A Functional Construction of Codes for Multiple Access and Broadcast Channels

Shouvik Ganguly
University of California, San Diego
La Jolla, CA 92093, USA
shgangul@eng.ucsd.edu

Lele Wang
University of British Columbia
Vancouver, BC V6T1Z4, Canada
lelewang@ece.ubc.ca

Young-Han Kim
University of California, San Diego
La Jolla, CA 92093, USA
yhk@ucsd.edu

***Abstract*—Codes are developed for two-user multiple access and broadcast channels starting from Gelfand–Pinsker codes with known block lengths, rates, and error performances. Guarantees are provided on the block error rates of the MAC and BC codes in terms of the parameters of the constituent Gelfand–Pinsker codes. These guarantees hold as long as the constituent codes satisfy the assumed properties on rate, codeword weights, and performances, irrespective of the basic structure and other properties.**

I. Introduction

The channel coding problem has been studied extensively ever since Shannon, in his seminal 1948 paper [1], modeled a point-to-point communication channel as a collection of conditional probability distributions. The point-to-point channel coding theorem was established, among others, by Shannon [1] and Gallager [2]. In a parallel direction of research, the practical problem of coding for point-to-point channels has seen enormous advances in recent years, with the advent and extensive studies of several low-complexity coding schemes that approach or achieve the Shannon capacity. Of particular note among these are the turbo codes [3], low-density parity-check (LDPC) codes [4]–[6], and polar codes [7]. Generalizing the point-to-point case, communication scenarios involving multiple senders and/or receivers were similarly modeled and studied by Ahlswede [8], Liao [9], Cover [10], Slepian and Wolf [11], Bergmans [12], and others. While a multitude of information-theoretic results on achievable rate tradeoffs (*inner bounds*) for multi-user channels exist in the literature (see, for example, [13, Chapter 1] and [14] for comprehensive reviews), we are far from achieving known inner bounds with low complexity coding schemes due to the high computational complexity in implementing some sophisticated multi-user coding schemes, such as Marton coding [13, Theorem 8.3 and Proposition 8.1] for broadcast channels and simultaneous decoding [13, Chapter 4.5.1] for multiple access channels.

At such a juncture, this work attempts to answer the following question: what happens when a code, whose performance is known in some setting through simulations or theoretical studies, is used for a different problem? More specifically, given point-to-point channel codes with certain known parameters such as block length n, rate R, and block error probability ϵ, we attempt to come up with coding schemes for multi-user channels whose performance can be directly obtained as a simple function of parameters of the original code, without redoing extensive studies in the new setting. In essence, we treat encoders and decoders of known codes as "black boxes" (or "Lego-bricks") satisfying some primitive properties and assemble them (potentially with other simple "bricks" such as interleavers or dithers) to build a bigger "box" for a different, and potentially more complicated, scenario, with performance guarantees. Such a theory enables one to leverage commercial off-the-shelf codes (such as those studied in [3]–[7]) for single user channels, or even hypothetical codes to be invented in future, to build codes for multi-user communication.

What are the minimum primitive properties these "Lego-bricks" should satisfy while being versatile in building various network communication codes? Given such "Lego-bricks", how do we assemble them in different network communication scenarios? How does the performance guarantee translate between different communication settings? These questions were studied between channel coding and Slepian–Wolf coding first by Wyner for binary symmetric channels and doubly symmetric binary source [15] and later for general binary-input channels and general Slepian–Wolf problems [16], [17]. In this paper, we propose another "Lego-brick", which is capable in building Gelfand–Pinsker codes for channel with states, channel codes for (asymmetric) point-to-point channels, and Marton coding for broadcast channels, among others. The focus of the paper is on how the performance of one code in a certain communication setting can be translated into the performance of another code in a different setting. The discussion on how to achieve an inner bound under random coding is beyond our scope.

We start out with primitive Gelfand–Pinsker (GP) codes [18] for binary-input, binary-state channels and construct codes for binary-input multiple access channels (MAC) and finite-alphabet broadcast channels (BC). In addition to primitive GP codes, we use a random interleaver that applies to a length-n sequence, a permutation chosen uniformly at random from the $n!$ possible permutations, as well as shared random bits between transmitters and receivers.

The rest of the paper is organized as follows. Section II-A

This work was supported in part by Institute of Information & communications Technology Planning & Evaluation (IITP) grant 2018-0-01410, in part by the NSERC Discovery Grant No. RGPIN-2019-05448, in part by the Rogers–UBC Collaborative Research Grant "Interference Mitigation for 5G Networks", and in part by the NSERC Collaborative Research and Development Grant No. CRDPJ 543676-19.

978-1-7281-6432-8/20/$31.00 ©2020 IEEE

introduces the primitive GP encoding and decoding blocks we use throughout the paper and establishes the performance of this code when a random interleaver is used on the output of the encoder and additional random bits are shared between the encoder and the decoder. Section II-B shows how to construct an ordinary point-to-point channel code from the primitive GP code. Section III develops a coding scheme for the 2-user binary-input MAC using two channel codes (ultimately derived from primitive GP codes). Section IV develops a coding scheme for the 2-user BC using a primitive GP code and a channel code. Throughout the paper, we follow the notation in [13], with the exception that for a natural number n, that we use $[n]$ to denote the set $\{1,\dots,n\}$. In addition, $|x^n| := |\{i \in [n] : x_i = 1\}|$ denotes the Hamming weight of a binary sequence $x^n \in \{0,1\}^n$ and for two binary sequences x^n, y^n, we denote by $x^n \oplus y^n := \{z^n : z_i = x_i \oplus y_i, i \in [n]\}$ the bitwise XOR operation or equivalently, binary addition without carry.

II. Gelfand–Pinsker codes to channel codes

A. Gelfand–Pinsker coding

A Gelfand–Pinsker problem $p(y\,|\,u,s)p(s)$ consists of finite alphabets $\mathcal{U} = \mathcal{S} = \{0,1\}$ and $\mathcal{Y}$, a collection of conditional probability mass functions $p(y,s\,|\,u)$ on $\mathcal{Y} \times \mathcal{S}$ for $u \in \mathcal{U}$ (referred to as the "channel" with input u and state s), and a probability mass function $p(s)$ on $\mathcal{S}$.

A $(R,n,\alpha,\epsilon,\delta)$ code (g,ψ) depicted in Fig. 1 for the Gelfand–Pinsker problem $p(y\,|\,u,s)p(s)$ consists of

- an encoder $g : [2^{nR}] \times \mathcal{S}^n \to \mathcal{U}^n$ that maps each message m and each state sequence s^n to a codeword $u^n = g(m,s^n)$ such that $|g(m,s^n) \oplus s^n| = n\alpha$ for every $m \in [2^{nR}]$, $s^n \in \mathcal{S}^n$,
- a decoder $\psi : \mathcal{Y}^n \to [2^{nR}]$ that assigns a message estimate $\hat{m} = \psi(y^n)$ to each received sequence y^n.

The average probability of error of the code with a *perturbed input* is defined as

$$P_e^{(n)}(z^n) := \sum_m \sum_{s^n} \left(2^{-nR} \prod_{i=1}^n p_S(s_i) \times \right.$$
$$\left. \mathsf{P}(\hat{M} \neq m \,|\, U^n = g(m,s^n) \oplus z^n, S^n = s^n) \right)$$

for $z^n \in \{0,1\}^n$. The *maximal average probability of error under sublinear perturbation* is

$$\max_{\substack{z^n \in \{0,1\}^n: \\ |z^n| \le n^{1/2+\delta}}} P_e^{(n)}(z^n) = \epsilon. \tag{1}$$

The condition (1) states that the message is decoded correctly w.h.p. as long as the Hamming weight of the perturbation z^n is not larger than $n^{1/2+\delta}$. This condition is motivated by the existence of practical codes with low decoding complexity and large block lengths n, such as Reed–Muller codes [19], [20] and BCH codes [21], for which the minimum distance can be made to grow as $n^{1/2+\delta}$ or faster by choosing code parameters appropriately.

Remark 1. For the rest of the paper, we assume that such a code satisfying (1) exists for every channel $p(y\,|\,u,s)$, every $\alpha \in (0,1)$, and every $\epsilon > 0$, however small, for some large-enough block length n and some $\delta \in (0,1/2)$.

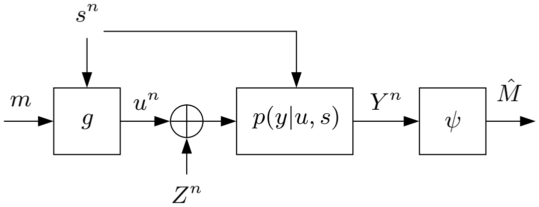

Fig. 1. Primitive GP code.

We now adapt the primitive Gelfand–Pinsker code to a form that is more useful in coding for multi-user channels. Specifically, we add a random dither $W^n \sim$ i.i.d. $\mathrm{Bern}(1/2)$ to the codeword as well as to the observed state sequence, and apply a uniform interleaver Γ_n (i.e., a permutation of n objects chosen uniformly at random) to both of them. To compensate for the interleaver, we apply the operation Γ_n^{-1} to the channel output Y^n and try to decode the message M assuming W^n is also available at the decoder. This arrangement is shown in Fig. 2. The following result connects the probability of error of this scheme to that of a slightly different Gelfand–Pinsker problem.

Lemma 2. *Consider a Gelfand–Pinsker problem $q(y,w\,|\,u,s)p(s)$, where the channel q is obtained by adding a* random dither $W^n \sim$ i.i.d. $\mathrm{Bern}(1/2)$ *to the input U^n as well as the state S^n of the channel $p(y\,|\,u,s)$, and taking (Y^n, W^n) as the output, as shown in Fig. 2. Suppose that we have a $(R,n,\alpha,\epsilon,\delta)$ code (g,ψ) for the problem $q(y,w\,|\,u,s)p(s)$. Then,*

$$\sum_{m,s^n} \left(\prod_{i=1}^n p_S(s_i) 2^{-nR} \cdot \mathsf{P}\Big(\psi(\Gamma_n^{-1}(Y^n), W^n) \neq m \Big| \right.$$
$$\left. U^n = \Gamma_n(g(m,s^n)), S^n = \Gamma_n(s^n)\Big) \right) \le \epsilon.$$

Lemma 2 illustrates that the arrangement in Fig. 2 can achieve the same rate and probability of error as a usual primitive Gelfand–Pinsker code adapted to the channel $q(y,w\,|\,u,s) \equiv p(w)p_{Y|U,S}(y\,|\,u \oplus w, s \oplus w)$.

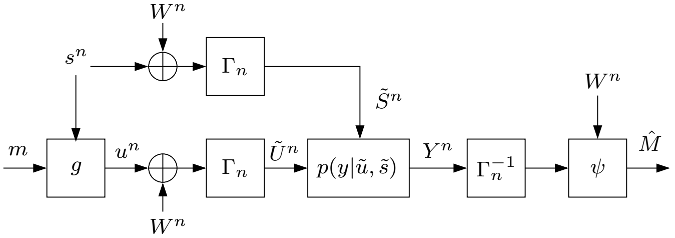

Fig. 2. GP code with interleavers and dithers.

B. Point-to-point channel coding

The Gelfand–Pinsker code described in Section II-A can be easily converted into a point-to-point channel code, as described in this section. Similar to [16], we define a binary-input discrete memoryless channel $p(y\,|\,u)$ as consisting of an input alphabet $\mathcal{U} = \{0,1\}$, a finite output alphabet $\mathcal{Y}$, and a collection of conditional probability mass functions $p(y\,|\,u)$ on $\mathcal{Y}$ for $u \in \mathcal{U}$. A (R, n, ϵ) code (f, ϕ) for the channel $p(y\,|\,u)$ consists of

- an encoder $f : [2^{nR}] \to \mathcal{U}^n$ that maps each message m to a codeword $u^n = f(m)$,
- a decoder $\phi : \mathcal{Y}^n \to [2^{nR}]$ that assigns a message estimate $\hat{m} = \phi(y^n)$ to each received sequence y^n.

The average probability of error of this code is

$$\sum_m 2^{-nR} \cdot \mathsf{P}(\phi(Y^n) \neq m \,|\, U^n = f(m)) = \epsilon.$$

Now, suppose we have a $(R, n, \alpha, \epsilon, \delta)$ code (g, ψ) for the Gelfand–Pinsker problem $p(y\,|\,u,s)p(s)$, where $p(y\,|\,u,s) \equiv p(y\,|\,u)$ (i.e., the channel output is independent of the state given the channel input) and $p_S(0) = 1$. Define $f : [2^{nR}] \to \mathcal{U}^n$ by $f(m) = g(m, \mathbf{0})$, where $\mathbf{0}$ is the all-zero sequence. Then, (f, ψ) forms a code for the channel $p(y\,|\,u)$ with length n and rate R, and has average probability of error

$$\begin{aligned}&\sum_m 2^{-nR} \cdot \mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = f(m)) \\ &= \sum_m 2^{-nR} \cdot \mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = g(m, \mathbf{0})).\end{aligned}$$

We write

$$\begin{aligned}&\mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = g(m, \mathbf{0})) \\ &= \mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = g(m, \mathbf{0}), S^n = \mathbf{0}) \prod_{i=1}^n p_S(0) \\ &+ \sum_{s^n \neq \mathbf{0}} \mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = g(m, \mathbf{0}), S^n = s^n) \prod_{i=1}^n p_S(s_i) \\ &\overset{(a)}{=} \mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = g(m, \mathbf{0}), S^n = \mathbf{0}) \\ &\overset{(b)}{=} \mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = g(m, \mathbf{0}), S^n = \mathbf{0}) \prod_{i=1}^n p_S(0) \\ &+ \sum_{s^n \neq \mathbf{0}} \mathsf{P}(\psi(Y^n) \neq m \,|\, U^n = g(m, s^n), S^n = s^n) \prod_{i=1}^n p_S(s_i) \\ &\le \epsilon,\end{aligned}$$

where (a) and (b) follow since $\prod_{i=1}^n p_S(s_i) = 0$ for $s^n \neq \mathbf{0}$, and the last step follows from condition (1). Thus, we have obtained a (R, n, ϵ') channel code from a $(R, n, \alpha, \epsilon, \delta)$ Gelfand–Pinsker code, where $\epsilon' \le \epsilon$. Similar to Lemma 2, the following result adapts the Gelfand–Pinsker code to a channel with dithered and permuted input.

Lemma 3. *Consider a Gelfand–Pinsker problem* $q(y, w\,|\,u, s)p(s)$, *where* $p_S(0) = 1$ *and the channel* $q(y, w\,|\,u, s) \equiv q(y, w\,|\,u)$ *is obtained by adding a* random dither $W^n \sim$ i.i.d. Bern$(1/2)$ *to the input* U^n *to the channel* $p(y\,|\,u)$ *and taking* (Y^n, W^n) *as the output, as shown in Fig. Consider a* $(R, n, \alpha, \epsilon, \delta)$ *code* (g, ψ) *for the problem* $q(y, w\,|\,u, s)p(s)$. *Then,*

$$\sum_m \left(2^{-nR} \cdot \mathsf{P}\left(\psi(\Gamma_n^{-1}(Y^n), W^n) \neq m \,\middle|\, U^n = \Gamma_n(g(m, \mathbf{0}))\right)\right) \le \epsilon.$$

Remark 4. While this approach of coding for a point-to-point channel by using a Gelfand–Pinsker code may seem redundant since a multitude of good codes already exist for several different classes of channels, this enables us to code for multi-user channels by using only Gelfand–Pinsker blocks, rather than adding separate channel coding blocks as and when we need it. This aligns with the spirit of this work, which aims to eventually code for complex channels, starting out with a minimal number of elementary blocks.

III. Coding for two-user MAC

In this section, we put together two channel codes, derived, as described in Section II-B, from Gelfand–Pinsker codes, to build a code for a binary-input discrete memoryless multiple-access channel (DM-MAC) $p(y\,|\,u_1, u_2)$, defined as consisting of input alphabets $\mathcal{U}_1 = \mathcal{U}_2 = \{0,1\}$, a finite output alphabet $\mathcal{Y}$, and a collection of conditional probability mass functions $p(y\,|\,u_1, u_2)$ on $\mathcal{Y}$ for $(u_1, u_2) \in \mathcal{U}_1 \times \mathcal{U}_2$. A (R_1, R_2, n, ϵ) code (f_1, f_2, ϕ) for the channel $p(y\,|\,u_1, u_2)$ consists of

- encoders $f_j : [2^{nR_j}] \to \mathcal{U}_j^n$ for $j = 1, 2$, that map each message $m_1 \in [2^{nR_1}]$ to a codeword $u_1^n = f_1(m_1)$ and each message $m_2 \in [2^{nR_2}]$ to a codeword $u_2^n = f_2(m_2)$,
- a decoder $\phi : \mathcal{Y}^n \to [2^{nR_1}] \times [2^{nR_2}]$ that assigns message estimates $(\hat{m}_1, \hat{m}_2) = \phi(y^n)$ to each received sequence y^n.

The average probability of error of this code is

$$2^{-n(R_1+R_2)} \cdot \sum_{m_1, m_2} \left(\mathsf{P}\left(\phi(Y^n) \neq (m_1, m_2) \,\middle|\, (U_1^n, U_2^n) = (f_1(m_1), f_2(m_2))\right)\right) = \epsilon.$$

Now, let us add independent dithering sequences $W_1^n, W_2^n \sim$ i.i.d. Bern$(1/2)$ to the codewords U_1^n and U_2^n and make W_1^n, W_2^n available at the decoder. We will construct a code for this modified MAC

$$\begin{aligned}&q(y, w_1, w_2\,|\,u_1, u_2) \\ &:= p(w_1)p(w_2)p_{Y|U_1,U_2}(y\,|\,u_1 \oplus w_1, u_2 \oplus w_2)\end{aligned}$$

using codes for the point-to-point channels

$$\begin{aligned}&q(y, w_1\,|\,u_1) \\ &= \sum_{u_2 \in \mathcal{U}_2} q(y, w_1\,|\,u_1, u_2)p(u_2) \\ &= \sum_{\substack{u_2 \in \mathcal{U}_2 \\ w_2 \in \mathcal{W}_2}} \left(p(w_1)p(w_2)p(u_2) \times p_{Y|U_1,U_2}(y\,|\,u_1 \oplus w_1, u_2 \oplus w_2)\right)\end{aligned}$$

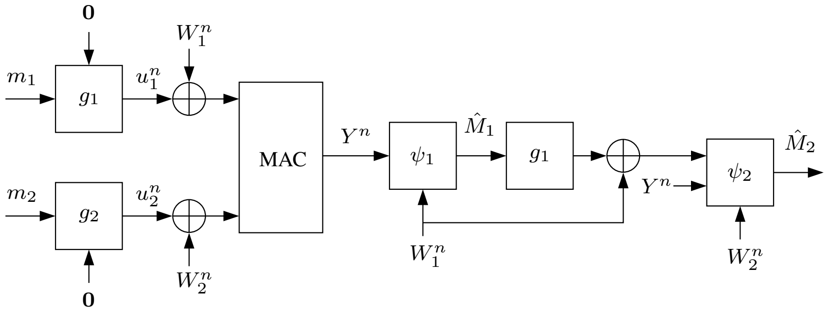

Fig. 3. Coding for MAC using primitive GP codes.

and

$$q(y, u_1 \oplus w_1, w_2 \,|\, u_2)$$
$$= p(w_2)p(u_1 \oplus w_1)p_{Y|U_1,U_2}(y \,|\, u_1 \oplus w_1, u_2 \oplus w_2).$$

Suppose that we have a $(R_1, n, \alpha, \epsilon_1, \delta)$ code (g_1, ψ_1) for the Gelfand–Pinsker problem $q(y, w_1 \,|\, u_1, s_1)p(s_1) \equiv q(y, w_1 \,|\, u_1)p(s_1)$ where $p_{S_1}(0) = 1$. Similarly, let us consider a $(R_2, n, \alpha, \epsilon_2, \delta)$ code (g_2, ψ_2) for the Gelfand–Pinsker problem $q(y, u_1 \oplus w_1, w_2 \,|\, u_2, s_2)p(s_2) \equiv q(y, u_1 \oplus w_1, w_2 \,|\, u_2)p(s_2)$, where $p_{S_2}(0) = 1$. The following result demonstrates that combining these two codes in the manner shown in Fig. 3 yields a (R_1, R_2, n, ϵ') code for the MAC $\tilde{p}(y, w_1, w_2 \,|\, u_1, u_2)$, where $\epsilon' \leq \epsilon_1 + \epsilon_2$.

Proposition 5. *Define* $f_1(m_1) := g_1(m_1, \mathbf{0})$, $f_2(m_2) := g_2(m_2, \mathbf{0})$, *and*

$$\phi(y^n, w_1^n, w_2^n)$$
$$:= (\psi_1(y^n, w_1^n), \psi_2(y^n, g_1(\psi_1((y^n, w_1^n))) \oplus w_1^n, w_2^n)),$$

i.e., the decoding function ϕ *yields message estimates* $\hat{m}_1 = \psi_1(y^n, w_1^n)$ *and* $\hat{m}_2 = \psi_2(y^n, f_1(\hat{m}_1) \oplus w_1^n, w_2^n)$. *Then, the average probability of error is bounded as*

$$\epsilon' := \mathsf{P}((\hat{M}_1, \hat{M}_2) \neq (M_1, M_2))$$
$$= 2^{-n(R_1+R_2)} \sum_{m_1,m_2} \mathsf{P}\Big(\phi(Y^n, W_1^n, W_2^n) \neq (m_1, m_2)\Big| \; U_1^n = f_1(m_1), U_2^n = f_2(m_2)\Big) \leq \epsilon_1 + \epsilon_2.$$

Proof sketch: One can show that addition of the dithering sequences ensures that the channel $U_1^n \to (Y^n, W_1^n)$ obtained by averaging over $M_2 \sim \mathrm{Unif}([2^{nR_2}])$ and W_2^n is discrete memoryless and charaterized by the conditional probability distribution

$$\prod_{i=1}^{n} p_{W_1}(w_{1i})p_{Y|U_1}(y_i \,|\, u_{1i} \oplus w_{1i}). \tag{2}$$

This implies that the probability of incorrectly decoding the first message is bounded as

$$\mathsf{P}(\psi_1(Y^n, W_1^n) \neq M_1) \leq \epsilon_1 \tag{3}$$

by our assumption on the $(R_1, n, \alpha, \epsilon_1, \delta)$ Gelfand–Pinsker code. Similarly, the channel $U_2^n \to (Y^n, U_1^n \oplus W_1^n, W_2^n)$ is discrete memoryless and therefore, if the first message is decoded correctly, the probability of incorrectly decoding the second message is bounded as

$$\mathsf{P}(\psi_2(Y^n, U_1^n \oplus W_1^n, W_2^n) \neq M_2) \leq \epsilon_2. \tag{4}$$

We can now combine (3) and (4) to bound the average probability of error $\epsilon' = \tilde{\mathsf{P}}((\hat{M}_1, \hat{M}_2) \neq (M_1, M_2))$. We write

$$\mathsf{P}\left(\{\hat{M}_1 \neq M_1\} \cup \{\hat{M}_2 \neq M_2\}\right)$$
$$= \mathsf{P}\left(\{\hat{M}_1 \neq M_1\} \cup \{\hat{M}_1 = M_1, \hat{M}_2 \neq M_2\}\right)$$
$$\leq \mathsf{P}(\hat{M}_1 \neq M_1) + \mathsf{P}(\hat{M}_1 = M_1, \hat{M}_2 \neq M_2)$$
$$= \mathsf{P}(\psi_1(Y^n, W_1^n) \neq M_1) +$$
$$\mathsf{P}(\psi_1(Y^n, W_1^n) = M_1, \psi_2(Y^n, f_1(M_1) \oplus W_1^n, W_2^n) \neq M_2)$$
$$\leq \mathsf{P}(\psi_1(Y^n, W_1^n) \neq M_1)$$
$$+ \mathsf{P}(\psi_2(Y^n, U_1^n \oplus W_1^n, W_2^n) \neq M_2) \leq \epsilon_1 + \epsilon_2. \qquad \blacksquare$$

Remark 6. The coding scheme used here corresponds to the well-known *successive cancellation decoding* [13, Chapter 4.5.1], where the message of user 1 is decoded first and the corresponding message estimate is used to decode the message of user 2. One can also implement the decoding order 2 → 1 to achieve a different rate pair for the same MAC.

IV. Coding for two-user BC

In this section, we put together a Gelfand–Pinsker code and another channel code derived from a Gelfand–Pinsker code, to build a code for a discrete memoryless broadcast channel (DM-BC) $p(y_1, y_2 \,|\, x)$, defined as consisting of a finite input alphabet $\mathcal{X}$, finite output alphabets $\mathcal{Y}_1$, $\mathcal{Y}_2$, and a collection of conditional probability mass functions $p(y_1, y_2 \,|\, x)$ on $\mathcal{Y}_1 \times \mathcal{Y}_2$ for $x \in \mathcal{X}$. A (R_1, R_2, n, ϵ) code (f, ξ_1, ξ_2) for the channel $p(y_1, y_2 \,|\, x)$ consists of

- an encoder $f : [2^{nR_1}] \times [2^{nR_2}] \to \mathcal{X}$ that maps each message pair (m_1, m_2) to a codeword $x^n = f(m_1, m_2)$,
- decoders $\xi_j : \mathcal{Y}_j^n \to [2^{nR_j}]$ for $j = 1, 2$, that assign message estimates $\hat{m}_1 = \xi_1(y_1^n)$ and $\hat{m}_2 = \xi_2(y_2^n)$ to received sequences y_1^n and y_2^n, respectively.

The average probability of error of this code is

$$2^{-n(R_1+R_2)} \sum_{m_1,m_2} \mathsf{P}\left(\{\xi_1(Y_1^n) \neq m_1\} \cup \{\xi_2(Y_2^n) \neq m_2\}\right| \; X^n = f(m_1, m_2)\Big) = \epsilon.$$

Now, let us implement the Marton coding scheme for the broadcast channel using primitive GP codes. Take $\mathcal{U}_1 = \mathcal{U}_2 =$

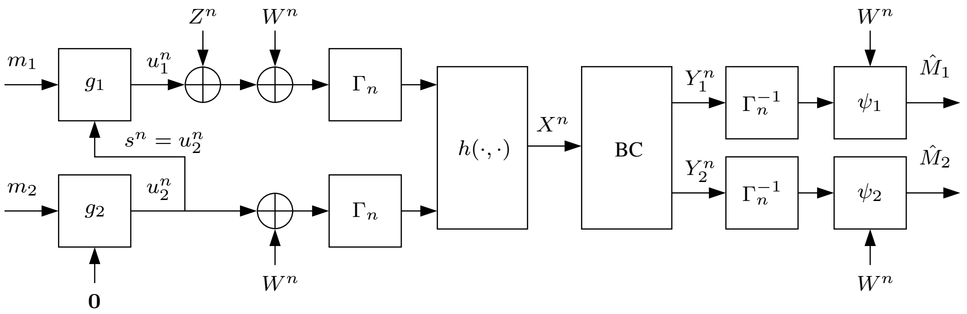

Fig. 4. Coding for BC using primitive GP codes.

$\{0,1\}$ and fix a mapping $h : \mathcal{U}_1 \times \mathcal{U}_2 \to \mathcal{X}$. We will use $u_1^n \in \mathcal{U}_1^n$ and $u_2^n \in \mathcal{U}_2^n$ to carry the messages m_1 and m_2, respectively. We will use a (R_2, n, ϵ_2) channel code (f_2, ϕ_2) for carrying the message m_2 and a $(R_1, n, \alpha, \epsilon_1, \delta)$ Gelfand–Pinsker code (g_1, ψ_1) for carrying the message m_1, using the codeword carrying m_2 as the state sequence known to the encoder, as in Marton coding [13, Chapter 8.3]. Analogous to Remark 6, we could also have flipped the *encoding order* and used a Gelfand–Pinsker code for m_2 and a channel code for m_1. In this section, we will finally add a perturbative noise Z^n to the codeword u_1^n and make use of condition (1). Z^n is generated as follows. Let $\mathcal{I}_0 \subseteq [n]$ and $\mathcal{I}_1 \subseteq [n]$ be the indices of the positions where 0 and 1, respectively, occur in $u_1^n \oplus u_2^n$, i.e., $\mathcal{I}_0 = \{i \in [n] : u_{1i} \oplus u_{2i} = 0\}$ and $\mathcal{I}_1 = \{i \in [n] : u_{1i} \oplus u_{2i} = 1\}$. Denote the (sorted) indices in $\mathcal{I}_0$ by $j_1, j_2, \dots, j_{n(1-\alpha)}$, and the sorted indices in $\mathcal{I}_1$ by $l_1, \dots, l_{n\alpha}$. Let Q be a $\mathrm{Binom}(n, \alpha)$ random variable. If $Q = k$ for $n\alpha < k \le n$, we choose Z^n to have 1s at the positions $j_1, \dots, j_{k-n\alpha}$, and 0s everywhere else. If $Q = k$ for $0 \le k < n\alpha$, we choose Z^n to have 1s at the positions $l_1, \dots, l_{n\alpha-k}$, and 0s everywhere else. Finally, if $Q = n\alpha$, we take $Z^n = \mathbf{0}$. It can be shown that with this choice of Z^n, $|Z^n \oplus u^n(m, \tilde{s}^n) \oplus \tilde{s}^n|$ is distributed as $\mathrm{Binom}(n, \alpha)$. The perturbative noise Z^n is crucial to introducing correlation among the codewords carrying the two messages. We will also add the same dithering sequence $W^n \sim$ i.i.d. $\mathrm{Bern}(1/2)$ to each of $(u_1^n \oplus Z^n)$ and u_2^n and make W^n available to both decoders as common randomness. We will then apply a random permutation $\Gamma_n(\cdot)$ on the sequences $(u_1^n \oplus Z^n \oplus W^n)$ and $u_2^n \oplus W^n$, similar to Lemma 2, and finally, generate the transmitted codeword X^n as

$$X_i = h(\Gamma_n(u_{1i} \oplus Z_i \oplus W_i), \Gamma_n(u_{2i} \oplus W_i)), \quad i \in [n].$$

The arrangement is put together as shown in Fig. 4. We note here that the Gelfand–Pinsker code used is for the effective channel $q_1(y_1, w \,|\, u_1, u_2)$, obtained by adding a random dither $W \sim \mathrm{Bern}(1/2)$ to the input U_1 and state U_2 of the channel $p(y_1 \,|\, u_1, u_2)$ defined using the BC $p(y_1, y_2 \,|\, x)$ and the map $x = h(u_1, u_2)$. Similarly, the channel code used is for the channel $q_2(y_2, w \,|\, u_2)$ obtained by adding W to the input U_2 of the channel $p(y_2 \,|\, u_2)$.

Lemma 7. *For the arrangement shown in Fig. 4,*

$$(\Gamma_n(g_1(m_1, f_2(m_2)) \oplus Z^n \oplus W^n), \Gamma_n(f_2(m_2) \oplus W^n)) \sim \text{i.i.d.} \quad \mathrm{DSBS}(\alpha)$$

for every $(m_1, m_2) \in [2^{nR_1}] \times [2^{nR_2}]$.

The following result demonstrates that combining the Gelfand–Pinsker code and the point-to-point channel code as shown in Fig. 4 yields a (R_1, R_2, n, ϵ') code for the BC $p(y_1, y_2, w \,|\, x)$, where $\epsilon' \le \epsilon_1 + \epsilon_2 + 2e^{-2n^{2\delta}}$.

Proposition 8. *For the coding scheme depicted in Fig. 4, the average probability of error is bounded as*

$$\begin{aligned}\epsilon' &:= \mathsf{P}((\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)) \\ &= 2^{-n(R_1+R_2)} \sum_{m_1, m_2} \mathsf{P}\Big(\{\psi_1(\Gamma_n^{-1}(Y_1^n), W^n) \neq m_1\} \cup \\ &\qquad\qquad \{\phi_2(\Gamma_n^{-1}(Y_2^n), W^n) \neq m_2\} \Big| \\ &\qquad\qquad U_1^n = g_1(m_1, f_2(m_2)), U_2^n = f_2(m_2)\Big) \\ &\le \epsilon_1 + \epsilon_2 + 2e^{-2n^{2\delta}}.\end{aligned}$$

Proof sketch: We first note that similar to Section III, the channel $U_2^n \to (Y_2^n, W^n)$ obtained by averaging over $M_1 \sim \mathrm{Unif}([2^{nR_1}])$ and Z^n, as well as the channel $U_1^n \to (Y_1^n, W^n)$ obtained by averaging over Z^n, is discrete memoryless. Therefore, we have

$$\begin{aligned}&\mathsf{P}(\psi_1(\Gamma_n^{-1}(Y_1^n), W^n) \neq M_1) \\ &= \sum_{|z^n| \le n^{1/2+\delta}} p(z^n)\, \mathsf{P}\left(\psi_1(\Gamma_n^{-1}(Y_1^n), W^n) \neq M_1 \middle| Z^n = z^n\right) \\ &+ \sum_{|z^n| > n^{1/2+\delta}} p(z^n)\, \mathsf{P}\left(\psi_1(\Gamma_n^{-1}(Y_1^n), W^n) \neq M_1 \middle| Z^n = z^n\right) \\ &\overset{(a)}{\le} \epsilon_1\, \mathsf{P}\left(|Z^n| \le n^{1/2+\delta}\right) + \mathsf{P}\left(|Z^n| > n^{1/2+\delta}\right) \\ &\le \epsilon_1 + \mathsf{P}(|Q - n\alpha| > n \cdot n^{-1/2+\delta}) \\ &\overset{(b)}{\le} \epsilon_1 + 2e^{-2n \cdot (n^{-1/2+\delta})^2} = \epsilon_1 + 2e^{-2n^{2\delta}},\end{aligned}$$

where in (a), we use the error probability bound (1) for the $(R_1, n, \alpha, \epsilon_1, \delta)$ Gelfand–Pinsker code and in (b), we use the Hoeffding bound [22] $\mathsf{P}(|Q - n\alpha| > n\beta) \le 2\exp(-2n\beta^2)$ for $Q \sim \mathrm{Binom}(n, \alpha)$. Similarly, by the memorylessness of $U_2^n \to (Y_2^n, W^n)$ and the assumed property of the (R_2, n, ϵ_2) channel code, we conclude that $\mathsf{P}(\phi_2(\Gamma_n^{-1}(Y_2^n), W^n) \neq M_2) \le \epsilon_2$. The result is then established by the union bound. ■

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[2] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, pp. 3–18, Jan. 1965.

[3] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *IEEE International Conference on Communications*, May 1993, pp. 1064–1070.

[4] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[5] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[6] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7761–7813, Dec. 2013.

[7] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[8] R. Ahlswede, "Multi-way communication channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Sep. 1971, pp. 23–51.

[9] H. Liao, "Multiple access channels," Ph.D. dissertation, University of Hawaii, Honolulu, 1972.

[10] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[11] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1037–1076, Sep. 1973.

[12] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, Mar. 1973.

[13] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[14] A. El Gamal and T. M. Cover, "Multiple user information theory," *Proc. IEEE*, vol. 68, no. 12, pp. 1466–1483, Dec. 1980.

[15] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 1, pp. 2–10, Jan. 1974.

[16] L. Wang and Y.-H. Kim, "Linear code duality between channel coding and Slepian–Wolf coding," in *Proc. 53rd Ann. Allerton Conf. Comm. Control Comput.*, Sep.–Oct. 2015, pp. 147–152.

[17] L. Wang, "Channel coding techniques for network communication," Ph.D. dissertation, University of California, San Diego, 2015.

[18] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Control Inf.Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[19] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *Transactions of the IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 38–49, Sep. 1954.

[20] D. E. Muller, "Application of boolean algebra to switching circuit design and to error detection," *Transactions of the I.R.E. Professional Group on Electronic Computers*, vol. EC-3, no. 3, pp. 6–12, Sep. 1954.

[21] R. Bose and D. Ray-Chaudhuri, "On a class of error correcting binary group codes," *Information and Control*, vol. 3, no. 1, pp. 68 – 79, 1960.

[22] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Am. Stat. Assoc*, vol. 58, no. 301, pp. 13–30, 1963.