{X_1|X}$ for fixed p_X , (32) comes from the i.i.d. property of the source, (33) comes from the Markov chain $X^n \circlearrowleft X_0^n \circlearrowleft X_1^n$ and the codebook cardinality $H(X_0^n) \leq \log |\mathcal{C}| = n\mathbb{R}$.

Mismatched encoding constraint. For all $f_{X'|X} \in \mathcal{H}$, we consider the encoding function

$$\tilde{\phi}(m|x^n) = \sum_{x'^n} f_{X'|X}^{\otimes n}(x'^n|x^n) \phi(m|x'^n). \quad (34)$$

As in (29), the encoding $\tilde{\phi}(m|x^n)$ induces a sequence of marginal distributions $(\tilde{p}_{X_t X'_t X_{1,t}})_{t=\{1, \dots, n\}}$, and an average distribution $\tilde{q}_{X X' X_1}$ defined, for all (x, x', x_1) , by

$$\tilde{q}_{X X' X_1}(x, x', x_1) = \frac{1}{n} \sum_{t=1}^n \tilde{p}_{X_t X'_t X_{1,t}}(x, x', x_1) \quad (35)$$

$$= p_X(x) f_{X'|X}(x'|x) q_{X_1|X}(x_1|x'). \quad (36)$$

Equation (36) holds since (X^n, X'^n) are generated i.i.d. according to $p_X^{\otimes n} f_{X'|X}^{\otimes n}$, and both marginal on X and X' are equal, i.e. $\sum_x p_X(x) f_{X'|X}(x'|x) = p_X(x')$ for all $x' \in \mathcal{X}$.

Therefore, $\tilde{\phi}(m|x^n)$ induces the expected distortion

$$\sum_{x, x', x_1} p_X(x) f_{X'|X}(x'|x) q_{X_1|X}(x_1|x') d_0(x, x_1) \quad (37)$$

$$= \mathbb{E}_{\mathcal{C}, \tilde{\phi}, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_0(X_t, X_{1,t}) \right] \geq \mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_0(X_t, X_{1,t}) \right] \quad (38)$$

$$= \sum_{x, x_1} p_X(x) q_{X_1|X}(x_1|x) d_0(x, x_1). \quad (39)$$

Equation (38) comes from the hypothesis $\phi \in \mathcal{A}(\mathcal{C}, \psi)$.

Tie-breaking rule. Let assume that $f_{X'|X} \in \mathcal{H}$ satisfies (18) with equality. The encoding function $\tilde{\phi}(m|x^n)$ constructed in (34) satisfies (38) with equality, thus $\tilde{\phi} \in \mathcal{A}(\mathcal{C}, \psi)$. Moreover,

$$\sum_{x, x', x_1} p_X(x) f_{X'|X}(x'|x) q_{X_1|X}(x_1|x') d_1(x, x_1) \quad (40)$$

$$= \mathbb{E}_{\mathcal{C}, \tilde{\phi}, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] \leq \mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] \quad (41)$$

$$= \sum_{x, x_1} p_X(x) q_{X_1|X}(x_1|x) d_1(x, x_1). \quad (42)$$

Equation (41) holds since $\tilde{\phi}$ achieves the maximum in (28). This concludes the proof of Lemma 2. \square

Since $q_{X_1|X} \in \mathcal{G}(\mathbb{R})$, we have

$$\mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] = \mathbb{E}_{q_{X_1|X}} [d_1(X, X_1)] \quad (43)$$

$$\geq \min_{p_{X_1|X} \in \mathcal{G}(\mathbb{R})} \mathbb{E}_{p_{X_1|X}} [d_1(X, X_1)] = \tilde{D}_1(\mathbb{R}). \quad (44)$$

Equation (44) holds for all $n \in \mathbb{N}^*$, \mathcal{C} , ψ , and $\phi \in \mathcal{A}(\mathcal{C}, \psi)$ that achieves the maximum in (28). We conclude that

$$D_1(\mathbb{R}) = \inf_{n \in \mathbb{N}^*} D_1^n(\mathbb{R}) \quad (45)$$

$$= \inf_{n \in \mathbb{N}^*} \inf_{\mathcal{C}, \psi} \max_{\phi \in \mathcal{A}(\mathcal{C}, \psi)} \mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] \quad (46)$$

$$\geq \min_{p_{X_1|X} \in \mathcal{G}(\mathbb{R})} \mathbb{E}_{p_{X_1|X}} \left[d_1(X, X_1) \right] = \tilde{D}_1(\mathbb{R}). \quad (47)$$

V. CASES WHERE THE TWO BOUNDS MATCH

Although we do not know whether $D_1(\mathbb{R}) = \tilde{D}_1(\mathbb{R})$ holds in general, we present four cases where these bounds coincide.

A. Large Information Rate

When $\mathbb{R} \geq \log_2 |\mathcal{X}|$, our model is similar to the ‘‘mechanism design problem’’ under study in [8], where $\tilde{D}_1(\mathbb{R}) = D_1^*(\mathbb{R})$.

Theorem 2 *If $\mathbb{R} \geq H(X)$, then $\tilde{D}_1(\mathbb{R}) = D_1(\mathbb{R}) = D_1^*(\mathbb{R})$.*

Proof. [Theorem 2] The idea is to take the optimal distribution $p_{X_1|X}^* \in \mathcal{G}(\mathbb{R})$ for (16) and to identify $X_0 = X$. It is optimal for the encoder to reveal truthfully the source sequence.

We consider the distributions $f_{X X_0} = p_X \mathbb{1}_{X_0|X}$ and $p_{X_0 X_1} = p_X p_{X_1|X}^*$, and we show that $f_{X X_0} \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})$ achieves the maximum in (15).

The information constraint is satisfied since $I(X; X_0) = H(X) \leq \mathbb{R}$. As a consequence, there is a one-to-one correspondence between $\mathcal{D}(p_{X_0}, \mathbb{R})$ and \mathcal{H} . According to (18),

$$\sum_{x, x_0, x_1} p_X(x) \mathbb{1}_{X_0|X}(x_0|x) p_{X_1|X}^*(x_1|x_0) d_0(x, x_1) \quad (48)$$

$$= \sum_{x, x_1} p_X(x) p_{X_1|X}^*(x_1|x) d_0(x, x_1) \quad (49)$$

$$= \min_{\tilde{f}_{X'|X} \in \mathcal{H}} \mathbb{E}_{\tilde{f}_{X'|X} p_{X_1|X}^*} \left[d_0(X, X_1) \right] \quad (50)$$

$$= \min_{f'_{X X_0} \in \mathcal{D}(p_{X_0}, \mathbb{R})} \mathbb{E}_{f'_{X X_0} p_{X_1|X_0}} \left[d_0(X, X_1) \right]. \quad (51)$$

Hence $f_{X X_0} = p_X \mathbb{1}_{X_0|X} \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})$. According to (20), we now show that $f_{X X_0}$ achieves the maximum in (15).

By contradiction, suppose $f'_{X X_0} \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})$ satisfies

$$\mathbb{E}_{f'_{X X_0} p_{X_1|X_0}} \left[d_1(X, X_1) \right] > \mathbb{E}_{f_{X X_0} p_{X_1|X_0}} \left[d_1(X, X_1) \right].$$

By identification of X_0 with X' , the distribution $\tilde{f}_{X'|X} = f'_{X_0|X}$ belongs to \mathcal{H} , satisfies (18) with equality and moreover

$$\mathbb{E}_{\tilde{f}_{X'|X} p_{X_1|X}^*} \left[d_1(X, X_1) \right] > \mathbb{E}_{p_{X_1|X}^*} \left[d_1(X, X_1) \right], \quad (52)$$

which contradict $p_{X_1|X}^* \in \mathcal{G}(\mathbb{R})$. We conclude that

$$\tilde{D}_1(\mathbb{R}) = \mathbb{E}_{p_{X_1|X}^*} \left[d_1(X, X_1) \right] \quad (53)$$

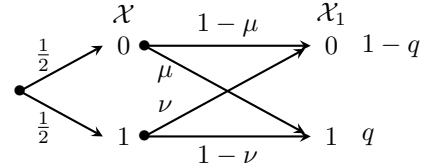
$$= \max_{\substack{f_{X X_0} \\ \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})}} \mathbb{E}_{f_{X X_0} p_{X_1|X_0}} \left[d_1(X, X_1) \right] \geq D_1^*(\mathbb{R}). \quad (54)$$

□

B. Binary Uniform Source and Binary Decoder Output

Theorem 3 *Suppose that $|\mathcal{X}| = |\mathcal{X}_1| = 2$ and $p_X = (\frac{1}{2}, \frac{1}{2})$. Then $\tilde{D}_1(\mathbb{R}) = D_1(\mathbb{R}) = D_1^*(\mathbb{R})$.*

Proof. [Theorem 3] We consider $\mathcal{X} = \mathcal{X}_1 = \{0, 1\}$ and $p_X = (\frac{1}{2}, \frac{1}{2})$. Without loss of generality, we consider $p_{X_1|X}^* \in \mathcal{G}(\mathbb{R})$ optimal in (16) that satisfies $I_{p_{X_1|X}^*}(X; X_1) = \mathbb{R}$, see Fig. with $(\mu, \nu) \in [0, 1]^2$. The idea is to identify $X_0 = X_1$. It is optimal that the encoder tells the decoder which symbol to output.



We define $f_{X X_0}(x, x_0) = p_X(x) p_{X_1|X}^*(x_0|x)$ for all (x, x_0) , with marginal $f_{X_0} = p_{X_1}^* = (1-q, q)$. Note that $I_{f_{X X_0}}(X; X_0) = I_{p_{X_1|X}^*}(X; X_1) = \mathbb{R}$.

Lemma 3 *Let $\sigma_{X'|X}$ be the permutation of the binary state.*

$$\mathcal{D}(f_{X_0}, \mathbb{R}) = \text{conv} \left\{ \{p_X p_{X_1|X}^*\} \cup \{p_X \sigma_{X'|X} p_{X_1|X}^*\} \right\} \quad (55)$$

$$= \text{conv} \left\{ \left(\frac{1-\mu}{2}, \frac{\mu}{2} \right), \left(\frac{\nu}{2}, \frac{1-\nu}{2} \right) \right\}. \quad (56)$$

The proof of Lemma 3 is stated in App. B.

Since both distributions $p_X p_{X_1|X}^*$ and $p_X \sigma_{X'|X} p_{X_1|X}^*$ belongs to the set \mathcal{H} , by Lemma 3 we have

$$\mathcal{D}(f_{X_0}, \mathbb{R}) \subseteq \quad (57)$$

$$\mathcal{J} = \left\{ \sum_{x'} p_X f_{X'|X}(x'|\cdot) p_{X_1|X'}^*(\cdot|x'), f_{X'|X} \in \mathcal{H} \right\}.$$

Moreover, the Markov chain $X \leftrightarrow X' \leftrightarrow X_1 = X_0$ implies that all distribution $p_{X X_0} \in \mathcal{J}$ satisfy $I_{p_{X X_0}}(X; X_0) \geq \mathbb{R} = I_{f_{X X_0}}(X; X_0)$ and has the same marginal distribution $p_{X_0} = f_{X_0}$. Thus $\mathcal{J} \subseteq \mathcal{D}(f_{X_0}, \mathbb{R})$.

Since $p_{X_1|X}^* \in \mathcal{G}(\mathbb{R})$ and $\mathcal{D}(f_{X_0}, \mathbb{R}) = \mathcal{J}$, the distribution $f_{X X_0} = p_X p_{X_1|X}^*$ belongs to $\mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})$ and moreover it achieves the maximum in

$$\max_{f_{X X_0} \in \mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R})} \mathbb{E}_{f_{X X_0} \mathbb{1}_{X_1|X_0}} \left[d_1(X, X_1) \right] = \tilde{D}_1(\mathbb{R}) \geq D_1^*(\mathbb{R}).$$

□

C. Distortion Measure d_0 Only Depends on the Symbols X_1

Following the ‘‘transparent motives’’ terminology of [9], we consider that the encoder distortion measure $d_0 : \mathcal{X} \times \mathcal{X}_1 \rightarrow \mathbb{R}$ does not depends on the source symbol $x \in \mathcal{X}$.

$$\forall (x, x') \in \mathcal{X}^2, \forall x_1 \in \mathcal{X}_1, \quad d_0(x, x_1) = d_0(x', x_1). \quad (58)$$

Theorem 4 *Suppose (58), then $\tilde{D}_1(\mathbb{R}) = D_1(\mathbb{R}) = D_1^*(\mathbb{R})$.*

Proof. [Theorem 4] When (58) is satisfied, the encoder distortion $\mathbb{E}[d_0(X, X_1)]$ only depends on the marginal distribution

p_{X_1} . Because of the tie-breaking rule, the encoder may select the worst encoding function for $\mathbb{E}[d_1(X, X_1)]$, for a given p_{X_1} .

We consider a distribution $p_{X_1|X} \in \mathcal{G}(\mathbb{R})$. Equation (58) implies that (18) is satisfied with equality for all $f_{X'|X} \in \mathcal{H}$. The tie-breaking rule (20) reformulates as

$$\begin{aligned} & \sum_{x, x', x_1} p_X(x) f_{X'|X}(x'|x) p_{X_1|X}(x_1|x') d_1(x, x_1) \\ & \leq \sum_{x, x_1} p_X(x) p_{X_1|X}(x_1|x) d_1(x, x_1), \quad \forall f_{X'|X} \in \mathcal{H}. \end{aligned} \quad (59)$$

Therefore, the solution reformulates

$$\tilde{D}_1(\mathbb{R}) = \min_{p_{X_1|X}, I(X; X_1) \leq \mathbb{R}} \max_{f_{X'|X} \in \mathcal{H}} \mathbb{E}_{f_{X'|X} p_{X_1|X'}} [d_1(X, X_1)] \quad (60)$$

$$= \min_{p_{X_1} \in \mathcal{P}(\mathcal{X}_1)} \mathbb{E}_{p_X p_{X_1}} [d_1(X, X_1)]. \quad (61)$$

Similarly, we can show that $\mathcal{F}(p_X p_{X_1|X}^*, \mathbb{R}) = \mathcal{D}(p_{X_1}^*, \mathbb{R})$.

$$D_1^*(\mathbb{R}) = \min_{p_{X_1} \in \mathcal{P}(\mathcal{X}_1)} \mathbb{E}_{p_X p_{X_1}} [d_1(X, X_1)] = \tilde{D}_1(\mathbb{R}). \quad (62)$$

□

D. Zero-Sum Distortion Measures

In this section, we consider that the distortion measures d_0 and d_1 satisfy the zero-sum property.

$$\forall (x, x_1) \in \mathcal{X} \times \mathcal{X}_1, \quad d_0(x, x_1) + d_1(x, x_1) = 0. \quad (63)$$

Theorem 5 Suppose (63), then $\tilde{D}_1(\mathbb{R}) = D_1(\mathbb{R}) = D_1^*(\mathbb{R})$.

Proof. [Theorem 5] We consider a distribution $p_{X_1|X} \in \mathcal{G}(\mathbb{R})$. Since $d_0(x, x_1) = -d_1(x, x_1)$, the mismatched encoding constraint (18) is equivalent to the tie-breaking rule (59). Also,

$$\mathcal{F}(p_{X_0}, p_{X_1|X_0}, \mathbb{R}) = \operatorname{argmax}_{f_{X X_0} \in \mathcal{D}(p_{X_0}, \mathbb{R})} \mathbb{E}_{f_{X X_0} p_{X_1|X_0}} [d_1(X, X_1)]. \quad (64)$$

Similarly to (60)-(62), both solutions reformulate as

$$\tilde{D}_1(\mathbb{R}) = \min_{p_{X_1} \in \mathcal{P}(\mathcal{X}_1)} \mathbb{E}_{p_X p_{X_1}} [d_1(X, X_1)] = D_1^*(\mathbb{R}). \quad (65)$$

□

APPENDIX A PROOF OF LEMMA 1

According to Def. 2, (\mathbb{R}, D_1) is achievable if for all $\varepsilon > 0$,

$$\inf_{n \in \mathbb{N}^*} \inf_{\mathcal{C}, \psi} \max_{\phi \in \mathcal{A}(\mathcal{C}, \psi)} \mathbb{E}_{\mathcal{C}, \phi, \psi} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] \leq D_1 + \varepsilon.$$

Therefore, $D_1(\mathbb{R}) = \inf_{n \in \mathbb{N}^*} D_1^n(\mathbb{R})$.

Lemma 4 The sequence $(n \cdot D_1^n(\mathbb{R}))_{n \in \mathbb{N}^*}$ is subadditive.

Fekete's Lemma [10] for subadditive sequence concludes that

$$\inf_{n \in \mathbb{N}^*} D_1^n(\mathbb{R}) = \lim_{n \rightarrow +\infty} D_1^n(\mathbb{R}). \quad (66)$$

Proof. [Lemma 4] For all $\varepsilon > 0$, $n \in \mathbb{N}^*$ and $m \in \mathbb{N}^*$, we consider a pairs of codebook of rate $\mathbb{R} \geq 0$ and decoding functions (\mathcal{C}^n, ψ^n) , (\mathcal{C}^m, ψ^m) and a pair of encoding functions $\phi^n \in \mathcal{A}(\mathcal{C}^n, \psi^n)$, $\phi^m \in \mathcal{A}(\mathcal{C}^m, \psi^m)$ that achieve the maximum in

$$\begin{aligned} & \max_{\substack{\phi^n, \phi^m \in \\ \mathcal{A}(\mathcal{C}^n, \psi^n)}} \mathbb{E}_{\mathcal{C}^n, \phi^n, \psi^n} \left[\frac{1}{n} \sum_{t=1}^n d_1(X_t, X_{1,t}) \right] \leq D_1^n(\mathbb{R}) + \varepsilon, \\ & \max_{\substack{\phi^m \in \\ \mathcal{A}(\mathcal{C}^m, \psi^m)}} \mathbb{E}_{\mathcal{C}^m, \phi^m, \psi^m} \left[\frac{1}{m} \sum_{t=1}^m d_1(X_t, X_{1,t}) \right] \leq D_1^m(\mathbb{R}) + \varepsilon. \end{aligned}$$

We construct a codebook $\tilde{\mathcal{C}} = \mathcal{C}^n \times \mathcal{C}^m$ composed of sequences of length $n + m$ defined by

$$\begin{aligned} \tilde{\mathcal{C}} = \{ & (x_0^n(1), x_0^m(1)), \dots, (x_0^n(1), x_0^m(2^{m\mathbb{R}})), \dots, \\ & (x_0^n(2^{n\mathbb{R}}), x_0^m(1)), \dots, (x_0^n(2^{n\mathbb{R}}), x_0^m(2^{m\mathbb{R}})) \} \subseteq \mathcal{X}_0^{n+m}. \end{aligned}$$

We concatenate the encoding $\tilde{\phi} = (\phi^n, \phi^m)$ and decoding functions $\tilde{\psi} = (\psi^n, \psi^m)$. Since the blocks of n and m symbols are treated independently, the concatenated encoding function satisfies $\tilde{\phi} \in \mathcal{A}(\tilde{\mathcal{C}}, \tilde{\psi})$ and achieves the maximum in

$$\max_{\tilde{\phi} \in \mathcal{A}(\tilde{\mathcal{C}}, \tilde{\psi})} \mathbb{E}_{\tilde{\mathcal{C}}, \tilde{\phi}, \tilde{\psi}} \left[\frac{1}{n+m} \sum_{t=1}^{n+m} d_1(X_t, X_{1,t}) \right] \quad (67)$$

$$= (n+m)\varepsilon + n \cdot D_1^n(\mathbb{R}) + m \cdot D_1^m(\mathbb{R}) \quad (68)$$

$$\geq (n+m) \cdot D_1^{n+m}(\mathbb{R}). \quad (69)$$

The subadditivity property of $(n \cdot D_1^n(\mathbb{R}))_{n \in \mathbb{N}^*}$ holds for all (m, n) since equations (67)-(69) are valid for all $\varepsilon > 0$, □

APPENDIX B PROOF OF LEMMA 3

Recall that we consider the distributions p_X and $p_{X_1|X}^*$, we identify $X_0 = X_1$ and we define $f_{X X_0}(x, x_0) = p_X(x) p_{X_1|X}^*(x_0|x)$ for all (x, x_0) , with marginal $f_{X_0} = p_{X_1}^* = (1-q, q)$, with $I_{f_{X X_0}}(X; X_0) = I_{p_{X_1|X}^*}(X; X_1) = \mathbb{R}$. By convexity of the mutual information,

$$\begin{aligned} \mathcal{D}(f_{X_0}, \mathbb{R}) = \operatorname{conv} \{ & f'_{X X_0} \in \mathcal{P}(\mathcal{X} \times \mathcal{X}_0), f'_X = p_X, \\ & f'_{X_0} = f_{X_0}, I(X; X_0) = \mathbb{R} \}. \end{aligned} \quad (70)$$

The distributions $f'_{X X_0}$ with marginal $f'_X = p_X = (\frac{1}{2}, \frac{1}{2})$ and $f'_{X_0} = p_{X_1} = (1-q, q)$ have parameters $(\mu, \nu) \in [0, 1]^2$ that satisfy $q = \frac{1}{2}(\mu + 1 - \nu) \Leftrightarrow \nu = \mu + 1 - 2q$.

Suppose that $q \leq \frac{1}{2}$. The function $\mu \mapsto I(X; X_0)$ is strictly decreasing on the interval $[0, q]$ and strictly increasing on $[q, 2q]$, and it takes values $H_b(q)$ when $\mu \in \{0, 2q\}$ and 0 when $\mu = q$. Thus there exists only two values μ^1 and μ^2 are such that $I(X; X_0) = \mathbb{R} \in [0, H_b(q)]$. These two values correspond to the two distributions $p_X p_{X_1|X}^*$ with $\mu^1 = \mu$ and $\sum_{x'} p_X \sigma_{X'|X}(x'|\cdot) p_{X_1|X}^*(\cdot|x')$ with $\mu^2 = \nu$, where $\sigma_{X'|X}$ denotes the binary permutation. Therefore any distribution $f'_{X X_0} \in \mathcal{D}(f_{X_0}, \mathbb{R})$ is a convex combination of the two distributions $p_X p_{X_1|X}^*$ and $\sum_{x'} p_X \sigma_{X'|X}(x'|\cdot) p_{X_1|X}^*(\cdot|x')$. Similar arguments lead to the same result for $q \geq \frac{1}{2}$. This concludes the proof of Lemma 3.

REFERENCES

- [1] A. Lapidoth, "On the role of mismatch in rate distortion theory," *IEEE Transactions on Information Theory*, vol. 43, no. 1, pp. 38–47, Jan. 1997.
- [2] H. von Stackelberg, *Market Structure and Equilibrium*. Springer-Verlag Berlin and Heidelberg, 2011.
- [3] A. S. Vora and A. A. Kulkarni, "Achievable rates for strategic communication," in *2020 IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 1379–1384.
- [4] M. Kanabar and J. Scarlett, "Mismatched rate-distortion theory: Ensembles, bounds, and general alphabets," *IEEE Transactions on Information Theory*, pp. 1–1, 2024.
- [5] J. Scarlett, A. Guillén i Fàbregas, A. Somekh-Baruch, and A. Martinez, "Information-theoretic foundations of mismatched decoding," *Foundations and Trends in Communications and Information Theory*, vol. 17, no. 2–3, pp. 149–401, 2020.
- [6] A. Somekh-Baruch, "A single-letter upper bound on the mismatch capacity via multicast transmission," *IEEE Transactions on Information Theory*, vol. 68, no. 5, pp. 2801–2812, 2022.
- [7] E. Asadi Kangarshahi and A. Guillén i Fàbregas, "A single-letter upper bound to the mismatch capacity," *IEEE Transactions on Information Theory*, vol. 67, no. 4, pp. 2013–2033, 2021.
- [8] M. O. Jackson and H. F. Sonnenschein, "Overcoming incentive constraints by linking decisions," *Econometrica*, vol. 75, no. 1, pp. 241 – 257, Jan. 2007.
- [9] E. Lipnowski and D. Ravid, "Cheap talk with transparent motives," *Econometrica*, vol. 88, no. 4, pp. 1631–1660, 2020.
- [10] M. Fekete, "Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten," *Mathematische Zeitschrift*, vol. 17, no. 1, pp. 228–249, 1923.