# Point-to-Point Strategic Communication

Maël Le Treust
ETIS UMR 8051, CY Cergy Paris Université, ENSEA, CNRS,
6, avenue du Ponceau,
95014 Cergy-Pontoise CEDEX, France
Email: mael.le-treust@ensea.fr

Tristan Tomala
HEC Paris, GREGHEC UMR 2959
1 rue de la Libération,
78351 Jouy-en-Josas CEDEX, France
Email: tomala@hec.fr



***Abstract*—We propose a strategic formulation for the joint source-channel coding problem in which the encoder and the decoder are endowed with distinct distortion functions. We provide the solutions in four different scenarios. First, we assume that the encoder and the decoder cooperate in order to achieve a certain pair of distortion values. Second, we suppose that the encoder commits to a strategy whereas the decoder implements a best response, as in the persuasion game where the encoder is the Stackelberg leader. Third, we consider that the decoder commits to a strategy, as in the mismatched rate-distortion problem or as in the mechanism design framework. Fourth, we investigate the cheap talk game in which the encoding and the decoding strategies form a Nash equilibrium.**

## I. Introduction

Strategic communication takes place when an informed sender communicates with a receiver that takes an action, given that the sender and the receiver optimize different metrics. This question was originally formulated in the game theory literature were the messages are costless and the communication is unrestricted. Crawford and Sobel [1] investigate the Nash equilibrium of the cheap talk game, whereas Kamenica and Gentzkow [2] introduce the Bayesian persuasion game in which the sender commits to an information disclosure policy, as the leader of the Stackelberg game. In a previous work [3], we characterize the solution of the Bayesian persuasion game when the communication channel is noisy.

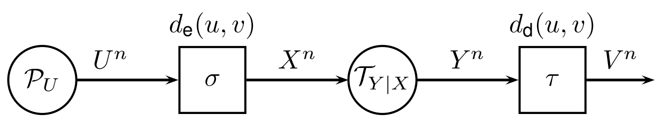


Fig. 1. The source is i.i.d. and the channel is memoryless. The encoder and the decoder have mismatched distortion functions $d_{\mathsf{e}}(u,v) \neq d_{\mathsf{d}}(u,v)$.

The strategic communication problem has attracted attention in computer science [4], in control theory [5], in information theory [6], [7], [8], [9] and is related to the lossy source coding with mismatch distortion functions [10], [11]. Recently, Vora and Kulkarni investigate a strategic communication problem in which the receiver is the Stackelberg leader that should recover the source sequence [12]. The authors introduce the notion of the "information extraction capacity" and formulate an elegant solution in terms of the zero error capacity of "the sender graph" [13].

In this paper, we compare four different solutions for the point-to-point strategic communication problem, and we characterize the set of Nash equilibrium distortions.

## II. System model

We denote by $\mathcal{U}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{V}$, the finite sets of information source, channel inputs, channel outputs and decoder's outputs. Uppercase letters $U^n = (U_1,\ldots,U_n) \in \mathcal{U}^n$ and $X^n$, $Y^n$, $V^n$ stand for $n$-length sequences of random variables with $n \in \mathbb{N}^\star = \mathbb{N} \setminus \{0\}$, whereas lowercase letters $u^n = (u_1,\ldots,u_n) \in \mathcal{U}^n$ and $x^n$, $y^n$, $v^n$, stand for sequences of realizations. We denote by $\Delta(\mathcal{X})$ the set of probability distributions $\mathcal{Q}_X$ over $\mathcal{X}$, i.e. the probability simplex. We consider an i.i.d. information source and a memoryless channel distributed according to $\mathcal{P}_U \in \Delta(\mathcal{U})$ and $\mathcal{T}_{Y|X} : \mathcal{X} \to \Delta(\mathcal{Y})$, as depicted in Fig. 1.

**Definition 1** *We define the encoding strategy $\sigma : \mathcal{U}^n \longrightarrow \Delta(\mathcal{X}^n)$ and the decoding strategy $\tau : \mathcal{Y}^n \longrightarrow \Delta(\mathcal{V}^n)$, and we denote by $\mathcal{P}^{\sigma,\tau}$ the distribution defined by*

$$\mathcal{P}^{\sigma,\tau} = \bigg(\prod_{t=1}^{n} \mathcal{P}_{U_t}\bigg)\sigma_{X^n|U^n}\bigg(\prod_{t=1}^{n} \mathcal{T}_{Y_t|X_t}\bigg)\tau_{V^n|Y^n}, \qquad (1)$$

*where $\sigma_{X^n|U^n}$, $\tau_{V^n|Y^n}$ denote the distributions of $\sigma$, $\tau$.*

**Definition 2** *The encoder and decoder distortion functions $d_{\mathsf{e}} : \mathcal{U} \times \mathcal{V} \longrightarrow \mathbb{R}$ and $d_{\mathsf{d}} : \mathcal{U} \times \mathcal{V} \longrightarrow \mathbb{R}$ induce long-run distortion functions $d_{\mathsf{e}}^n(\sigma,\tau)$ and $d_{\mathsf{d}}^n(\sigma,\tau)$ defined by*

$$d_{\mathsf{d}}^n(\sigma,\tau) = \sum_{u^n,v^n} \mathcal{P}^{\sigma,\tau}\big(u^n,v^n\big) \cdot \Bigg[\frac{1}{n}\sum_{t=1}^{n} d_{\mathsf{d}}(u_t,v_t)\Bigg]. \qquad (2)$$

## III. Cooperative scenario

**Definition 3** *The pair $(D_{\mathsf{e}}, D_{\mathsf{d}})$ is achievable if*

$$\forall \varepsilon > 0,\ \ \exists \bar{n} \in \mathbb{N}^\star,\ \ \forall n \geq \bar{n},\ \ \exists (\sigma,\tau) \qquad (3)$$

$$\text{s.t.} \quad |D_{\mathsf{e}} - d_{\mathsf{e}}^n(\sigma,\tau)| + |D_{\mathsf{d}} - d_{\mathsf{d}}^n(\sigma,\tau)| \leq \varepsilon \qquad (4)$$

*We denote by $\mathcal{C}$ the set of achievable pairs $(D_{\mathsf{e}}, D_{\mathsf{d}})$.*

We define the set of distributions

$$\mathbb{Q}_1 = \Big\{\mathcal{P}_U\mathcal{Q}_{V|U} \text{ s.t. } \max_{\mathcal{P}_X} I(X;Y) - I(U;V) \geq 0\Big\}. \qquad (5)$$

**Theorem 1 (Cooperative scenario)**

$$\mathcal{C} = \Big\{\big(\mathbb{E}_{\mathcal{Q}}[d_{\mathsf{e}}(U,V)], \mathbb{E}_{\mathcal{Q}}[d_{\mathsf{d}}(U,V)]\big) \quad \mathcal{Q} \in \mathbb{Q}_1\Big\}. \qquad (6)$$

The proof of Theorem 1 follows from Shannon's separation result [15, Theorem 3.7], with two distortion functions.

## IV. Persuasion game: encoder commitment

In this section, the encoder chooses first a strategy $\sigma$, and the decoder selects a best response strategy $\tau$ accordingly. This corresponds to the Bayesian persuasion game [2], where the encoder is the Stackelberg leader.

**Definition 4** *Given $n \in \mathbb{N}^\star$, we define*
*1. the set of decoder best responses to strategy $\sigma$ by*

$$\mathsf{BR}_{\mathsf{d}}(\sigma) = \underset{\tau}{\operatorname{argmin}}\ d_{\mathsf{d}}^n(\sigma,\tau), \tag{7}$$

*2. the long-run encoder distortion value by*

$$D_{\mathsf{e}}^n = \inf_{\sigma} \max_{\tau \in \mathsf{BR}_{\mathsf{d}}(\sigma)} d_{\mathsf{e}}^n(\sigma,\tau). \tag{8}$$

In case $\mathsf{BR}_{\mathsf{d}}(\sigma)$ is not a singleton, we assume that the decoder selects the worst strategy for the encoder distortion $\max_{\tau \in \mathsf{BR}_{\mathsf{d}}(\sigma)} d_{\mathsf{e}}^n(\sigma,\tau)$, so that the solution is robust to the exact specification of the decoding strategy.

We aim at characterizing the asymptotic behavior of $D_{\mathsf{e}}^n$.

**Definition 5** *We consider an auxiliary random variable $W \in \mathcal{W}$ with $|\mathcal{W}| = \min\big(|\mathcal{U}|+1, |\mathcal{V}|\big)$ and we define*

$$\mathbb{Q}_2 = \Big\{\mathcal{P}_U \mathcal{Q}_{W|U} \text{ s.t. } \max_{\mathcal{P}_X} I(X;Y) - I(U;W) \geq 0\Big\}. \tag{9}$$

*Given $\mathcal{Q}_{UW}$, we define the single-letter decoder best responses*

$$\mathbb{Q}_{\mathsf{d}}\big(\mathcal{Q}_{UW}\big) = \underset{\mathcal{Q}_{V|W}}{\operatorname{argmin}}\ \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{Q}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big]. \tag{10}$$

*The encoder optimal distortion $D_{\mathsf{e}}^\star$ is given by*

$$D_{\mathsf{e}}^\star = \inf_{\mathcal{Q}_{UW} \in \mathbb{Q}_2} \max_{\substack{\mathcal{Q}_{V|W} \in \\ \mathbb{Q}_{\mathsf{d}}(\mathcal{Q}_{UW})}} \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{Q}_{V|W}}}\Big[d_{\mathsf{e}}(U,V)\Big]. \tag{11}$$

**Theorem 2 (Encoder commitment, Theorem 3.1 in [3])**

$$\forall n \in \mathbb{N}^\star, \qquad D_{\mathsf{e}}^n \geq D_{\mathsf{e}}^\star, \tag{12}$$

$$\forall \varepsilon > 0,\ \exists \bar{n} \in \mathbb{N}^\star,\ \forall n \geq \bar{n}, \qquad D_{\mathsf{e}}^n \leq D_{\mathsf{e}}^\star + \varepsilon. \tag{13}$$

Theorem 2 is a particular case of [9, Theorem III.3] when no side information is available at the decoder. Note that the sequence $(D_{\mathsf{e}}^n)_{n\in\mathbb{N}^\star}$ is sub-additive. Indeed, when $\sigma$ is the concatenation of several encoding strategies, the concatenation of the corresponding optimal decoding strategies still belongs to $\mathsf{BR}_{\mathsf{d}}(\sigma)$. Theorem 2 and Fekete's lemma, show that

$$D_{\mathsf{e}}^\star = \lim_{n\to+\infty} D_{\mathsf{e}}^n = \inf_{n\in\mathbb{N}^\star} D_{\mathsf{e}}^n. \tag{14}$$

**Remark 1** *The decoder long-run distortion $d_{\mathsf{d}}^n(\sigma,\tau)$ obtained with $\sigma$ asymptotically optimal for* (8) *and $\tau \in \mathsf{BR}_{\mathsf{d}}(\sigma)$ converges to $\mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{Q}_{V|W}}}\big[d_{\mathsf{d}}(U,V)\big]$, where $\mathcal{Q}_{V|W} \in \mathbb{Q}_{\mathsf{d}}\big(\mathcal{Q}_{UW}\big)$ and $\mathcal{Q}_{UW}$ is a limit of a minimizing sequence of* (11).

## V. Mechanism design: decoder commitment

In this section, it is the decoder which chooses first a strategy $\tau$, and then the encoder selects a strategy $\sigma$ accordingly. This corresponds to the mismatched rate-distortion problem in information theory [10], [11], and to the Mechanism design problem [14] in game theory, where the decoder is the Stackelberg leader.

**Definition 6** *Given $n \in \mathbb{N}^\star$, we define*
*1. the set of encoder best responses to strategy $\tau$ by*

$$\mathsf{BR}_{\mathsf{e}}(\tau) = \underset{\sigma}{\operatorname{argmin}}\ d_{\mathsf{e}}^n(\sigma,\tau), \tag{15}$$

*2. the long-run decoder distortion value by*

$$D_{\mathsf{d}}^n = \inf_{\tau} \max_{\sigma \in \mathsf{BR}_{\mathsf{e}}(\tau)} d_{\mathsf{d}}^n(\sigma,\tau). \tag{16}$$

The value $D_{\mathsf{d}}^n$ corresponds to the best distortion the decoder can obtain for fixed $n \in \mathbb{N}^\star$. In case there are several best responses, we assume the encoder selects the worst strategy $\sigma$ for the decoder distortion.

We aim at characterizing the asymptotic behaviour of $D_{\mathsf{d}}^n$

**Definition 7** *Given an auxiliary random variable $W \in \mathcal{W}$ with $|\mathcal{W}| = \min\big(|\mathcal{U}|+1, |\mathcal{V}|\big)$ with distribution $\mathcal{P}_W$, we define*

$$\mathbb{Q}_3(\mathcal{P}_W) = \Big\{\mathcal{Q}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W}) \text{ s.t. } \mathcal{Q}_U = \mathcal{P}_U,$$
$$\mathcal{Q}_W = \mathcal{P}_W \text{ and } \max_{\mathcal{P}_X} I(X;Y) - I(U;W) \geq 0\Big\}. \tag{17}$$

*Given $\mathcal{P}_{WV}$, we define the single-letter encoder best responses*

$$\mathbb{Q}_{\mathsf{e}}(\mathcal{P}_{WV}) = \underset{\mathcal{Q}_{UW}\in\mathbb{Q}_3(\mathcal{P}_W)}{\operatorname{argmin}}\ \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{e}}(U,V)\Big]. \tag{18}$$

*The decoder optimal distortion $D_{\mathsf{d}}^\star$ is given by*

$$D_{\mathsf{d}}^\star = \inf_{\mathcal{P}_{WV}} \max_{\mathcal{Q}_{UW}\in\mathbb{Q}_{\mathsf{e}}(\mathcal{P}_{WV})} \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big]. \tag{19}$$

In both (11) and (19), it is the Stackelberg leader that selects the marginal distribution $\mathcal{P}_W$, whereas the incentive constraints affect the Stackelberg follower. Furthermore, the encoder selects the distribution $\mathcal{Q}_{UW} \in \mathbb{Q}_3(\mathcal{P}_W)$ that satisfies the information constraint and the decoder selects $\mathcal{P}_{V|W}$.

**Theorem 3 (Decoder commitment)**

$$\forall n \in \mathbb{N}^\star, \qquad D_{\mathsf{d}}^n \geq D_{\mathsf{d}}^\star, \tag{20}$$

$$\forall \varepsilon > 0,\ \exists \bar{n} \in \mathbb{N}^\star,\ \forall n \geq \bar{n}, \qquad D_{\mathsf{d}}^n \leq D_{\mathsf{d}}^\star + \varepsilon. \tag{21}$$

The achievability proof of Theorem 3 is provided in App. B, and relies on similar arguments as in [10, Step 1] and [11, Lemma 4.3]. The converse proof is based on standard arguments with the identification of the auxiliary random variable $W = (Y^{T-1}, Y_{T+1}^n, T)$, $T \in \{1,\ldots,n\}$. The sequence $(D_{\mathsf{d}}^n)_{n\in\mathbb{N}^\star}$ is sub-additive, thus Theorem 3 and Fekete's lemma show that

$$D_{\mathsf{d}}^\star = \lim_{n\to+\infty} D_{\mathsf{d}}^n = \inf_{n\in\mathbb{N}^\star} D_{\mathsf{d}}^n. \tag{22}$$

**Remark 2** *The encoder long-run distortion $d_{\mathsf{e}}^n(\sigma,\tau)$ obtained with $\tau$ asymptotically optimal for* (16) *and $\sigma \in \mathsf{BR}_{\mathsf{e}}(\tau)$ converges to $\mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{P}_{V|W}}}\big[d_{\mathsf{e}}(U,V)\big]$, where $\mathcal{Q}_{UW} \in \mathbb{Q}_{\mathsf{e}}(\mathcal{P}_{WV})$ and $\mathcal{P}_{WV}$ is a limit of a minimizing sequence of* (19).

## VI. Cheap talk game: no commitment

**Definition 8** *Given $\varepsilon \geq 0$ and $n \in \mathbb{N}^\star$, an $\varepsilon$-Nash equilibrium is a pair of strategies $(\sigma,\tau)$ such that*

$$\sigma \in \mathsf{BR}_{\mathsf{e}}^{\varepsilon}(\tau) \quad \text{and} \quad \tau \in \mathsf{BR}_{\mathsf{d}}^{\varepsilon}(\sigma) \quad \text{where,} \tag{23}$$

$$\mathsf{BR}_{\mathsf{e}}^{\varepsilon}(\tau) = \Big\{\sigma,\ \ d_{\mathsf{e}}^n(\sigma,\tau) \leq \min_{\tilde{\sigma}} d_{\mathsf{e}}^n(\tilde{\sigma},\tau) + \varepsilon\Big\}, \tag{24}$$

$$\mathsf{BR}_{\mathsf{d}}^{\varepsilon}(\sigma) = \Big\{\tau,\ \ d_{\mathsf{d}}^n(\sigma,\tau) \leq \min_{\tilde{\tau}} d_{\mathsf{d}}^n(\sigma,\tilde{\tau}) + \varepsilon\Big\}. \tag{25}$$

*We denote by $\mathsf{NE}_{\varepsilon}^n$ the set of distortion pairs $(D_{\mathsf{e}}^{\varepsilon}, D_{\mathsf{d}}^{\varepsilon})$ for which there exists a $\varepsilon$-Nash equilibrium $(\sigma,\tau)$ such that*

$$D_{\mathsf{e}}^{\varepsilon} = d_{\mathsf{e}}^n(\sigma,\tau) \quad \text{and} \quad D_{\mathsf{d}}^{\varepsilon} = d_{\mathsf{d}}^n(\sigma,\tau). \tag{26}$$

*We denote by $\mathsf{NE}^n$ the set of $\mathsf{NE}_{\varepsilon}^n$ with $\varepsilon = 0$.*

**Definition 9** *For $\varepsilon \geq 0$, we define the set of distributions that are $\varepsilon$-best responses for both encoder and decoder.*

$$\begin{aligned}\mathbb{Q}_4^{\varepsilon} = \Big\{&\mathcal{Q}_{UWV} = \mathcal{P}_U\mathcal{Q}_{W|U}\mathcal{Q}_{V|W} \quad \text{s.t.}\\ &\mathcal{Q}_{UW} \in \mathbb{Q}_{\mathsf{e}}^{\varepsilon}(\mathcal{Q}_{WV}), \quad \mathcal{Q}_{V|W} \in \mathbb{Q}_{\mathsf{d}}^{\varepsilon}(\mathcal{Q}_{UW})\Big\},\end{aligned} \tag{27}$$

$$\begin{aligned}\mathbb{Q}_{\mathsf{e}}^{\varepsilon}(\mathcal{Q}_{WV}) = \Big\{&\mathcal{Q}_{UW} \in \mathbb{Q}_3(\mathcal{Q}_W) \ \text{ s.t. } \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{Q}_{V|W}}}\Big[d_{\mathsf{e}}(U,V)\Big]\\ &\leq \min_{\substack{\tilde{\mathcal{Q}}_{UW}\\ \in\mathbb{Q}_3(\mathcal{Q}_W)}} \mathbb{E}_{\substack{\tilde{\mathcal{Q}}_{UW}\\ \mathcal{Q}_{V|W}}}\Big[d_{\mathsf{e}}(U,V)\Big] + \varepsilon\Big\},\end{aligned} \tag{28}$$

$$\begin{aligned}\mathbb{Q}_{\mathsf{d}}^{\varepsilon}(\mathcal{Q}_{UW}) = \Big\{&\mathcal{Q}_{V|W} \ \text{ s.t. } \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{Q}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big]\\ &\leq \min_{\tilde{\mathcal{P}}_{V|W}} \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \tilde{\mathcal{P}}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big] + \varepsilon\Big\}.\end{aligned} \tag{29}$$

*Then, we define*

$$\mathcal{N}^{\varepsilon} = \Big\{\big(\mathbb{E}_{\mathcal{Q}}[d_{\mathsf{e}}(U,V)], \mathbb{E}_{\mathcal{Q}}[d_{\mathsf{d}}(U,V)]\big) \quad \mathcal{Q} \in \mathbb{Q}_4^{\varepsilon}\Big\}. \tag{30}$$

*We denote by $\mathcal{N}$ the set $\mathcal{N}^{\varepsilon}$ with $\varepsilon = 0$.*

**Theorem 4 (Nash equilibrium distortions)**

$$\forall\varepsilon \geq 0,\ \forall n \in \mathbb{N}, \quad \mathsf{NE}_{\varepsilon}^n \subset \mathcal{N}^{\varepsilon}, \tag{31}$$

$$\lim_{\varepsilon\to 0}\lim_{n\to+\infty} \mathsf{NE}_{\varepsilon}^n = \mathcal{N}. \tag{32}$$

Theorem 4 is a consequence of Theorems 2 and 3. If the distribution $\mathcal{P}_U\mathcal{Q}_{W|U}\mathcal{Q}_{V|W}$ have marginals that belong to the sets $\mathbb{Q}_{\mathsf{d}}(\mathcal{Q}_{UW})$ and $\mathbb{Q}_{\mathsf{e}}(\mathcal{P}_{WV})$, then Shannon's encoding and decoding schemes form an $\varepsilon$-Nash equilibrium.

**Conjecture 1**

$$\lim_{n\to+\infty}\lim_{\varepsilon\to 0} \mathsf{NE}_{\varepsilon}^n = \mathcal{N}. \tag{33}$$

## Appendix A
## Preliminary results

**Definition 10** *Given $\mathcal{P}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W})$, tolerance $\delta > 0$, let*

$$B_\delta(\mathcal{P}_{UW}) = \Big\{\mathcal{Q}_{UW} \ \text{s.t.}\ ||\mathcal{Q}_{UW} - \mathcal{P}_{UW}||_1 \leq \delta\Big\}. \tag{34}$$

*We define the set of typical sequences by*

$$T_\delta(\mathcal{P}_{UW}) = \Big\{(u^n, w^n) \ \text{s.t.}\ Q_{UW}^n \in B_\delta(\mathcal{P}_{UW})\Big\}, \tag{35}$$

*where $Q_{UW}^n$ denotes the empirical distribution of $(u^n, w^n)$.*

**Definition 11** *We consider two distributions $\mathcal{P}_U \in \Delta(\mathcal{U})$, $\mathcal{P}_W \in \Delta(\mathcal{W})$, a rate parameter $\mathsf{R} \geq 0$ and a tolerance $\delta \geq 0$. We define the sets*

$$\begin{aligned}\mathbb{Q}_\delta^-(\mathsf{R}) = \Big\{&\mathcal{Q}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W}) \text{ s.t. } ||\mathcal{Q}_U - \mathcal{P}_U||_1 \leq \delta,\\ &||\mathcal{Q}_W - \mathcal{P}_W||_1 \leq \delta \ \text{ and } \ I(U;W) \leq \mathsf{R}\Big\},\end{aligned} \tag{36}$$

$$\begin{aligned}\mathbb{Q}_\delta^+(\mathsf{R}) = \Big\{&\mathcal{Q}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W}) \text{ s.t. } ||\mathcal{Q}_U - \mathcal{P}_U||_1 \leq \delta,\\ &||\mathcal{Q}_W - \mathcal{P}_W||_1 \leq \delta \ \text{ and } \ I(U;W) \geq \mathsf{R}\Big\}.\end{aligned} \tag{37}$$

*We use the notation $\mathbb{Q}_0^-(\mathsf{R})$ and $\mathbb{Q}_0^+(\mathsf{R})$ when $\delta = 0$.*

**Lemma 1 (see Step 1 in [10] and Lemma 4.3 in [11])** *We consider two distributions $\mathcal{P}_U \in \Delta(\mathcal{U})$ and $\mathcal{P}_W \in \Delta(\mathcal{W})$, a rate $\mathsf{R} \geq 0$, a small $\eta > 0$ and $n \in \mathbb{N}^\star$.*

- *We generate a sequence $U^n$ according to $\mathcal{P}_U^{\otimes n}$.*
- *Independently, we generate a family of sequences $\big(W^n(m)\big)_{m\in\{1,\ldots,2^{n\mathsf{R}}\}}$ according to $\mathcal{P}_W^{\otimes n}$.*

*There exists $\bar{\delta}$, for all $\delta < \bar{\delta}$ and for all $\varepsilon > 0$, there exists $\bar{n}$, for all $n \geq \bar{n}$,*

$$\mathbb{P}\bigg(\exists m \in \{1,\ldots,2^{n\mathsf{R}}\}, \quad Q_m^n \in \mathbb{Q}_\delta^+(\mathsf{R}+\eta)\bigg) \leq \varepsilon,$$

*where $Q_m^n$ denotes the empirical distribution of $(U^n, W^n(m))$.*

The provide the proof of Lemma 1 in App. C.

**Lemma 2 (Covering lemma, see Lemma 3.3 in [15])** *We consider a distribution $\mathcal{P}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W})$, a rate parameter $\mathsf{R} = I(U;W) + \eta$ with $\eta > 0$, $n \in \mathbb{N}$.*

- *We generate a sequence $U^n$ according to $\mathcal{P}_U^{\otimes n}$.*
- *Independently, we generate a family of sequences $\big(W^n(m)\big)_{m\in\{1,\ldots,2^{n\mathsf{R}}\}}$ according to $\mathcal{P}_W^{\otimes n}$.*

*There exists $\bar{\delta} > 0$, for all $\delta < \bar{\delta}$ and for all $\varepsilon > 0$, there exists $\bar{n}$, such that for all $n \geq \bar{n}$,*

$$\mathbb{P}\bigg(\exists m \in \{1,\ldots,2^{n\mathsf{R}}\}, \quad ||Q_m^n - \mathcal{P}_{UW}||_1 \leq \delta\bigg) \geq 1 - \varepsilon.$$

**Definition 12** *For $\mathcal{P}_U \in \Delta(\mathcal{U})$, $\mathcal{P}_W \in \Delta(\mathcal{W})$, $\delta > 0$, $\mathsf{R} \geq 0$, and $\mathsf{D} \geq 0$ we define*

$$\begin{aligned}&\mathbb{Q}_\delta(\mathsf{R},\mathsf{D}) = \Big\{\mathcal{Q}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W}) \text{ s.t. } ||\mathcal{Q}_U - \mathcal{P}_U||_1 \leq \delta,\\ &||\mathcal{Q}_W - \mathcal{P}_W||_1 \leq \delta,\ I(U;W) \leq \mathsf{R},\ \mathbb{E}\Big[d_{\mathsf{e}}(U,V)\Big] \leq \mathsf{D}\Big\}.\end{aligned} \tag{38}$$

*We have* $\mathbb{Q}_{\delta}(\mathsf{R},\mathsf{D}) = \mathbb{Q}_{\delta}^{-}(\mathsf{R}) \cap \mathbb{Q}_{\delta}^{\circ}(\mathsf{D})$ *with*

$$\mathbb{Q}_{\delta}^{\circ}(\mathsf{D}) = \Big\{ \mathcal{Q}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W}) \text{ s.t. } ||\mathcal{Q}_U - \mathcal{P}_U||_1 \leq \delta, \\ ||\mathcal{Q}_W - \mathcal{P}_W||_1 \leq \delta \text{ and } \mathbb{E}\Big[d_{\mathsf{e}}(U,V)\Big] \leq \mathsf{D}\Big\}. \quad (39)$$

## Appendix B
## Achievability proof of Theorem 3

If the channel capacity is equal to zero, then a trivial coding scheme satisfies (21). From now on, we assume that the channel capacity is strictly positive, therefore for all $\varepsilon_0 > 0$ there exists $\eta_0 > 0$ and a distribution $\mathcal{P}_{WV}$ such that

$$\Big| D_{\mathsf{d}}^{\star} - \max_{\mathcal{Q}_{UW}\in\mathbb{Q}_{\mathsf{e}}^{\eta_0}(\mathcal{P}_{WV})} \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big]\Big| \leq \varepsilon_0, \quad (40)$$

where

$$\mathbb{Q}_{\mathsf{e}}^{\eta_0}(\mathcal{P}_{WV}) = \operatorname*{argmin}_{\mathcal{Q}_{UW}\in\mathbb{Q}_3^{\eta_0}(\mathcal{P}_W)} \mathbb{E}_{\substack{\mathcal{Q}_{UW}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{e}}(U,V)\Big], \quad (41)$$

$$\mathbb{Q}_3^{\eta_0}(\mathcal{P}_W) = \Big\{\mathcal{Q}_{UW}\in\Delta(\mathcal{U}\times\mathcal{W}) \text{ s.t. } \mathcal{Q}_U = \mathcal{P}_U, \\ \mathcal{Q}_W = \mathcal{P}_W \text{ and } \max_{\mathcal{P}_X} I(X;Y) - I(U;W) \geq 2\eta_0\Big\}. \quad (42)$$

We use the notation $\mathcal{Q}_{UW}$ to refer to the distribution that achieves the maximum in (40), and without loss of generality, we assume that $I(U;W) = \max_{\mathcal{P}_X} I(X;Y) - 2\eta_0$. We introduce the rate parameter $\mathsf{R} = I(U;W) + \eta_0$ and the tolerance of the typical sequences $\delta > 0$. We consider that the decoder implements Shannon's channel decoding and lossy source decoding, see [15, Sec. 3.1 and 3.6], that we denote by $\tau^{\star}$. We denote by $M$ and $m$ the index selected by the encoder, whereas $\hat{M}$ and $\hat{m}$ refer to the index selected by the decoder.

- The random codebooks $(W^n(m), X^n(m))_{m\in\{1,\ldots,2^{n\mathsf{R}}\}}$ are drawn independently according to $\mathcal{P}_W^{\otimes n}$ and $\mathcal{P}_X^{\otimes n}$, where $\mathcal{P}_X$ maximizes the channel capacity.
- The decoder observes the sequence of channel output $Y^n \in \mathcal{Y}^n$ and returns the unique index $\hat{m}$ such that the sequences $\big(Y^n, X^n(\hat{m})\big) \in T_{\delta}(\mathcal{P}_X\mathcal{T}_{Y|X})$ are jointly typical. Otherwise it returns the index 1.
- Then the decoder returns the sequence $W^n(\hat{m})$ corresponding to $\hat{m}$ and draws $V^n$ i.i.d. according to $\mathcal{P}_{V|W}$.

Standard channel coding arguments ensures that

$$\exists\bar{\delta}_1, \forall\delta<\bar{\delta}_1, \forall\varepsilon_1>0, \exists\bar{n}_1\in\mathbb{N}^{\star}, \forall n\geq\bar{n}_1,\ \mathbb{P}(\hat{M}\neq M)\leq\varepsilon_1. \quad (43)$$

Since the encoder is strategic, it selects a best response $\sigma \in \mathsf{BR}_{\mathsf{e}}(\tau^{\star})$ that, for a given $u^n$, returns $x^n$ in order to minimize

$$\sum_{\substack{y^n,v^n\\ \hat{m}}} \mathcal{T}(y^n|x^n)\mathbb{P}(\hat{m}|y^n)\mathcal{P}^{\otimes n}(v^n|w^n(\hat{m}))\frac{1}{n}\sum_{t=1}^{n} d_{\mathsf{e}}(u_t,v_t) \\ = \sum_{\hat{m}} \mathbb{P}(\hat{m}|x^n)\cdot\sum_{u,w} Q_{\hat{m}}^n(u,w)\sum_{v}\mathcal{P}(v|w)d_{\mathsf{e}}(u,v), \quad (44)$$

where $Q_{\hat{m}}^n \in \Delta(\mathcal{U}\times\mathcal{W})$ denotes the empirical distribution of $(u^n, w^n(\hat{m}))$. We denote by $x^{n\star}$ the sequence that minimizes (44) and we denote by

$$Q^{x^n} = \sum_{\hat{m}} \mathbb{P}(\hat{m}|x^n)\cdot Q_{\hat{m}}^n \in \Delta(\mathcal{U}\times\mathcal{W}), \quad (45)$$

the average empirical distribution induced by the input sequence $x^n$. By Lemma 1, for all $\eta_2 > 0$, there exists $\bar{\delta}_2$, for all $\delta < \bar{\delta}_2$ and for all $\varepsilon_2 > 0$, there exists $\bar{n}_2$, for all $n \geq \bar{n}_2$,

$$\mathbb{P}\Big(Q^{X^{n\star}} \notin \mathbb{Q}_{\delta}^{-}(\mathsf{R}+\eta_2)\Big) \leq \mathbb{P}\Big(Q^{X^{n\star}} \in \mathbb{Q}_{\delta}^{+}(\mathsf{R}+\eta_2)\Big) \quad (46)$$

$$+ \mathbb{P}\Big(||Q_U^{X^{n\star}} - \mathcal{P}_U||_1 + ||Q_W^{X^{n\star}} - \mathcal{P}_W||_1 > \delta\Big) \quad (47)$$

$$\leq \mathbb{P}\Big(\exists x^n\in\mathcal{X}^n, \quad Q^{x^n}\in\mathbb{Q}_{\delta}^{+}(\mathsf{R}+\eta_2)\Big) + \varepsilon_2 \quad (48)$$

$$\leq \mathbb{P}\Big(\exists m\in\{1,\ldots,2^{n\mathsf{R}}\},\ Q_m^n\in\mathbb{Q}_{\delta}^{+}(\mathsf{R}+\eta_2)\Big) + \varepsilon_2 \quad (49)$$

$$\leq 2\varepsilon_2. \quad (50)$$

On the other hand, we assume that the encoder implements Shannon's coding scheme $\sigma_{\mathsf{c}}$, by selecting the unique $m$ such that $(U^n, W^n(m)) \in T_{\delta}(\mathcal{Q}_{UW})$, and 1 otherwise. By Lemma 2, there exists $\bar{\delta}_3 > 0$, for all $\delta < \bar{\delta}_3$ and for all $\varepsilon_3 > 0$, there exists $\bar{n}_3$, such that for all $n \geq \bar{n}_3$,

$$\mathbb{P}\Big(\forall m\in\{1,\ldots,2^{n\mathsf{R}}\}, \quad ||Q_m^n - \mathcal{Q}_{UW}||_1 > \delta\Big) \leq \varepsilon_3. \quad (51)$$

The bounds given in (43), (51) imply

$$1-\varepsilon_1-\varepsilon_3 \leq \mathbb{P}\Big(Q^{X^n(m)}\in\mathbb{Q}_{\delta}^{\circ}(\mathsf{D}+\mu)\Big) \quad (52)$$

$$\leq \mathbb{P}\Big(Q^{X^{n\star}}\in\mathbb{Q}_{\delta}^{\circ}(\mathsf{D}+\mu)\Big), \quad (53)$$

with $\mathsf{D} = \min_{\mathcal{Q}_{UW}\in\mathbb{Q}_3^{\eta_0}(\mathcal{P}_W)} \mathbb{E}\big[d_{\mathsf{e}}(U,V)\big]$ and $\mu = \delta\overline{d_{\mathsf{e}}}$ where $\overline{d_{\mathsf{e}}} = \max_{u,v} d_{\mathsf{e}}(u,v)$. Thus for all $\delta \leq \min(\bar{\delta}_1,\bar{\delta}_2,\bar{\delta}_3)$ and $n \geq \max(\bar{n}_1,\bar{n}_2,\bar{n}_3)$ we have

$$\mathbb{P}\Big(Q^{X^{n\star}}\in\mathbb{Q}_{\delta}(\mathsf{R}+\eta_2,\mathsf{D}+\mu)\Big) \quad (54)$$

$$\geq 1-\mathbb{P}\Big(Q^{X^{n\star}}\notin\mathbb{Q}_{\delta}^{-}(\mathsf{R}+\eta_2)\Big) - \mathbb{P}\Big(Q^{X^{n\star}}\notin\mathbb{Q}_{\delta}^{\circ}(\mathsf{D}+\mu)\Big) \quad (55)$$

$$\geq 1-\varepsilon_1-2\varepsilon_2-\varepsilon_3. \quad (56)$$

This shows the existence of a strategy $\tau^{\star}$ with codebook $(w^n(m), x^n(m))_{m\in\{1,\ldots,2^{n\mathsf{R}}\}}$ such that (56) is satisfied. We consider $\sigma \in \mathsf{BR}_{\mathsf{e}}(\tau^{\star})$ that achieves the maximum in (16) and we denote $\overline{d_{\mathsf{d}}} = \max_{u,v} d_{\mathsf{d}}(u,v)$. Form Berge's Maximum Theorem the correspondance $(\delta,\mathsf{R},\mathsf{D}) \mapsto \mathbb{Q}_{\delta}(\mathsf{R},\mathsf{D})$ is continuous, and therefore

$$d_{\mathsf{d}}^n(\sigma,\tau^{\star}) = \mathbb{E}_{\substack{Q^{X^{n\star}}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big] \quad (57)$$

$$\leq \sup_{\substack{\mathcal{P}_{UW}\in\\ \mathbb{Q}_{\delta}(\mathsf{R}+\eta_2,\mathsf{D}+\mu)}} \mathbb{E}_{\substack{\mathcal{P}_{UW}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big] + (\varepsilon_1+2\varepsilon_2+\varepsilon_3)\overline{d_{\mathsf{d}}} \quad (58)$$

$$\leq \sup_{\substack{\mathcal{P}_{UW}\in\\ \mathbb{Q}(\mathsf{R}-\eta_0,\mathsf{D})}} \mathbb{E}_{\substack{\mathcal{P}_{UW}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big] + (\varepsilon_1+2\varepsilon_2+\varepsilon_3+\varepsilon_4)\overline{d_{\mathsf{d}}} \quad (59)$$

$$= \max_{\substack{\mathcal{P}_{UW}\in\\ \mathbb{Q}_{\mathsf{e}}^{\eta_0}(\mathcal{P}_{WV})}} \mathbb{E}_{\substack{\mathcal{P}_{UW}\\ \mathcal{P}_{V|W}}}\Big[d_{\mathsf{d}}(U,V)\Big] + (\varepsilon_1+2\varepsilon_2+\varepsilon_3+\varepsilon_4)\overline{d_{\mathsf{d}}} \quad (60)$$

$$\leq D_{\mathsf{d}}^{\star} + \varepsilon_0 + (\varepsilon_1+2\varepsilon_2+\varepsilon_3+\varepsilon_4)\overline{d_{\mathsf{d}}}. \quad (61)$$

We take $\varepsilon_0$, $\varepsilon_1$, $\varepsilon_2$, $\varepsilon_3$, $\varepsilon_4$, $\delta$, $\eta_2$, $\eta_0$ small and $n \in \mathbb{N}^\star$ large and the achievability result of Theorem 3 follows.

## APPENDIX C
## PROOF OF LEMMA 1

Lemma 3 below ensures for all $\delta > 0$, there exists a family of distributions $(\mathcal{Q}^k_{UW})_{k\in\mathcal{K}} \subset \operatorname{int}\Delta(\mathcal{U}\times\mathcal{W})$ with $|\mathcal{K}| < +\infty$ such that

$$\Delta(\mathcal{U}\times\mathcal{W}) \subset \bigcup_{k\in\mathcal{K}} T_\delta(\mathcal{Q}^k_{UW}), \tag{62}$$

$$\min_{k\in\mathcal{K}} \min_{(u,w)\in\mathcal{U}\times\mathcal{W}} \mathcal{Q}^k(u,w) \geq \frac{\delta}{4(|\mathcal{U}\times\mathcal{W}|-1)}. \tag{63}$$

Thus for all $\delta > 0$, there exists a family of distributions $(\mathcal{Q}^{\tilde{k}}_{UW})_{\tilde{k}\in\widetilde{\mathcal{K}}} \subset \mathbb{Q}^+_\delta(\mathsf{R}+\eta) \cap \operatorname{int}\Delta(\mathcal{U}\times\mathcal{W})$ with $|\widetilde{\mathcal{K}}| < +\infty$ such that (63) is satisfied and

$$\mathbb{Q}^+_\delta(\mathsf{R}+\eta) \subset \bigcup_{\tilde{k}\in\widetilde{\mathcal{K}}} T_\delta(\mathcal{Q}^{\tilde{k}}_{UW}). \tag{64}$$

We choose $\delta < \bar{\delta}$ such that $3\bar{\delta}\log\frac{4(|\mathcal{U}\times\mathcal{W}|-1)}{\bar{\delta}} < \eta$.

$$\begin{aligned}
&\mathbb{P}\bigg(\exists m \in \{1,\ldots,2^{n\mathsf{R}}\} \text{ s.t. } Q^n_m \in \mathbb{Q}^+_\delta(\mathsf{R}+\eta)\bigg) && (65)\\
\leq&\mathbb{P}\bigg(\exists m \text{ s.t. } Q^n_m \in \bigcup_{\tilde{k}\in\widetilde{\mathcal{K}}} T_\delta(\mathcal{Q}^{\tilde{k}}_{UW})\bigg) && (66)\\
=&\mathbb{P}\bigg(\exists \tilde{k}\in\widetilde{\mathcal{K}}, \exists m \text{ s.t. } Q^n_m \in T_\delta(\mathcal{Q}^{\tilde{k}}_{UW})\bigg) && (67)\\
\leq&\sum_{\tilde{k}\in\widetilde{\mathcal{K}}} \sum_{m\in\{1,\ldots,2^{n\mathsf{R}}\}} \sum_{\substack{(u^n,w^n)\in\\ T_\delta(\mathcal{Q}^{\tilde{k}}_{UW})}} \mathcal{P}^{\otimes n}_U(u^n)\mathcal{P}^{\otimes n}_W(w^n) && (68)\\
\leq&|\widetilde{\mathcal{K}}| \cdot 2^{n(\mathsf{R}-I(U;W)+3\delta\log\frac{4(|\mathcal{U}\times\mathcal{W}|-1)}{\delta})} && (69)\\
\leq&|\widetilde{\mathcal{K}}| \cdot 2^{-n(\eta-3\delta\log\frac{4(|\mathcal{U}\times\mathcal{W}|-1)}{\delta})}. && (70)
\end{aligned}$$

Equation (66) comes from (64). Equation (69) comes from (63) with $\min_{u,w}\mathcal{Q}^{\tilde{k}}(u,w) \geq \frac{\delta}{4(|\mathcal{U}\times\mathcal{W}|-1)}$, and Proposition 1 and 2 below. Equation (70) comes from $\mathcal{Q}^{\tilde{k}}_{UW} \in \mathbb{Q}^+_\delta(\mathsf{R}+\eta)$, that induce $\mathsf{R} \leq I(U;W)-\eta$.

Since $|\widetilde{\mathcal{K}}| < +\infty$ and $\eta - 3\delta\log\frac{4(|\mathcal{U}\times\mathcal{W}|-1)}{\delta} > 0$, we choose $n$ large such that $|\widetilde{\mathcal{K}}|\cdot 2^{-n(\eta-3\delta\log\frac{4(|\mathcal{U}\times\mathcal{W}|-1)}{\delta})} \leq \varepsilon$. This concludes the proof of Lemma 1.

**Proposition 1 (see 1. pp. 27 in [15])** *We consider $\mathcal{P}_U \in \Delta(\mathcal{U})$, $n\in\mathbb{N}$, $\delta>0$. For all $u^n \in T_\delta(\mathcal{P}_U)$ we have*

$$2^{-n(H(U)+\delta_1)} \leq \mathcal{P}^{\otimes n}_U(u^n) \leq 2^{-n(H(U)-\delta_1)}, \tag{71}$$

*with $\delta_1 = \log\frac{1}{\min_{u\in\operatorname{supp}\mathcal{P}_U}\mathcal{P}(u)}\cdot\delta$.*

**Proposition 2 (see 2. pp. 27 in [15])** *We consider $\mathcal{P}_{UW} \in \Delta(\mathcal{U}\times\mathcal{W})$, $n\in\mathbb{N}$, $\delta>0$. Then $\big|T_\delta(\mathcal{P}_{UW})\big| \leq 2^{n(H(U,W)+\delta_2)}$ with $\delta_2 = \log\frac{1}{\min_{(u,w)\in\operatorname{supp}\mathcal{P}_{UW}}\mathcal{P}(u,w)}\cdot\delta$.*

**Lemma 3** *We consider a set $\mathcal{U}$ such that $2 \leq |\mathcal{U}| < +\infty$. For all $\delta > 0$, there exists a family of distributions $(\mathcal{Q}^k_U)_{k\in\mathcal{K}} \subset \operatorname{int}\Delta(\mathcal{U})$ with $|\mathcal{K}| < +\infty$ such that*

$$\Delta(\mathcal{U}) \subset \bigcup_{k\in\mathcal{K}} T_\delta(\mathcal{Q}^k_U), \quad \min_{k\in\mathcal{K}}\min_{u\in\mathcal{U}} \mathcal{Q}^k(u) \geq \frac{\delta}{4(|\mathcal{U}|-1)}.$$

*Proof.* [Lemma 3] We consider a symbols $\tilde{u}\in\mathcal{U}$ and we define the distributions

$$\mathcal{P}_U = \begin{cases} 1 & \text{if } U=\tilde{u},\\ 0 & \text{otherwise,}\end{cases} \quad \mathcal{Q}^{\tilde{u}}_U = \begin{cases} 1-\frac{\delta}{4} & \text{if } U=\tilde{u},\\ \frac{\delta}{4(|\mathcal{U}|-1)} & \text{otherwise.}\end{cases}$$

Then,

$$\begin{aligned}
||\mathcal{Q}^{\tilde{u}}_U - \mathcal{P}_U||_1 =& \sum_u |\mathcal{Q}^{\tilde{u}}(u) - \mathcal{P}(u)|\\
=&\frac{\delta}{4} + \frac{\delta}{4(|\mathcal{U}|-1)}(|\mathcal{U}|-1) = \frac{\delta}{2} < \delta. \qquad (72)
\end{aligned}$$

This shows that $\mathcal{P}_U \in T_\delta(\mathcal{Q}^{\tilde{u}}_U)$. The same construction applies to any other symbol $\hat{u}\in\mathcal{U}$, and this generates a collection of distributions $(\mathcal{Q}^{\hat{u}}_U)_{\hat{u}\in\mathcal{U}}$. We construct a family of distributions $(\mathcal{Q}^k_U)_{k\in\mathcal{K}} \subset \operatorname{int}\Delta(\mathcal{U})$ based on the lattice with steps $\frac{\delta}{4(|\mathcal{U}|-1)}$ that connects the elements of $(\mathcal{Q}^{\hat{u}}_U)_{\hat{u}\in\mathcal{U}}$. Since $\Delta(\mathcal{U}) \subset [0,1]^{|\mathcal{U}|-1}$, we have $|\mathcal{K}| \leq \Big(\frac{4(|\mathcal{U}|-1)}{\delta}\Big)^{|\mathcal{U}|-1} < +\infty$. $\square$

## REFERENCES


[1] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, vol. 50, no. 6, pp. 1431–1451, 1982.

[2] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, pp. 2590–2615, 2011.

[3] M. Le Treust and T. Tomala, "Persuasion with limited communication capacity," *Journal of Economic Theory*, vol. 184, p. 104940, 2019.

[4] S. Dughmi, R. Niazadeh, A. Psomas, and S. M. Weinberg, "Persuasion and incentives through the lens of duality," in *International Conference on Web and Internet Economics*, pp. 142–155, Springer, 2019.

[5] S. Sarıtaş, S. Yüksel, and S. Gezici, "Dynamic signaling games with quadratic criteria under nash and stackelberg equilibria," *Automatica*, vol. 115, p. 108883, 2020.

[6] E. Akyol, C. Langbort, and T. Başar, "Information-theoretic approach to strategic communication as a hierarchical game," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 205–218, 2017.

[7] M. Le Treust and T. Tomala, "Information design for strategic coordination of autonomous devices with non-aligned utilities," *IEEE Proc. of the 54th Allerton conference, Monticello, Illinois*, pp. 233–242, 2016.

[8] M. Le Treust and T. Tomala, "Strategic coordination with state information at the decoder," *Proc. of 2018 International Zurich Seminar on Information and Communication*, 2018.

[9] M. Le Treust and T. Tomala, "Strategic communication with side information at the decoder," *[on-line] available: https://arxiv.org/abs/1911.04950*, Nov. 2019.

[10] A. Lapidoth, "On the role of mismatch in rate distortion theory," *IEEE Transactions on Information Theory*, vol. 43, pp. 38–47, Jan. 1997.

[11] J. Scarlett, A. G. 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.

[12] A. S. Vora and A. A. Kulkarni, "Achievable rates for strategic communication," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1379–1384, 2020.

[13] A. S. Vora and A. A. Kulkarni, "Information extraction from a strategic sender: The zero error case," *[on-line] available: https://arxiv.org/abs/2006.10641*, 2020.

[14] M. O. Jackson and H. F. Sonnenschein, "Overcoming incentive constraints by linking decisions," *Econometrica*, vol. 75, pp. 241 – 257, January 2007.

[15] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, Dec. 2011.