Strategic Communication via Cascade Multiple-Description Network

Rony Bou Rouphael, Maël Le Treust

Strategic Communication via Cascade Multiple-Description Network

Rony Bou Rouphael and Maël Le Treust

ETIS UMR 8051, CY Cergy-Paris Université, ENSEA, CNRS,
6, avenue du Ponceau, 95014 Cergy-Pontoise CEDEX, FRANCE
Email: {rony.bou-rouphael ; mael.le-treust}@ensea.fr

***Abstract*—Strategic information compression and transmission is under study. We consider a three-node cascade network with an encoder, a relay and a decoder, endowed with distinct objectives captured by cost functions. In such a cascade network, agents choose their respective strategies sequentially, as a response to the former agent's strategy and in a way to influence the decision of the latter agent in the network. We assume the encoder commits to a strategy before the communication takes place. Upon revelation of the encoding strategy, the relay commits to a strategy and reveals it. The communication starts, the source sequence is drawn and processed by the encoder and relay. Then, the decoder observes a sequences of symbols, updates its Bayesian posterior beliefs accordingly, and takes the optimal action. This is an extension of the Bayesian persuasion problem in the Game Theory literature. In this work, we provide an information-theoretic approach to study the fundamental limit of the strategic communication via three-node cascade network. Our goal is to characterize the optimal strategies of the encoder, the relay and the decoder, and study the asymptotic behavior of the encoder's minimal long-run cost function.**

I. Introduction

We study a decentralized decision-making problem with restricted communication between three agents with non-aligned objectives. As depicted in Fig. 1, we consider a Cascade channel where information travels from a strategic encoder to a decoder through a strategic relay. We are interested in designing an achievable multiple description coding scheme that minimizes the encoder's long run cost function subject to the challenges imposed by the Cascade channel.

The problem of strategic communication originally emerged in the game theory literature to address situations in economics (lobbying, advertising, sales, negotiations, etc.). The game was referred to as the sender-receiver game, and communication was assumed to be perfect and unconstrained by any limits on the amount of information transmitted. The Nash equilibrium solution of the cheap talk game was investigated by Crawford and Sobel in their seminal paper [1], in which the encoder and the decoder are endowed with distinct objectives and choose their coding strategies simultaneously. The Stackelberg version of the strategic communication game, referred to as

Maël Le Treust gratefully acknowledges financial support from INS2I CNRS, DIM-RFSI, SRV ENSEA, UFR-ST UCP, INEX Paris Seine Initiative and IEA Cergy-Pontoise. This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01).

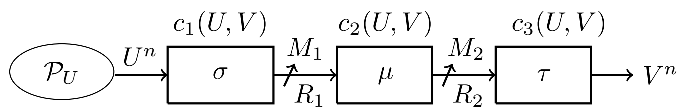

Fig. 1: Strategic Source Coding for Cascade Channel with Successive Commitment.

the Bayesian persuasion game, was formulated by Kamenica and Gentzkow in [2], where the encoder is the Stackelberg leader and the decoder is the Stackelberg follower choosing its strategies as a response to the encoder's strategy . In this paper, we assume that the encoder commits to an encoding and announces its commitment before observing the source. Then, the relay commits to and announces a strategy accordingly. If the relay was assumed to commit to a strategy before the encoder, then the problem boils down to a strategic joint source-channel coding of Shannon, like the one investigated in [3]. Cascade source coding consists of compressing a source sequence through an intermediate or relay node which then reconstructs the source and transmits it to the next node. In [4], Yamamoto considered the source coding problem for cascade and branching communication systems, and established the region of achievable rates for cascade systems and bounds for the branching systems. Lossy source coding for cascade communication systems was also considered in [5] where both the relay and the terminal node have access to side information and wish to reconstruct the source with certain fidelities.

In Bayesian persuasion, the encoder is considered to be an information designer. In [6], information design with multiple designers interacting with a set of agents is studied. In [7], [8], the Nash equilibrium solution is investigated for multi-dimensional sources and quadratic cost functions, whereas the Stackelberg solution is studied in [9]. The computational aspects of the persuasion game are considered in [10]. In several recent contributions [11], [12], [13], Vora and Kulkarni addressed the problem of extracting truthful information from a strategic sender with an incentive to misreport information. Modeled as a Stackelberg game in which the decoder is the Stackelberg leader, authors investigate the region of achievable rates of strategic communication between agents with distinct utility functions. The strategic communication problem with a noisy channel is investigated in [14], [15], [16], [3], and [17]. The case where the decoder privately observes a signal correlated to the state, also referred to as the Wyner-Ziv setting [18], is studied in [19], [20] and [21].

In this paper, we study the Bayesian persuasion game via a Cascade multiple description network. The objectives of the players are captures by distinct cost functions that depend on the source and the action taken by the decoder. For some particular cases, we are able to characterize the encoder's optimal cost obtained with strategic cascade multiple description coding that satisfies the decoder's incentives constraints.

A. Notations

Let $n \in \mathbb{N}^\star = \mathbb{N}\backslash\{0\}$ denote a sequence block length. We denote by U^n the n-sequences of random variables of source information $u^n = (u_1, ..., u_n) \in \mathcal{U}^n$, and by V^n the sequences of decoders' actions $v^n \in \mathcal{V}^n$. Sequences M_1 and M_2 denote the channel inputs of encoders 1 and 2 respectively. Calligraphic fonts $\mathcal{U}$, and $\mathcal{V}$ denote the alphabets and lowercase letters u and v denote the realizations. For a discrete random variable X, we denote by $\Delta(\mathcal{X})$ the probability simplex, i.e. the set of probability distributions over $\mathcal{X}$, and by $\mathcal{P}_X(x)$ the probability mass function $\mathbb{P}\{X = x\}$. Notation $X -\!\!\circ\!\!- Y -\!\!\circ\!\!- Z$ stands for the Markov chain property $\mathcal{P}_{Z|XY} = \mathcal{P}_{Z|Y}$.

II. System Model

In this section, we formulate the coding problem. Let $n \in \mathbb{N}^\star$, and $(R_1, R_2) \in \mathbb{R}_+^2$ denote the rate pair. We assume that the information source U follows the independent and identically distributed (i.i.d) probability distribution $\mathcal{P}_U \in \Delta(\mathcal{U})$.

Definition 1. *The coding strategies σ, μ and τ of the encoder, relay, and decoder respectively are defined by*

$$\sigma : \mathcal{U}^n \longrightarrow \Delta\Big(\{1, ..2^{\lfloor nR_1 \rfloor}\}\Big), \tag{1}$$

$$\mu : \{1, ..2^{\lfloor nR_1 \rfloor}\} \longrightarrow \Delta\Big(\{1, ..2^{\lfloor nR_2 \rfloor}\}\Big), \tag{2}$$

$$\tau : \{1, ..2^{\lfloor nR_2 \rfloor}\} \longrightarrow \Delta\Big(\mathcal{V}^n\Big). \tag{3}$$

The stochastic coding strategies (σ, μ, τ) induce a joint probability distribution $\mathcal{P}^{\sigma\mu\tau} \in \Delta\big(\mathcal{U}^n \times \{1, 2, ..2^{\lfloor nR_1 \rfloor}\} \times \{1, 2, ..2^{\lfloor nR_2 \rfloor}\} \times \mathcal{V}^n\big)$ defined for all (u^n, m_1, m_2, v^n) by

$$\mathcal{P}^{\sigma\mu\tau}(u^n, m_1, m_2, v^n) = \left(\prod_{t=1}^{n} \mathcal{P}_U(u_t)\right)\sigma(m_1|u^n)\mu(m_2|m_1)\tau(v^n|m_2). \tag{4}$$

Definition 2. *We consider arbitrary single-letter cost functions $c_1 : \mathcal{U} \times \mathcal{V} \longrightarrow \mathbb{R}$ for the encoder $\mathcal{E}$, $c_2 : \mathcal{U} \times \mathcal{V} \longrightarrow \mathbb{R}$ for the relay, and $c_3 : \mathcal{U} \times \mathcal{V} \longrightarrow \mathbb{R}$ for the decoder. The long-run cost functions are defined by*

$$\begin{aligned} c_1^n(\sigma, \mu, \tau) =& \mathbb{E}_{\sigma,\mu,\tau}\left[\frac{1}{n}\sum_{t=1}^{n} c_1(U_t, V_t)\right] \\ =& \sum_{u^n, v^n} \mathcal{P}_{U^nV^n}^{\sigma,\mu,\tau}(u^n, v^n) \cdot \left[\frac{1}{n}\sum_{t=1}^{n} c_1(u_t, v_t)\right], \\ c_2^n(\sigma, \mu, \tau) =& \mathbb{E}_{\sigma,\mu,\tau}\left[\frac{1}{n}\sum_{t=1}^{n} c_2(U_t, V_t)\right] \\ =& \sum_{u^n, v^n} \mathcal{P}_{U^nV^n}^{\sigma,\mu,\tau}(u^n, v^n) \cdot \left[\frac{1}{n}\sum_{t=1}^{n} c_2(u_t, v_t)\right], \\ c_3^n(\sigma, \mu, \tau) =& \mathbb{E}_{\sigma,\mu,\tau}\left[\frac{1}{n}\sum_{t=1}^{n} c_3(U_t, V_t)\right] \\ =& \sum_{u^n, v^n} \mathcal{P}_{U^nV^n}^{\sigma,\mu,\tau}(u^n, v^n) \cdot \left[\frac{1}{n}\sum_{t=1}^{n} c_3(u_t, v_t)\right], \end{aligned}$$

In the above equations, $\mathcal{P}_{U^nV^n}^{\sigma\mu\tau}$ denote the marginal distributions over the sequences (U^n, V^n) of $\mathcal{P}^{\sigma\mu\tau}$ defined in (4) over the n-sequences (U^n, M_1, M_2, V^n).

Definition 3. *For any strategy pair (σ, μ), the set of best-response strategies τ of the decoder is defined by*

$$\mathbb{A}_3(\sigma, \mu) = \arg\min_{\tau} c_3^n(\sigma, \mu, \tau). \tag{5}$$

For any strategy σ, the set of best-response strategies μ of the relay is defined by

$$\mathbb{A}_2(\sigma) = \underset{\substack{(\mu,\tau) s.t. \\ \tau \in \mathbb{A}_3(\sigma,\mu)}}{\arg\min}\, c_2^n(\sigma, \mu, \tau). \tag{6}$$

Therefore, the encoder has to solve the following coding problem,

$$\Gamma_e^n(R_1, R_2) = \inf_{\sigma} \max_{\substack{(\mu,\tau) \in \\ \mathbb{A}_2(\sigma)}} c_1^n(\sigma, \mu, \tau). \tag{7}$$

Remark 1. *In order to get a robust solution concept, we assume that the encoder solves the problem for the worst case scenario, i.e. if more than one pair of strategies are available in $\mathbb{A}_2(\sigma)$, we consider the one that maximizes the encoder's cost.*

The operational significance of (7) corresponds to the persuasion game that is played in the following steps:

- The encoder chooses, announces the encoding σ.
- knowing σ, the relay chooses, announces the encoding μ.
- Knowing (σ, μ), the decoder compute its best-response strategy τ.
- Sequences U^n are drawn i.i.d with distribution $\mathcal{P}_U$.
- Message sequence M_1 are encoded according to $\sigma_{M_1|U^n}$.
- Message sequence M_2 are encoded according to $\mu_{M_2|M_1}$.
- The decoder observes M_2 and draws V^n according to $\tau_{V^n|M_2}$.
- Cost functions $c_1^n(\sigma, \mu, \tau)$, $c_2^n(\sigma, \mu, \tau)$, $c_3^n(\sigma, \mu, \tau)$ are computed.

III. Cascade Multiple Description Coding

Consider the cooperative communication scenario where $c_1 = c_2 = c_3$, and all three agents share the objective of minimizing the same cost function. This setting corresponds to the standard coding setup of a cascade multiple description network [22], under the assumption that the relay does not reconstruct the source, but only relays a message M_2, and the cost functions of the three players depend on the source and the decoder's action.

Consider an auxiliary random variables $W \in \mathcal{W}$ such that $|\mathcal{W}| = |\mathcal{U}|$. The set $\mathbb{Q}_0^c(R_1, R_2)$ of target distributions is defined by:

$$\mathbb{Q}_0^c(R_1,R_2) =\{\mathcal{Q}_{W_2|U};\quad \min(R_1,R_2) \geq I(U;W_2)\}. \tag{8}$$

The single-letter best-response of the decoder is defined by:

$$\mathbb{Q}_3^c(\mathcal{Q}_{W_2|U}) = \underset{\mathcal{Q}_{V|W_2}}{\arg\min}\, \mathbb{E}[c_3(U,V)]. \tag{9}$$

The single-letter optimal cost $\Gamma_e^c(R_1,R_2)$ of the encoder is given by

$$\Gamma_e^c(R_1,R_2) = \inf_{\substack{\mathcal{Q}_{W_2|U}\in\\ \mathbb{Q}_0^c(R_1,R_2)}} \max_{\substack{\mathcal{Q}_{V|W_2}\in\\ \mathbb{Q}_3^c(\mathcal{Q}_{W_2|U})}} \mathbb{E}\Big[c_1(U,V)\Big]. \tag{10}$$

Theorem 1. *Let $(R_1,R_2) \in \mathbb{R}_+^2$. If $c_1 = c_2 = c_3$, then*

$$\lim_{n\longrightarrow\infty} \Gamma_e^n(R_1,R_2) = \inf_{n\in\mathbb{N}^\star} \Gamma_e^n(R_1,R_2) = \Gamma_e^c(R_1,R_2). \tag{11}$$

The proof of Theorem 1 can be directly derived from the proof of [22, Theorem 20.4] by considering the relay's estimate to be a constant and its role is to only transition the message received from the encoder.

IV. Bayesian Persuasion with no Information Constraint

We assume that the communication is perfect and unrestricted. Fix $\mathcal{Q}_{W_1|U}\mathcal{Q}_{W_2|W_1}$. Consider two auxiliary random variables $W_1 \in \mathcal{W}_1$ and $W_2 \in \mathcal{W}_2$ such that $|\mathcal{W}_1| = |\mathcal{W}_2| = |\mathcal{U}|$ and

$$U -\!\!\circ\!\!- W_1 -\!\!\circ\!\!- W_2, \quad W_1 -\!\!\circ\!\!- W_2 -\!\!\circ\!\!- V.$$

The single-letter best-responses are defined by:

$$\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1}) = \underset{\mathcal{Q}_{V|W_2}}{\arg\min}\, \mathbb{E}[c_3(U,V)],$$
$$\mathbb{Q}_2(\mathcal{Q}_{W_1|U}) = \underset{\substack{(\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}),\\ \mathcal{Q}_{V|W_2}\in\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1})}}{\arg\min} \mathbb{E}[c_2(U,V)],$$

The single-letter optimal cost Γ_e of the encoder is given by

$$\Gamma_e = \inf_{\mathcal{Q}_{W_1|U}} \max_{\substack{\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}\\ \in\mathbb{Q}_2(\mathcal{Q}_{W_1|U})}} \mathbb{E}_{\substack{\mathcal{P}_U\mathcal{Q}_{W_1|U}\\ \mathcal{Q}_{W_2|W_1}\mathcal{Q}_{V|W_2}}} \Big[c_1(U,V)\Big].$$

Theorem 2. *If $R_1 = R_2 = \log|\mathcal{U}|$, then*

$$\lim_{n\longrightarrow\infty} \Gamma_e^n = \inf_{n\in\mathbb{N}^\star} \Gamma_e^n = \Gamma_e \tag{12}$$

A. Achievability of Theorem 1

Let $R_1 = R_2 = \log|\mathcal{U}|$, and fix a joint probability distribution $\mathcal{Q}_{W_1|U}\mathcal{Q}_{W_2|W_1}$. The sequences U^n are drawn according to the i.i.d. distribution $\mathcal{P}_{U^n}$. Randomly and independently generate 2^{nR_1} sequences $w_1^n(m_1)$ for each $m_1 \in \{1,..2^{\lfloor nR_1\rfloor}\}$, according to the i.i.d distribution $\mathcal{Q}_{W_1^n|U^n} = \Pi_{t=1}^n \mathcal{Q}_{W_1|U}(w_{1t}|u_t)$. Similarly, generate 2^{nR_2} sequences $w_2^n(m_2)$ for $m_2 \in \{1,..2^{\lfloor nR_2\rfloor}\}$ randomly and independently according to the i.i.d distribution $\mathcal{Q}_{W_2^n|W_1^n} = \Pi_{t=1}^n \mathcal{Q}_{W_2|W_1}(w_{2t}|w_{1t})$.

Since $R_1 = \log|\mathcal{U}| = \log|\mathcal{W}_1|$ and $R_2 = \log|\mathcal{U}| = \log|\mathcal{W}_2|$, encoder $\mathcal{E}$ observes u^n and looks in the codebook for the corresponding sequences $w_1^n(m_1)$ and sends m_1 to the relay. The relay observes m_1 and sends m_2 to the decoder. Then, the decoder $\mathcal{D}$ observes m_2 and declares v^n according to τ.

B. Converse Proof

Given a triple (σ,μ,τ) and a random variable T uniformly distributed over $\{1,2,...,n\}$ and independent of (U^n, M_1, M_2, V^n). We identify the auxiliary random variables $W_1 = (M_1,T)$, $W_2 = M_2$, $(U,V) = (U_T,V_T)$, distributed according to $\mathcal{P}^{\sigma\mu\tau}_{UW_1W_2V}$ defined for all $(u,w_1,w_2,v) = (u_t,x_1,x_2,t,v_t)$ by

$$\mathcal{P}^{\sigma\mu\tau}_{UW_1W_2V}(u,w_1,w_2,v) = \mathcal{P}^{\sigma\mu\tau}_{U_TW_1W_2TV_T}(u_t,x_1,x_2,t,v_t)$$
$$=\frac{1}{n}\sum_{\substack{u^{t-1}\\ u^n_{t+1}}}\ \sum_{\substack{x_1^{t-1},x^n_{1,t+1}\\ x_2^{t-1},x^n_{2,t+1}}}\ \sum_{v^{t-1},v^n_{t+1}}\bigg(\prod_{t=1}^n \mathcal{P}_U(u_t)\bigg)\mathcal{P}^\sigma_{M_1|U^n}(m_1|u^n)$$
$$\times\, \mathcal{P}^\mu_{M_2|M_1}(m_2|m_1)\mathcal{P}^\tau_{V^n|M_2}(v^n|m_2).$$

Lemma 1. *The distribution $\mathcal{P}^{\sigma\mu\tau}_{UW_1W_2V}$ has marginal on $\Delta(\mathcal{U})$ given by $\mathcal{P}_U$ and satisfies the following Markov chain property*

$$U -\!\!\circ\!\!- W_1 -\!\!\circ\!\!- W_2, \quad W_1 -\!\!\circ\!\!- W_2 -\!\!\circ\!\!- V.$$

Proof. [Lemma 1] The i.i.d. property of the source ensures that the marginal distribution is $\mathcal{P}_U$. By the definition of the coding functions σ, μ and τ we have

$$(U_T) -\!\!\circ\!\!- (M_1,T) -\!\!\circ\!\!- M_2,$$
$$(M_1,T) -\!\!\circ\!\!- M_2 -\!\!\circ\!\!- V_T.$$

$\square$ *Therefore $\mathcal{P}^{\sigma\mu\tau}_{UW_1W_2V} = \mathcal{P}_U\mathcal{P}^\sigma_{W_1|U}\mathcal{P}^\mu_{W_2|W_1}\mathcal{P}^\tau_{V|W_2}$.*

Lemma 2. *For all (σ,τ_1,τ_2) and $i\in\{1,2,3\}$, we have*

$$c_i^n(\sigma,\mu,\tau) = \mathbb{E}\big[c_i(U,V)\big], \tag{13}$$

evaluated with respect to $\mathcal{P}_U\mathcal{P}^\sigma_{W_1|U}\mathcal{P}^\mu_{W_2|W_1}\mathcal{P}^\tau_{V|W_2}$. Moreover for all σ,μ we have

$$\mathbb{Q}_3(\mathcal{P}^\sigma_{W_1|U},\mathcal{P}^\mu_{W_2|W_1}) = \Big\{\mathcal{Q}_{V|W_2},\ \exists\tau\in\mathbb{A}_3(\sigma,\mu),\ \mathcal{Q}_{V|W_2} = \mathcal{P}^\tau_{V|W_2}\Big\}, \tag{14}$$

$$\mathbb{Q}_2(\mathcal{P}^\sigma_{W_1|U}) = \Big\{(\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}),\ \exists(\mu,\tau)\in\mathbb{A}_2(\sigma),\ \mathcal{Q}_{W_2|W_1} = \mathcal{P}^\mu_{W_1|W_2}, \mathcal{Q}_{V|W_2} = \mathcal{P}^\tau_{V|W_2}\Big\}. \tag{15}$$

Proof. [Lemma 2] By Definition 2 we have for $i \in \{1,2,3\}$

$$c_i^n(\sigma,\mu,\tau) = \sum_{\substack{u^n,m_1,\\ m_2,v^n}} \bigg(\prod_{t=1}^n \mathcal{P}_U(u_t)\bigg)\mathcal{P}^{\sigma}_{M_1|U^n}(m_1|u^n)\mathcal{P}^{\mu}_{M_2|M_1}(m_2|m_1) \times \mathcal{P}^{\tau}_{V^n|M_2}(v^n|m_2)\cdot\bigg[\frac{1}{n}\sum_{t=1}^n c_i(u_t,v_t)\bigg] \tag{16}$$

$$=\sum_{t=1}^n \sum_{\substack{u_t,x_1,\\ x_2,t,v_t}} \mathcal{P}^{\sigma,\mu,\tau}(u_t,x_1,x_2,t,v_t)\times c_i(u_t,v_t) = \mathbb{E}\big[c_i(U,V)\big].$$

Given $\mathcal{Q}_{V|W_2} \in \mathbb{Q}_3(\mathcal{P}^{\sigma}_{W_1|U},\mathcal{P}^{\mu}_{W_2|W_1})$, we consider τ such that

$$\mathcal{P}^{\tau}_{V^n|M_2}(v^n|m_2) = \prod_{t=1}^n \mathcal{Q}_{V|W_2}(v_{1,t}|m_2).$$

Given $(\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}) \in \mathbb{Q}_2(\mathcal{P}^{\sigma}_{W_1|U})$, we consider (μ,τ) such that

$$\mathcal{P}^{\mu}_{M_2|M_1}(m_2|m_1) = \prod_{t=1}^n \mathcal{Q}_{W_2|W_1}(m_2|m_1,t),$$
$$\mathcal{P}^{\tau}_{V^n|M_2}(v^n|m_2) = \prod_{t=1}^n \mathcal{Q}_{V|W_2}(v_{1,t}|m_2).$$

Therefore

$$c_3^n(\sigma,\mu,\tau) = \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{P}^{\mu}_{W_2|W_1}\mathcal{Q}_{V|W_2}}}\big[c_3(U,V)\big] \tag{17}$$
$$= \min_{\mathcal{P}_{V|W_2}} \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{P}^{\mu}_{W_2|W_1},\mathcal{P}_{V|W_2}}}\big[c_3(U,V)\big] \tag{18}$$
$$\leq \min_{\tilde{\tau}} \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{P}^{\mu}_{W_2|W_1},\mathcal{P}^{\tilde{\tau}}_{V|W_2}}}\big[c_3(U,V)\big] = \min_{\tilde{\tau}} c_3^n(\sigma,\mu,\tilde{\tau}),$$

hence $\tau \in \mathbb{A}_3(\sigma,\mu)$. Similarly,

$$c_2^n(\sigma,\mu,\tau) = \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{Q}_{W_2|W_1}\mathcal{Q}_{V|W_2}}}\big[c_2(U,V)\big] \tag{19}$$
$$= \min_{(\mathcal{P}_{W_2|W_1},\mathcal{P}_{V|W_2})} \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{P}_{W_2|W_1},\mathcal{P}_{V|W_2}}}\big[c_2(U,V)\big] \tag{20}$$
$$\leq \min_{(\tilde{\mu},\tilde{\tau})} \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{P}^{\tilde{\mu}}_{W_2|W_1},\mathcal{P}^{\tilde{\tau}}_{V|W_2}}}\big[c_2(U,V)\big] = \min_{(\tilde{\mu},\tilde{\tau})} c_2^n(\sigma,\tilde{\mu},\tilde{\tau}), \tag{21}$$

and thus $(\mu,\tau) \in \mathbb{A}_2(\sigma)$. The other inclusions are direct and the same arguments imply (15) *and* (14). □

For any strategy σ, we have

$$\max_{\mu,\tau} c_1^n(\sigma,\mu,\tau) = \max_{\mu,\tau} \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{P}^{\mu}_{W_2|W_1}\mathcal{P}^{\tau}_{V|W_2}}}\Big[c_1(U,V)\Big] \tag{22}$$
$$\geq \max_{\substack{\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}\\ \in\mathbb{Q}_2(\mathcal{Q}_{W_1|U})}} \mathbb{E}_{\substack{\mathcal{P}^{\sigma}_{W_1|U}\\ \mathcal{Q}_{W_2|W_1}\mathcal{Q}_{V|W_2}}}\Big[c_1(U,V)\Big] \tag{23}$$
$$\geq \inf_{\mathcal{Q}_{W_1|U}} \max_{\substack{\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}\\ \in\mathbb{Q}_2(\mathcal{Q}_{W_1|U})}} \mathbb{E}\Big[c_1(U,V)\Big] \tag{24}$$
$$= \Gamma_e(R_1,R_2). \tag{25}$$

Equations (22) *and* (23) *comes from Lemma 2, whereas* (24) *comes from taking the infimun over $\mathcal{Q}_{W_1|U}$. This concludes the converse proof of Theorem 2.*

V. Locally Restricted Communication

A. Relay's Restriction

Assume that the encoder can send messages at large enough rate $R_1 = \log|\mathcal{U}|$, but the relay sends at a fixed smaller rate R_2. Fix $\mathcal{Q}_{W_1|U}$. In this setting, the single-letter best-responses are defined by:

$$\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1}) = \arg\min_{\mathcal{Q}_{V|W_2}} \mathbb{E}[c_3(U,V)],$$
$$\mathbb{Q}_2^r(\mathcal{Q}_{W_1|U}) = \arg\min_{\substack{(\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}))s.t. R_2\geq I(W_1;W_2),\\ \mathcal{Q}_{V|W_2}\in\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1})}} \mathbb{E}[c_2(U,V)],$$

The single-letter optimal cost $\Gamma_e^r(R_2)$ of the encoder is given by

$$\Gamma_e^r(R_2) = \inf_{\mathcal{Q}_{W_1|U}} \max_{\substack{\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}\\ \in\mathbb{Q}_2(\mathcal{Q}_{W_1|U})}} \mathbb{E}\Big[c_1(U,V)\Big].$$

Theorem 3. *Let $R_2 \in \mathbb{R}^+$. If $R_1 = \log|\mathcal{U}|$, then*

$$\lim_{n\longrightarrow\infty} \Gamma_e^n(R_2) = \inf_{n\in\mathbb{N}^\star} \Gamma_e^n(R_2) = \Gamma_e^r(R_2). \tag{26}$$

The proof of Theorem 3 relies on the lossy source coding at the relay by considering the source to be the observed message which is uniformly drawn from the codebook of size 2^{nR_1}. This slight modification does not affect the condition on the covering lemma as the coding will only depend on the size 2^{nR_2} of the message set of the relay.

B. Encoder's Restriction

Now assume that $R_2 = \log|\mathcal{U}|$, i.e. the encoder is restricted to a limited amount of bits per transmission, but the relay can transmit with no information constraints. Therefore, the set $\mathbb{Q}_0^e(R_1)$ of the encoder's target distributions is given by

$$\mathbb{Q}_0^e(R_1) = \{\mathcal{Q}_{W_1|U};\quad R_1 \geq I(U;W_1)\}.$$

Single-letter best-responses are defined by:

$$\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1}) = \arg\min_{\mathcal{Q}_{V|W_2}} \mathbb{E}[c_3(U,V)],$$
$$\mathbb{Q}_2^e(\mathcal{Q}_{W_1|U}) = \arg\min_{\substack{(\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2})),\\ \mathcal{Q}_{V|W_2}\in\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1})}} \mathbb{E}[c_2(U,V)],$$

The single-letter optimal cost $\Gamma_e^\star(R_1)$ of the encoder is given by

$$\tilde{\Gamma}_e(R_1) = \inf_{\mathcal{Q}_{W_1|U}\in\mathbb{Q}_0^e(R_1)} \max_{\substack{\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}\\ \in\mathbb{Q}_2^e(\mathcal{Q}_{W_1|U})}} \mathbb{E}\Big[c_1(U,V)\Big].$$

Theorem 4. *Let $R_1 \in \mathbb{R}^+$. If $R_2 = \log|\mathcal{U}|$, then*

$$\lim_{n\longrightarrow\infty} \Gamma_e^n(R_2) = \inf_{n\in\mathbb{N}^\star} \Gamma_e^n(R_2) = \tilde{\Gamma}_e(R_2).$$

VI. Locally Cooperating Agents

Consider now that either the relay and the encoder or the relay and the decoder are cooperating. In other words, we assume that either $c_1 = c_2$ or $c_2 = c_3$ holds.

A. Encoder-Relay Cooperation

Assume $c_1 = c_2$ the encoder and the relay are cooperating. The encoder will reveal information using the maximal rate R_1. The set $\mathbb{Q}_0^s(R_1, R_2)$ of target distributions:

$$\mathbb{Q}_0^s(R_1,R_2) = \{\mathcal{Q}_{W_1|U};\quad R_1 \geq I(U;W_1)\}.$$

Single-letter best-responses are defined by:

$$\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1}) = \operatorname*{arg\,min}_{\mathcal{Q}_{V|W_2}} \mathbb{E}[c_3(U,V)],$$
$$\mathbb{Q}_2^s(\mathcal{Q}_{W_1|U}) = \operatorname*{arg\,min}_{\substack{(\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2})),R_2\geq I(W_1;W_2)\\ \mathcal{Q}_{V|W_2}\in\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1})}} \mathbb{E}[c_2(U,V)],$$

The single-letter optimal cost $\Gamma_e^s(R_1,R_2)$ of the encoder is given by

$$\Gamma_e^s(R_1,R_2) = \inf_{\mathcal{Q}_{W_1|U}\in\mathbb{Q}_0^s(R_1,R_2)} \max_{\substack{\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}\\ \in\mathbb{Q}_2^s(\mathcal{Q}_{W_1|U})}} \mathbb{E}\Big[c_1(U,V)\Big].$$

Theorem 5. *Let $(R_1,R_2)\in\mathbb{R}_+^2$. If $c_1 = c_2$, then*

$$\lim_{n\longrightarrow\infty}\Gamma_e^n(R_1,R_2) = \inf_{n\in\mathbb{N}^\star}\Gamma_e^n(R_1,R_2) = \Gamma_e^s(R_1,R_2).$$

The proof relies on considering that the relay observes the source as the encoder can fully reveal it, and the Bayesian persuasion setting between the relay and the decoder.

B. Relay-Decoder Cooperation

Assume now that the decoder cooperates with the relay because $c_2 = c_3$. The set $\mathbb{Q}_0^d(R_1,R_2)$ of target distributions:

$$\mathbb{Q}_0^d(R_1,R_2) = \{\mathcal{Q}_{W_1|U};\quad R_1 \geq I(U;W_1)\}.$$

Single-letter best-responses are defined by:

$$\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1}) = \operatorname*{arg\,min}_{\mathcal{Q}_{V|W_2}} \mathbb{E}[c_3(U,V)],$$
$$\mathbb{Q}_2^d(\mathcal{Q}_{W_1|U}) = \operatorname*{arg\,min}_{\substack{(\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2})),R_2\geq I(W_1;W_2)\\ \mathcal{Q}_{V|W_2}\in\mathbb{Q}_3(\mathcal{Q}_{W_1|U},\mathcal{Q}_{W_2|W_1})}} \mathbb{E}[c_2(U,V)].$$

The single-letter optimal cost $\Gamma_e^d(R_1,R_2)$ of the encoder is given by

$$\Gamma_e^d(R_1,R_2) = \inf_{\mathcal{Q}_{W_1|U}\in\mathbb{Q}_0^d(R_1,R_2)} \max_{\substack{\mathcal{Q}_{W_2|W_1},\mathcal{Q}_{V|W_2}\\ \in\mathbb{Q}_2^d(\mathcal{Q}_{W_1|U})}} \mathbb{E}\Big[c_1(U,V)\Big].$$

Theorem 6.

$$\lim_{n\longrightarrow\infty}\Gamma_e^n(R_1,R_2) = \inf_{n\in\mathbb{N}^\star}\Gamma_e^n(R_1,R_2) = \Gamma_e^d(R_1,R_2).$$

The proof follows by considering the relay and the decoder as one party, and lossy source coding at the encoder.

VII. Binary Example

Assume that $R_1 = R_2 = \log|\mathcal{U}|$, and $c_1 = c_3$. We illustrate the problem using a binary source information $\mathcal{U} = \{u_0,u_1\}$, binary channel inputs $\mathcal{X}_2 = \{x_0^2, x_1^2\}$, $\mathcal{X}_1 = \{x_0^1,x_1^1\}$ and binary action set $\mathcal{V} = \{v_0,v_1\}$. The prior belief $\mathcal{P}_U(u_1)$ is given by the parameter $p_0\in[0,1]$. Single-letter cost functions are given in the tables below.

TABLE I: $c_1(u,v)$

	v_0	v_1
u_0	9	0
u_1	4	10

TABLE II: $c_2(u,v)$

	v_0	v_1
u_0	1	0
u_1	1	0

TABLE III: $c_3(u,v)$

	v_0	v_1
u_0	9	0
u_1	4	10

Let $(\alpha,\beta),(\gamma,\delta)$,and $(\epsilon,\eta)\in[0,1]^2$.

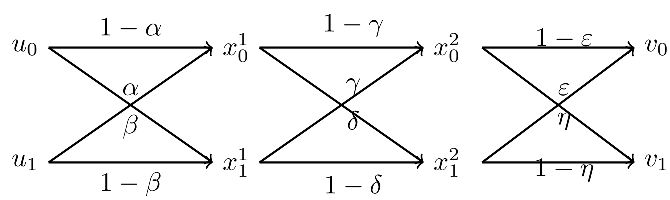

Fig. 2: Encoders' Joint Strategies σ_1 and σ_2 and decoder's strategy σ_3.

Using Baye's rule, we compute the following

$$q_0^1 = \mathcal{P}(u_1|x_0^2) = \frac{\mathcal{P}(u_1,x_0^2)}{\mathcal{P}(x_0^2)} = \frac{(\beta(1-\gamma)+(1-\beta)\delta)\cdot p_0}{(\beta(1-\gamma)+(1-\beta)\delta)\cdot p_0 + ((1-\alpha)(1-\gamma)+\alpha\delta)\cdot(1-p_0)},$$

Similarly, one can compute $q_1^1 = \mathcal{P}(u_1|x_1^2), q_0^2 = \mathcal{P}(x_1^1|x_0^2)$ and $q_1^2 = \mathcal{P}(x_1^1|x_1^2)$ can be computed. Let $p_1(\alpha,\beta)\in[0,1]$ denote the belief parameter of the decoder about X^1. In other words, $p_1(\alpha,\beta) = \mathcal{P}_{X^1}(x_1^1) = \sum_u \mathcal{P}_U(u)\cdot\mathcal{P}_{X^1|U}(x_1^1|u) = (1-p_0)\alpha + p_0(1-\beta)$. Beliefs q_0^2 and q_1^2 can be reformulated as follows:

$$q_0^2 = \frac{p_1(\alpha,\beta)\cdot\delta}{p_1(\alpha,\beta)\cdot\delta + (1-p_1(\alpha,\beta))(1-\gamma)},$$
$$q_1^2 = \frac{p_1(\alpha,\beta)\cdot(1-\delta)}{p_1(\alpha,\beta)\cdot(1-\delta)+(1-p_1(\alpha,\beta))\gamma}.$$

Knowing (α,β), the relay will tune $(\gamma^\star(\alpha,\beta),\delta^\star(\alpha,\beta))$ so that posteriors (q_0^2,q_1^2) are an optimal splitting as follows:

$$\gamma^\star(\alpha,\beta) = \frac{(1-q_1^2)(p_1(\alpha,\beta)-q_0^2)}{(1-p_1(\alpha,\beta))(q_1^2-q_0^2)},$$
$$\delta^\star(\alpha,\beta) = \frac{q_0^2(q_1^2-p_1(\alpha,\beta))}{p_1(\alpha,\beta)\cdot(q_1^2-q_0^2)}.$$

Remark 2. *The order of commitment is crucial in this setting. If the relay commits to a strategy (γ,δ) and announces it before the encoder commits to and announces a strategy, thus the problem boils down to the one tackled in [3].*

Remark 3. *If $\alpha+\beta=1$, then the source U and channel's input X^1 are independent. In that case, the decoder will stick to its prior belief p_0 disregarding any information received from the relay, and play its default action v_1. The corresponding costs are $10\times 0.4=4$ for the encoder, and 1 for the relay.*

Using the convex closure of the decoder's expected cost, we aim to find the optimal splitting (q_0^2, q_1^2) for the relay. For that, we need to define the costs $c_i^x(x^1,v), i\in\{1,2,3\}$ of all players as functions of the channel input X_1 and the decoder's action V as follows

$$c_1^x(x^1,v)=\sum_u \mathcal{P}_{U|X^1}(u|x^1)c_1(u,v),\ \forall x^1,v.$$
$$c_2^x(x^1,v)=\sum_u \mathcal{P}_{U|X^1}(u|x^1)c_2(u,v),\ \forall x^1,v.$$
$$c_3^x(x^1,v)=\sum_u \mathcal{P}_{U|X^1}(u|x^1)c_3(u,v).\ \forall x^1,v.$$

The distributions $\mathcal{P}_{U|X^1}(u_0|x_0^1)$ and $\mathcal{P}_{U|X^1}(u_1|x_1^1)$ computed as follows

$$\mathcal{P}(u_0|x_0^1)=\frac{\mathcal{P}(u_0,x_0^1)}{\mathcal{P}(x_0^1)}=\frac{(1-\alpha)\cdot(1-p_0)}{\beta\cdot p_0+(1-\alpha)\cdot(1-p_0)},$$
$$\mathcal{P}(u_1|x_1^1)=\frac{\mathcal{P}(u_1,x_1^1)}{\mathcal{P}(x_1^1)}=\frac{(1-\beta)\cdot p_0}{(1-\beta)\cdot p_0+\alpha\cdot(1-p_0)}.$$

The threshold $g(\alpha,\beta)$ at which the decoder changes action is computed as follows

$$g(\alpha,\beta)=\frac{2-\mathcal{P}(u_0|x_0^1)\cdot 5}{5\cdot(1-\mathcal{P}(u_1|x_1^1)-\mathcal{P}(u_0|x_0^1))}$$

We define the single-letter cost of the encoder and the relay as a function of the belief parameter $q\in[0,1]$ about X^1 and for threshold g as follows:

$$c_1^x(q)=\sum_{x^1} q(x^1)c_1^x(x^1,v^\star(q(x^1))),$$
$$c_2^x(q)=\sum_{x^1} q(x^1)c_2^x(x^1,v^\star(q(x^1))),$$

where

$$v^\star(q(x^1))=\arg\min_v \sum_{x^1} q(x^1)c_3^x(x^1,v).$$

For a given (α,β), the optimal cost of the relay can be computed using the convexification method as follows:

$$C_2^\star(\alpha,\beta)=\inf_{(\lambda,q)_{x^2}}\Big\{\sum_{x^2}\lambda_{x^2}c_2^x(q_{x^2}),$$
$$\sum_k \lambda_k^2=1,\ \sum_{x^2}\lambda_{x^2}q_{x^2}=p_1(\alpha,\beta)\Big\}.$$

The optimal single-letter cost of the encoder is therefore given by

$$\Gamma_e^\star(R_1,R_2)=\inf_{\alpha,\beta}\Big\{\sum_{x^2}\lambda_{x^2}c_1^x(q_{x^2}),$$
$$(\lambda_{x^2},q_{x^2})_{x^2}\in\arg\min_{(\lambda_{x^2},q_{x^2})_{x^2}} C_2^\star(\alpha,\beta)\Big\}.$$

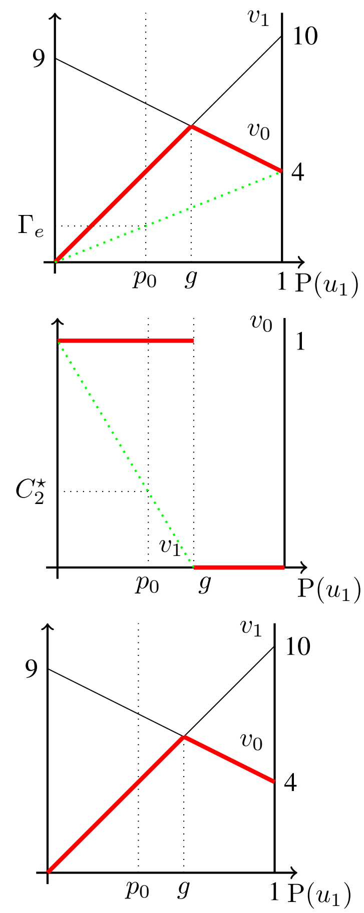

Fig. 3: Expected cost functions with $p_0=0.4$, $g=0.6$, $C_2^\star=0.33$ and $\Gamma_e=1.6$ for large enough rates $R_1,R_2\geq\log|\mathcal{U}|=1$.

References

[1] *V. Crawford and J. Sobel, "Strategic information transmission,"* Econometrica, *vol. 50, no. 6, pp. 1431–51, 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] *H. Yamamoto, "Source coding theory for cascade and branching communication systems,"* IEEE Transactions on Information Theory, *vol. 27, no. 3, pp. 299–308, 1981.*

[5] *D. Vasudevan, C. Tian, and S. N. Diggavi, "Lossy source coding for a cascade communication system with side-informations," 2006. [Online]. Available: http://infoscience.epfl.ch/record/105103*

[6] *F. Koessler, M. Laclau, and T. Tomala, "Interactive information design,"* Mathematics of Operations Research, *06 2021.*

[7] *S. Saritas, S. Yuksel, and S. Gezici, "Quadratic multi-dimensional signaling games and affine equilibria,"* IEEE Transactions on Automatic Control, *vol. 62, no. 2, p. 605–619, Feb 2017.*

[8] *S. Sarıtaş, P. Furrer, S. Gezici, T. Linder, and S. Yüksel, "On the number of bins in equilibria for signaling games," in* 2019 IEEE International Symposium on Information Theory (ISIT), *2019, pp. 972–976.*

[9] *S. Sarıtaş, S. Yüksel, and S. Gezici, "Dynamic signaling games with quadratic criteria under nash and stackelberg equilibria,"* Automatica, *vol. 115, no. C, May 2020.*

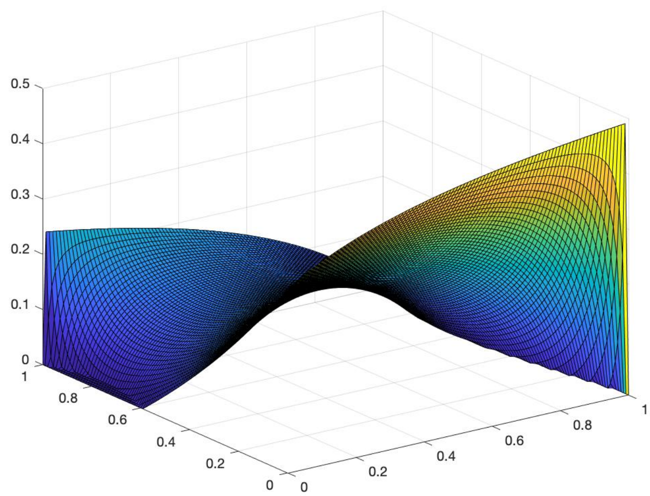

Fig. 4: *The relay's optimal cost* $C_2^{\star}(\alpha,\beta)$ *for* $p_0 = 0.4$ *and* $\alpha,\beta \in [0,1]$.

[10] S. Dughmi, D. Kempe, and R. Qiang, "Persuasion with limited communication," in Proceedings of the 2016 ACM Conference on Economics and Computation, ser. EC '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 663–680.

[11] 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.

[12] ——, "Information extraction from a strategic sender over a noisy channel," in 2020 59th IEEE Conference on Decision and Control (CDC), 2020, pp. 354–359.

[13] ——, "Optimal questionnaires for screening of strategic agents," in ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021, pp. 8173–8177.

[14] E. Akyol, C. Langbort, and T. Başar, "Strategic compression and transmission of information," in IEEE Information Theory Workshop - Fall (ITW), Oct 2015, pp. 219–223.

[15] 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.

[16] 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.

[17] ——, "Point-to-point strategic communication," IEEE Information Theory Workshop, 2020.

[18] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Transactions on Information Theory, vol. 22, no. 1, pp. 1–11, 1976.

[19] E. Akyol, C. Langbort, and T. Başar, "On the role of side information in strategic communication," in IEEE International Symposium on Information Theory (ISIT), July 2016, pp. 1626–1630.

[20] R. Bou Rouphael and M. Le Treust, "Impact of private observation in bayesian persuasion," International Conference on NETwork Games COntrol and OPtimization NetGCoop, Mar. 2020.

[21] M. Le Treust and T. Tomala, "Strategic communication with decoder side information," Information Symposium on Information Theory (ISIT), 2021.

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