[9] P. Zarrinkhat and A. H. Banihashemi, "Hybrid hard-decision iterative decoding of irregular low-density parity-check codes," *IEEE Trans Commun.*, vol. 55, no. 2, pp. 292–302, Feb. 2007.
[10] P. Zarrinkhat and A. H. Banihashemi, "Hybrid hard-decision iterative decoding of regular low-density parity-check codes," in *Proc. IEEE Int. Conf. Commun.*, 2004, pp. 435–439.
[11] P. Zarrinkhat, A. H. Banihashemi, and H. Xiao, "Time-invariant and switch-type hybrid iterative decoding of low-density parity-check codes," *Ann. Telecommun./Ann. Telecommun.*, vol. 60, no. 1-2, pp. 103–131, Jan.-Feb. 2005.
[12] S. t. Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
[13] M. Ardakani and F. R. Kschischang, "Designing irregular LDPC codes using EXIT charts based on message error rate," in *Proc. Int. Symp. Inf. Theory*, 2002, p. 454.
[14] M. Ardakani, T. H. Chan, and F. R. Kschischang, "EXIT chart properties of the highest-rate LDPC code with desired convergence behavior," *IEEE Commun. Lett.*, vol. 9, no. 1, pp. 52–54, Jan. 2005.
[15] W. Yu, M. Ardakani, B. Smith, and F. R. Kschischang, "Complexity-optimized low-density parity-check codes for Gallager decoding algorithm B," in *Proc. Int. Symp. Inf. Theory*, 2005, pp. 1488–1492.
[16] N. Miladinovic and M. P. C. Fossorier, "Improved bit-flipping decoding of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1594–1606, Apr. 2005.
[17] D. J. C. Mackay, Encyclopedia of Sparse Graph Codes, [Online]. Available: http://www.inference.phy.cam.ac.uk/mackay/codes/data.html
[18] X.-Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 386–398, Jan. 2005.
[19] R. M. Tanner, D. Sridhara, A. Sridharan, T. E. Fuja, and D. J. Costello, "LDPC block and convolutional codes based on circulant matrix," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 2966–2984, Dec. 2004.
[20] A. Amraoui, A. Montanari, and R. Urbanke, "Finite-length scaling of irregular LDPC code ensembles," in *Proc IEEE ISOC ITW Coding Complexity*, 2005, pp. 6–10.
[21] A. Amraoui, A. Montanari, and R. Urbanke, "Analytic determination of scaling parameters," in *Proc. Int. Symp. Inf. Theory*, Seattle, WA, 2006, pp. 562–566.
[22] A. Amraoui, R. Urbanke, A. Montanari, and T. Richardson, "Further results on finite-length scaling for iteratively decoded LDPC ensembles," in *Proc. Int. Symp. Inf. Theory*, Chicago, IL, 2004, pp. 101–101.
[23] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, "Finite-length scaling and finite-length shift for low-density parity-check codes," presented at the 42nd Annu. Allerton Conf. Commun. Control Comput., Monticello, IL, Sep.-Oct. 2004.
[24] P. Henrici, *Applied and Computational Complex Analysis*. New York: Wiley, 1974, vol. 1.

On The Han–Kobayashi Region for the Interference Channel

Hon-Fah Chong, *Member, IEEE*, Mehul Motani, *Member, IEEE*, Hari Krishna Garg, *Senior Member, IEEE*, and Hesham El Gamal, *Senior Member, IEEE*

***Abstract*—In this correspondence, we derive a simplified description of the Han–Kobayashi rate region for the general interference channel. Using this result, we establish that the recently discovered Chong–Motani–Garg rate region is a new representation of the Han–Kobayashi region. Moreover, a tighter bound for the cardinality of the time-sharing auxiliary random variable emerges from our simplified description.**

***Index Terms*—Han–Kobayashi region, information rates, interference channel.**

I. Background

The interference channel (IC) models the situation where M unrelated senders try to communicate their separate messages to M different receivers via a common channel. In this model, there is no cooperation between any of the senders or receivers, and hence, the transmission from each sender to its corresponding receiver is viewed as interference by the other sender–receiver pairs. In this correspondence, we limit ourselves to the two-user IC. The study of the IC was first initiated by Shannon [1] and was further studied by Ahlswede [2]. In [3], Carleial determined an improved achievable rate region for the IC. Later, Han and Kobayashi established the best achievable rate region to date for the general IC [4]. Except for the Gaussian IC under strong interference [4]–[6], the frequency-selective Gaussian IC under strong interference [7], a class of discrete degraded ICs [8] (which includes the discrete additive degraded IC studied by Benzel [9]), a class of deterministic ICs [10] (which includes the class of deterministic ICs studied by El Gamal and Costa [11]), and the discrete memoryless IC with strong interference [12], the capacity of the general IC remains unknown to date.

A. Definitions and Notations

In our notation, a discrete random variable U is assumed to take values u in a finite set $\mathcal{U}$. We use $\|\mathcal{U}\|$ to denote the cardinality of $\mathcal{U}$, and $p_U(u)$ to denote the probability distribution function of U on $\mathcal{U}$. Vectors are denoted with boldface letters, e.g., $\boldsymbol{x}^n$, where the ith element of a vector $\boldsymbol{x}^n$ is denoted by x_i. We use the symbol $\subsetneq$ to indicate proper subset.

Definition 1: [13, p. 384] Let $(X_1, X_2, \ldots, X_k)$ denote a finite collection of discrete random variables with some fixed joint distribution, $p_{X_1X_2\ldots X_k}(x_1, x_2, \ldots, x_k), (x_1, x_2, \ldots, x_k) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_k$.

Manuscript received August 7, 2006; revised December 14, 2007. This work was supported in part by the National University of Singapore under Grants R-263-000-293-112 and R-263-000-261-112.

H.-F. Chong, M. Motani, and H. K. Garg are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 119260 (e-mail: g0305962@nus.edu.sg; motani@nus.edu.sg; eleghk@nus.edu.sg).

H. El Gamal is with the Department of Electrical and Computer Engineering, The Ohio State University, 205 Dreese Laboratory, Columbus, OH 43210 USA (e-mail: helgamal@ece.osu.edu).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2008.924720

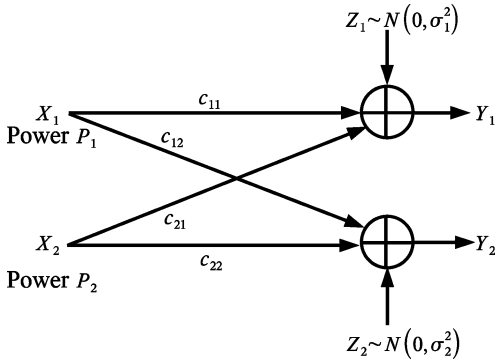

Fig. 1. The Gaussian IC.

Let S denote an ordered subset of these random variables and consider n independent copies of S. Thus

$$\Pr(\boldsymbol{S}^n = \boldsymbol{s}^n) = \prod_{i=1}^{n} \Pr(S_i = s_i), \quad \boldsymbol{s}^n \in \mathcal{S}^n.$$

The set $A_\epsilon^{(n)}$ of ϵ-typical n-sequences $(\boldsymbol{x}_1^n, \boldsymbol{x}_2^n, \ldots, \boldsymbol{x}_k^n)$ is defined as

$$A_\epsilon^{(n)}(X_1, X_2, \ldots, X_k) = \left\{ (\boldsymbol{x}_1^n, \boldsymbol{x}_2^n, \ldots, \boldsymbol{x}_k^n) : \left| -\frac{1}{n} \log p_{\boldsymbol{S}^n}(\boldsymbol{s}^n) - H(S) \right| < \epsilon, \ \forall\, S \subseteq \{X_1, X_2, \ldots, X_k\} \right\}.$$

Definition 2: A two-user discrete memoryless IC consists of two input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, two output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$, and a probability transition function $p_{Y_1Y_2|X_1X_2}(y_1, y_2|x_1, x_2)$. The conditional joint probability distribution of the memoryless IC used without feedback can be factored as

$$p_{\boldsymbol{Y}_1^n\boldsymbol{Y}_2^n|\boldsymbol{X}_1^n\boldsymbol{X}_2^n}(\boldsymbol{y}_1^n, \boldsymbol{y}_2^n|\boldsymbol{x}_1^n, \boldsymbol{x}_2^n) = \prod_{i=1}^{n} p_{Y_1Y_2|X_1X_2}(y_{1i}, y_{2i}|x_{1i}, x_{2i}).$$

Since there is no cooperation between the receivers, the capacity region of the discrete memoryless IC depends only on the conditional marginal distributions

$$p_{Y_1|X_1X_2}(y_1|x_1, x_2) = \sum_{y_2 \in \mathcal{Y}_2} p_{Y_1Y_2|X_1X_2}(y_1, y_2|x_1, x_2),$$
$$p_{Y_2|X_1X_2}(y_2|x_1, x_2) = \sum_{y_1 \in \mathcal{Y}_1} p_{Y_1Y_2|X_1X_2}(y_1, y_2|x_1, x_2).$$

Definition 3: A $\left(2^{nR_1}, 2^{nR_2}, n\right)$ code for an IC with independent information consists of two sets of integers $\mathcal{V}_1 = \left\{1, 2, \ldots, \left\lfloor 2^{nR_1} \right\rfloor\right\}$ and $\mathcal{V}_2 = \left\{1, 2, \ldots, \left\lfloor 2^{nR_2} \right\rfloor\right\}$ called the message sets, two encoding functions

$$f_1 : \mathcal{V}_1 \mapsto \mathcal{X}_1^n \quad \text{and} \quad f_2 : \mathcal{V}_2 \mapsto \mathcal{X}_2^n$$

and two decoding functions

$$g_1 : \mathcal{Y}_1^n \mapsto \mathcal{V}_1 \quad \text{and} \quad g_2 : \mathcal{Y}_2^n \mapsto \mathcal{V}_2.$$

Definition 4: The average probability of error is defined as the probability that at least one of the decoded messages is not equal to the corresponding transmitted message, i.e.,

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \times \sum_{\substack{(v_1, v_2) \\ \in \mathcal{V}_1 \times \mathcal{V}_2}} \Pr\left(g_1(Y_1^n) \neq v_1 \text{ or } g_2(Y_2^n) \neq v_2 | (v_1, v_2) \text{ sent} \right)$$

where (V_1, V_2) are assumed to be uniformly distributed over $\left\{1, 2, \ldots, \left\lfloor 2^{nR_1} \right\rfloor\right\} \times \left\{1, 2, \ldots, \left\lfloor 2^{nR_2} \right\rfloor\right\}$.

Definition 5: A rate pair (R_1, R_2) is said to be achievable for the IC if there exists a sequence of $\left(2^{nR_1}, 2^{nR_2}, n\right)$ codes with $P_e^{(n)} \to 0$.

Definition 6: The discrete-time additive white Gaussian IC, shown in Fig. 1, is described by

$$Y_1 = c_{11}X_1 + c_{21}X_2 + Z_1$$
$$Y_2 = c_{12}X_1 + c_{22}X_2 + Z_2$$

where the input and output signals are real, the coefficients c_{ij} are real constants, and the noise terms Z_1 and Z_2 are zero-mean Gaussian random variables. Also, the mean values of X_1^2 and X_2^2 cannot exceed P_1 and P_2 respectively, i.e.,

$$\mathbb{E}\left[X_1^2\right] \leq P_1 \quad \text{and} \quad \mathbb{E}\left[X_2^2\right] \leq P_2.$$

In [3], it was shown that any Gaussian IC can be reduced to a standard form by an appropriate transformation, where $c_{11}^2 = c_{22}^2 = 1$ and $\mathbb{E}\left[Z_1^2\right] = \mathbb{E}\left[Z_2^2\right] = 1$. The capacity of the Gaussian IC is not known, except for the case of no interference, where $c_{21}^2 = c_{12}^2 = 0$, for the case of strong interference, where $c_{21}^2 \geq 1$ and $c_{12}^2 \geq 1$, and for the one-sided Gaussian IC under strong interference, where $c_{12}^2 = 0$ and $c_{21}^2 \geq 1$ or $c_{21}^2 = 0$ and $c_{12}^2 \geq 1$ ([4]–[6]).

B. The Han–Kobayashi Region

In [4], Han and Kobayashi introduced five auxiliary random variables Q, U_1, W_1, U_2, and W_2, defined on arbitrary finite sets $\mathcal{Q}$, $\mathcal{U}_1$, $\mathcal{W}_1$, $\mathcal{U}_2$, and $\mathcal{W}_2$, respectively. In the Han–Kobayashi coding strategy, sender TX_1 splits the message V_1 into (V_{11}, V_{12}), where $\mathcal{V}_{11} = \left\{1, 2, \ldots, \left\lfloor 2^{nS_1} \right\rfloor\right\}$ and $\mathcal{V}_{12} = \left\{1, 2, \ldots, \left\lfloor 2^{nT_1} \right\rfloor\right\}$. Similarly, sender TX_2 splits the message V_2 into (V_{21}, V_{22}), where $\mathcal{V}_{21} = \left\{1, 2, \ldots, \left\lfloor 2^{nT_2} \right\rfloor\right\}$ and $\mathcal{V}_{22} = \left\{1, 2, \ldots, \left\lfloor 2^{nS_2} \right\rfloor\right\}$. This split aims at allowing each of the receivers to decode partial information from its nonintended sender. Hence, V_{12} represents the message intended for receiver RX_1 which can also be decoded by receiver RX_2, and similarly, V_{21} represents the message intended for receiver RX_2 which can also be decoded by receiver RX_1. Here, the auxiliary random variable W_1 serves to carry the message V_{12}, while the auxiliary random variable U_1 serves to carry the message V_{11}. The same applies to the auxiliary random variables W_2 and U_2. Hence, the encoding functions f_1 and f_2 are given by

$$f_1 : \mathcal{V}_1 \mapsto \mathcal{X}_1^n \quad \text{and} \quad f_2 : \mathcal{V}_2 \mapsto \mathcal{X}_2^n$$

where the function f_1 consists of three functions f_{11}, f_{12}, and f_{13} defined as follows:

$$f_{11} : \mathcal{V}_{11} \mapsto \mathcal{U}_1^n, \quad f_{12} : \mathcal{V}_{12} \mapsto \mathcal{W}_1^n \text{ and } f_{13} : \mathcal{U}_1^n \times \mathcal{W}_1^n \mapsto \mathcal{X}_1^n.$$

Similarly, f_2 decomposes into the following three components:

$$f_{21} : \mathcal{V}_{21} \mapsto \mathcal{W}_2^n, \quad f_{22} : \mathcal{V}_{22} \mapsto \mathcal{U}_2^n \text{ and } f_{23} : \mathcal{U}_2^n \times \mathcal{W}_2^n \mapsto \mathcal{X}_2^n.$$

In a nutshell, this strategy is basically an application of Cover's superposition coding technique [14] and was first used by Carleial [3] in the context of the Gaussian IC. Carleial made use of a sequential decoder, otherwise known as the stripping decoder. In this approach, receiver RX_1 decodes either W_1 or W_2 first before decoding U_1, whereas receiver RX_2 decodes either W_1 or W_2 first before decoding U_2. On the

other hand, Han and Kobayashi uses the more powerful joint decoder where receiver RX_1 decodes W_1, W_2, and U_1 simultaneously, while receiver RX_2 decodes W_1, W_2, and U_2 simultaneously. In addition, Han and Kobayashi introduced a time-sharing parameter Q instead of using the convex-hull operation. The time-sharing parameter Q includes, as a special case, the time-division multiplexing/frequency-division multiplexing (TDM/FDM) strategy introduced by Carleial [3] for the Gaussian IC. Next, we state the achievable rate region of Han and Kobayashi, $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$, as described in [4].[1]

Let $\mathcal{P}^*$ be the set of probability distributions $P^*(.)$ that factor as

$$\begin{aligned}&P^*(q,u_1,w_1,u_2,w_2,x_1,x_2)\\&\quad=p_Q(q)\,p_{U_1|Q}(u_1|q)\,p_{W_1|Q}(w_1|q)\,p_{U_2|Q}(u_2|q)\,p_{W_2|Q}(w_2|q)\\&\qquad\times p_{X_1|U_1W_1Q}(x_1|u_1,w_1,q)\,p_{X_2|U_2W_2Q}(x_2|u_2,w_2,q).\end{aligned}$$

and where $p(x_1|u_1,w_1,q)$ and $p(x_2|u_2,w_2,q)$ equals either 0 or 1. Suppose we fix $P^*(.)$. Consider receiver RX_1 and the set of nonnegative rate-tuples (S_1,T_1,S_2,T_2) denoted by $\mathcal{R}_{\mathrm{HK}}^{(o,1)}(P^*)$ that satisfy

$$S_1 \le I(U_1;Y_1|W_1W_2Q) \tag{1}$$
$$T_1 \le I(W_1;Y_1|U_1W_2Q) \tag{2}$$
$$T_2 \le I(W_2;Y_1|W_1U_1Q) \tag{3}$$
$$S_1+T_1 \le I(U_1W_1;Y_1|W_2Q) \tag{4}$$
$$S_1+T_2 \le I(U_1W_2;Y_1|W_1Q) \tag{5}$$
$$T_1+T_2 \le I(W_1W_2;Y_1|U_1Q) \tag{6}$$
$$S_1+T_1+T_2 \le I(U_1W_1W_2;Y_1|Q) \tag{7}$$
$$S_1,T_1,T_2 \ge 0. \tag{8}$$

Similarly, let $\mathcal{R}_{\mathrm{HK}}^{(o,2)}(P^*)$ be the set of nonnegative rate-tuples (S_1,T_1,S_2,T_2) that satisfy (1)–(8) with the indices 1 and 2 swapped everywhere. For a set $\mathcal{S}$ of 4-tuples (S_1,T_1,S_2,T_2), let $\prod(\mathcal{S})$ be the set of (R_1,R_2) such that $0\le R_1\le S_1+T_1$ and $0\le R_2\le S_2+T_2$ for some $(S_1,T_1,S_2,T_2)\in\mathcal{S}$. We have the following result.

Theorem 1 (Han–Kobayashi): The set

$$\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}} = \prod\left(\bigcup_{P^*\in\mathcal{P}^*}\mathcal{R}_{\mathrm{HK}}^{(o,1)}(P^*)\cap\mathcal{R}_{\mathrm{HK}}^{(o,2)}(P^*)\right) \tag{9}$$

is an achievable rate region for the discrete memoryless IC.

Proof: Refer to [4]. □

II. The Main Result

Our main contribution is the following compact description of the Han–Kobayashi achievable rate region. This form for the Chong–Motani–Garg representation of the Han–Kobayashi rate region [15] first appeared in [16, Theorem 3].

Theorem 2: Let $\mathcal{P}_1^*$ be the set of probability distributions $P_1^*(.)$ that factor as

$$\begin{aligned}&P_1^*(q,w_1,w_2,x_1,x_2)\\&\qquad=p_Q(q)\,p_{X_1W_1|Q}(x_1,w_1|q)\,p_{X_2W_2|Q}(x_2,w_2|q).\end{aligned} \tag{10}$$

For a fixed $P_1^*\in\mathcal{P}_1^*$, let $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*)$ be the set of (R_1,R_2) satisfying

$$R_1 \le I(X_1;Y_1|W_2Q) \tag{11}$$
$$R_2 \le I(X_2;Y_2|W_1Q) \tag{12}$$
$$R_1+R_2 \le I(X_1W_2;Y_1|Q)+I(X_2;Y_2|W_1W_2Q) \tag{13}$$
$$R_1+R_2 \le I(X_1;Y_1|W_1W_2Q)+I(X_2W_1;Y_2|Q) \tag{14}$$
$$R_1+R_2 \le I(X_1W_2;Y_1|W_1Q)+I(X_2W_1;Y_2|W_2Q) \tag{15}$$
$$2R_1+R_2 \le I(X_1W_2;Y_1|Q)+I(X_1;Y_1|W_1W_2Q)+I(X_2W_1;Y_2|W_2Q) \tag{16}$$
$$R_1+2R_2 \le I(X_2;Y_2|W_1W_2Q)+I(X_2W_1;Y_2|Q)+I(X_1W_2;Y_1|W_1Q) \tag{17}$$
$$R_1,R_2 \ge 0. \tag{18}$$

Then we have that

$$\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}} = \bigcup_{P_1^*\in\mathcal{P}_1^*}\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*) \tag{19}$$

is an achievable rate region for the interference channel. Furthermore, $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}=\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$ and the region remains invariant if we impose the following constraints on the cardinalities of the auxiliary sets:

$$\|\mathcal{W}_1\|\le\|\mathcal{X}_1\|+4,\quad \|\mathcal{W}_2\|\le\|\mathcal{X}_2\|+4,\quad\text{and}\ \|\mathcal{Q}\|\le 7. \tag{20}$$

Proof: The Han–Kobayashi rate region given in Theorem 1 can be reduced to that in Lemma 1 using Fourier–Motzkin elimination. It is then straightforward to see that $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}\subseteq\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$ since the bounds (30)–(38) are the same as (11)–(17) except for (31) and (33). In order to prove that $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}\subseteq\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$, we make use of Lemma 2. The assertion about the cardinalities of $\mathcal{W}_1$, $\mathcal{W}_2$, and $\mathcal{Q}$ follows directly from the application of Caratheodory's theorem to the expressions (11)–(17). □

Before proceeding to Lemmas 1 and 2, we need to derive a few simple results about $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$. The Han–Kobayashi rate region $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$ was derived by assuming deterministic encoding functions rather than probabilistic functions [4]. Hence, we can write the following:

$$\begin{aligned}&I(U_1;Y_1|W_1W_2Q)\\&\quad=H(Y_1|W_1W_2Q)-H(Y_1|U_1W_1W_2Q)\\&\quad=H(Y_1|W_1W_2Q)-H(Y_1|X_1U_1W_1W_2Q)\\&\quad=H(Y_1|W_1W_2Q)-H(Y_1|X_1W_1W_2Q)\\&\quad=I(X_1;Y_1|W_1W_2Q).\end{aligned} \tag{21}$$

Following along the same lines, we can write the following equalities:

$$I(U_2;Y_2|W_1W_2Q)=I(X_2;Y_2|W_1W_2Q) \tag{22}$$
$$I(U_1W_2;Y_1|W_1Q)=I(X_1W_2;Y_1|W_1Q) \tag{23}$$
$$I(U_2W_1;Y_2|W_2Q)=I(X_2W_1;Y_2|W_2Q) \tag{24}$$
$$I(U_1W_1;Y_1|W_2Q)=I(X_1;Y_1|W_2Q) \tag{25}$$
$$I(U_2W_2;Y_2|W_1Q)=I(X_2;Y_2|W_1Q) \tag{26}$$
$$I(U_1W_1W_2;Y_1|Q)=I(X_1W_2;Y_1|Q) \tag{27}$$
$$I(U_2W_1W_2;Y_2|Q)=I(X_2W_1;Y_2|Q). \tag{28}$$

Next, consider a pair of discrete and finite random variables (W,X) and define $\|\mathcal{W}\|$ independent random variables $U(w)$, $w\in\mathcal{W}$, where $P_{U(w)}(x)=P_{X|W}(x|w)$ for all $x\in\mathcal{X}$. We observe that X may be written as a function of W and $\underline{U}$, i.e., $X=f(W,\underline{U})$ where $\underline{U}=[U(x):x\in\mathcal{X}]$ is independent of W. Hence, if we set $W=W_1$ and $\underline{U}=U_1$, or $W=W_2$ and $\underline{U}=U_2$, we readily see that for a fixed $P_1^*\in\mathcal{P}_1^*$, there always exists a fixed $P^*\in\mathcal{P}^*$ such that

$$\begin{aligned}&P_1^*(q,w_1,w_2,x_1,x_2)\\&\qquad=\sum_{u_1\in\mathcal{U}_1,u_2\in\mathcal{U}_2}P^*(q,u_1,u_2,w_1,w_2,x_1,x_2).\end{aligned} \tag{29}$$

On applying Fourier–Motzkin elimination and the above equalities to Theorem 1, one may obtain the following result.

[1]We use superscript "o" and "c" to differentiate the original description of the Han–Kobayashi region from our compact description.

Lemma 1: For a fixed $P^* \in \mathcal{P}^*$, let $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*)$ be the set of all rate pairs (R_1, R_2) satisfying

$$\begin{aligned}
R_1 &\le I(X_1;Y_1|W_2Q) && (30)\\
R_1 &\le I(X_1;Y_1|W_1W_2Q) + I(W_1;Y_2|X_2Q) && (31)\\
R_2 &\le I(X_2;Y_2|W_1Q) && (32)\\
R_2 &\le I(X_2;Y_2|W_1W_2Q) + I(W_2;Y_1|X_1Q) && (33)\\
R_1 + R_2 &\le I(X_1W_2;Y_1|Q) + I(X_2;Y_2|W_1W_2Q) && (34)\\
R_1 + R_2 &\le I(X_1;Y_1|W_1W_2Q) + I(X_2W_1;Y_2|Q) && (35)\\
R_1 + R_2 &\le I(X_1W_2;Y_1|W_1Q) + I(X_2W_1;Y_2|W_2Q) && (36)\\
2R_1 + R_2 &\le I(X_1W_2;Y_1|Q) + I(X_1;Y_1|W_1W_2Q)\\
&\quad + I(X_2W_1;Y_2|W_2Q) && (37)\\
R_1 + 2R_2 &\le I(X_2;Y_2|W_1W_2Q) + I(X_2W_1;Y_2|Q)\\
&\quad + I(X_1W_2;Y_1|W_1Q) && (38)\\
R_1, R_2 &\ge 0. && (39)
\end{aligned}$$

Finally, we have

$$\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}} = \bigcup_{P^* \in \mathcal{P}^*} \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*). \quad (40)$$

Proof: We apply the equalities in (21)–(28) to [17, Theorem B] to obtain (30)–(39) in Lemma 1 and two additional inequalities

$$\begin{aligned}
2R_1 + R_2 &\le 2I(U_1;Y_1|W_1W_2Q) + I(W_1W_2;Y_2|U_2Q)\\
&\quad + I(U_2W_1;Y_2|W_2Q), && (41)\\
R_1 + 2R_2 &\le 2I(U_2;Y_2|W_1W_2Q) + I(W_1W_2;Y_1|U_1Q)\\
&\quad + I(U_1W_2;Y_1|W_1Q). && (42)
\end{aligned}$$

We note that since

$$\begin{aligned}
R_1 &\le I(U_1;Y_1|W_1W_2Q) + I(W_1;Y_2|U_2W_2Q) && (43)\\
R_1 + R_2 &\le I(U_1;Y_1|W_1W_2Q) + I(U_2W_2W_1;Y_2|Q) && (44)
\end{aligned}$$

we have

$$\begin{aligned}
2R_1 + R_2 &\le 2I(U_1;Y_1|W_1W_2Q) + I(W_1;Y_2|U_2W_2Q)\\
&\quad + I(U_2W_2W_1;Y_2|Q)\\
&= 2I(U_1;Y_1|W_1W_2Q) + I(W_1;Y_2|U_2W_2Q)\\
&\quad + I(U_2W_1;Y_2|W_2Q) + I(W_2;Y_2|Q)\\
&\le 2I(U_1;Y_1|W_1W_2Q) + I(W_1;Y_2|U_2W_2Q)\\
&\quad + I(U_2W_1;Y_2|W_2Q) + I(W_2;Y_2|U_2Q)\\
&= 2I(U_1;Y_1|W_1W_2Q) + I(W_1W_2;Y_2|U_2Q)\\
&\quad + I(U_2W_1;Y_2|W_2Q). && (45)
\end{aligned}$$

Similarly, since

$$\begin{aligned}
R_2 &\le I(U_2;Y_2|W_1W_2Q) + I(W_2;Y_1|U_1W_1Q) && (46)\\
R_1 + R_2 &\le I(U_2;Y_2|W_1W_2Q) + I(U_1W_1W_2;Y_1|Q) && (47)
\end{aligned}$$

we have

$$\begin{aligned}
R_1 + 2R_2 \le{}& 2I(U_2;Y_2|W_1W_2Q) + I(W_1W_2;Y_1|U_1Q)\\
&+ I(U_1W_2;Y_1|W_1Q). && (48)
\end{aligned}$$

Hence, the two additional inequalities (41) and (42) are redundant. □

The equivalence between $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$ and $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$ emerges from the following lemma.

Lemma 2: For a fixed $P_1^* \in \mathcal{P}_1^*$, there exists a fixed $P^* \in \mathcal{P}^*$ such that $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*) \subseteq \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \cup \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{**}) \cup \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{***})$ where

$$\begin{aligned}
&P_1^*(q, w_1, w_2, x_1, x_2)\\
&\quad = \sum_{u_1 \in \mathcal{U}_1, u_2 \in \mathcal{U}_2} P^*(q, u_1, u_2, w_1, w_2, x_1, x_2) && (49)\\
&P^{**} = \sum_{w_1 \in \mathcal{W}_1} P^* && (50)\\
&P^{***} = \sum_{w_2 \in \mathcal{W}_2} P^*. && (51)
\end{aligned}$$

Proof: Suppose (R_1, R_2) is in $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*)$ but not in $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*)$. Then either (31) or (33) is violated. If (31) is violated, we have

$$R_1 > I(W_1;Y_2|X_2Q) + I(X_1;Y_1|W_1W_2Q). \quad (52)$$

By substituting $W_1 = \phi$ in Lemma 1, we see that $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{**})$ consists of all rate pairs (R_1, R_2) such that

$$\begin{aligned}
R_1 &\le I(X_1;Y_1|W_2Q)\\
R_2 &\le I(X_2;Y_2|Q)\\
R_2 &\le I(W_2;Y_1|X_1Q) + I(X_2;Y_2|W_2Q)\\
R_1 + R_2 &\le I(X_1W_2;Y_1|Q) + I(X_2;Y_2|W_2Q).
\end{aligned}$$

However, from (11), we obtain

$$R_1 \le I(X_1;Y_1|W_2Q);$$

from (52) and (14), we obtain

$$R_2 < I(X_2;Y_2|Q);$$

from (52) and (15), we obtain

$$\begin{aligned}
R_2 &< I(W_2;Y_1|W_1Q) + I(X_2;Y_2|W_2Q)\\
&\le I(W_2;Y_1|X_1Q) + I(X_2;Y_2|W_2Q);
\end{aligned}$$

and from (52) and (16), we obtain

$$R_1 + R_2 \le I(X_1W_2;Y_1|Q) + I(X_2;Y_2|W_2Q).$$

We see that (R_1, R_2) satisfying the above constraints are in $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{**})$. The proof for the case of

$$R_2 > I(W_2;Y_1|X_1Q) + I(X_2;Y_2|W_1W_2Q)$$

follows exactly along the same lines. It readily follows that

$$\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*) \subseteq \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \cup \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{**}) \cup \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{***}). \quad \square$$

Finally, since

$$\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*) \subseteq \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \cup \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{**}) \cup \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^{***})$$

it immediately follows that $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}} \subseteq \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$ and since $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}} \subseteq \mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$, we obtain our result $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}} = \mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$.

III. Discussion

In this section, we make a few remarks about our results.

Remark 1: Han and Kobayashi make use of the polymatroidal structure underlying the collection of bounds that specify the region $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$, (1)–(7), and their counterparts at receiver RX_2, to convert them to a set of bounds on R_1, R_2, $R_1 + R_2$, $2R_1 + R_2$, and $R_1 + 2R_2$ [4, Theorem 4.1]. Even though Theorem 2 is just a different description of the Han–Kobayashi rate region, it gives the simplest description of the best rate region to date. From [4, Theorem 4.1], the cardinalities of the auxiliary sets is given by $\|\mathcal{W}_1\| \le \|\mathcal{X}_1\| + 7$, $\|\mathcal{W}_2\| \le \|\mathcal{X}_2\| + 7$,

$\|\mathcal{U}_1\| \leq \|\mathcal{X}_1\|+2$, $\|\mathcal{U}_2\| \leq \|\mathcal{X}_2\|+2$, and $\|\mathcal{Q}\| \leq 11$. Hence, Theorem 2 also gives us tighter bounds for the cardinalities of the auxiliary sets. Another interesting observation is that even though the coding technique requires the use of the auxiliary random variables U_1 and U_2, the rate region $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$ does not depend on these auxiliary random variables. Hence, cardinality bounds on $\mathcal{U}_1$ and $\mathcal{U}_2$ are unnecessary.

Remark 2: We observe that the Chong–Motani–Garg region, i.e., $\mathcal{R}_{\mathrm{CMG}}$, reported in [15], is in fact equivalent to the Han–Kobayashi region. This equivalence sheds light on the two main insights behind our compact description of the Han–Kobayashi region. We first observe that for receiver RX_1, no decoding error is committed if the message $V_1 = (V_{11}, V_{12})$ is decoded correctly but the message V_{21} is decoded wrongly. The same applies to receiver RX_2. This implies that constraint (3) and its counterpart for receiver RX_2, are unnecessary to drive the overall probability of error to ϵ. Moreover, the coding scheme considered in [15] uses only three auxiliary random variables Q, W_1, and W_2 defined on arbitrary finite sets $\mathcal{Q}$, $\mathcal{W}_1$, and $\mathcal{W}_2$. The auxiliary random variables W_1 and W_2 now serve as cloud centers that can be distinguished by both receivers. For sender TX_1, instead of generating two independent codebooks with codewords $\boldsymbol{w}_1^n(j)$ and $\boldsymbol{u}_1^n(k)$, for each codeword $\boldsymbol{w}_1^n(j)$, we generate a codebook with codewords $\boldsymbol{x}_1^n(j,k)$, where $j \in \left\{1, 2, \ldots, \left\lfloor 2^{nT_1} \right\rfloor\right\}$ and $k \in \left\{1, 2, \ldots, \left\lfloor 2^{nS_1} \right\rfloor\right\}$. This construction renders the constraints (2) and (6), and their counterparts for receiver RX_2, unnecessary. Combining these two observations yields the following result.

Lemma 3: Let $\mathcal{R}_{\mathrm{CMG}}^{(1)}(P_1^*)$ be the set of nonnegative rate-tuples (S_1, T_1, S_2, T_2) that satisfy

$$S_1 \leq I(X_1; Y_1|W_1W_2Q) \tag{53}$$
$$S_1 + T_2 \leq I(W_2X_1; Y_1|W_1Q) \tag{54}$$
$$S_1 + T_1 \leq I(X_1; Y_1|W_2Q) \tag{55}$$
$$S_1 + T_1 + T_2 \leq I(W_2X_1; Y_1|Q) \tag{56}$$
$$S_1, T_1, T_2 \geq 0. \tag{57}$$

Similarly, let $\mathcal{R}_{\mathrm{CMG}}^{(2)}(P_1^*)$ be the set of nonnegative rate-tuples (S_1, T_1, S_2, T_2) that satisfy (53)–(57) with the indices 1 and 2 swapped. Then, the set given by

$$\mathcal{R}_{\mathrm{CMG}} = \prod \left(\bigcup_{P_1^* \in \mathcal{P}_1^*} \mathcal{R}_{\mathrm{CMG}}^{(1)}(P_1^*) \cap \mathcal{R}_{\mathrm{CMG}}^{(2)}(P_1^*) \right) \tag{58}$$

is an achievable rate region for the discrete memoryless IC.

Proof: Refer to Appendix A. □

We can see that $\mathcal{R}_{\mathrm{CMG}} = \mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$ through the following simple argument. First, since we can choose a fixed P_1^* such that

$$P_1^*(q, w_1, w_2, x_1, x_2) = \sum_{u_1 \in \mathcal{U}_1, u_2 \in \mathcal{U}_2} P^*(q, u_1, u_2, w_1, w_2, x_1, x_2). \tag{59}$$

We readily see that $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \subseteq \mathcal{R}_{\mathrm{CMG}}(P_1^*)$, and hence, $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}} \subseteq \mathcal{R}_{\mathrm{CMG}}$. The bounds (53)–(57) and their counterparts at receiver RX_2 can be again simplified using Fourier–Motzkin elimination to obtain the following result.

Lemma 4: For a fixed $P_1^* \in \mathcal{P}_1^*$, let $\mathcal{R}_{\mathrm{CMG}}(P_1^*)$ be the set of (R_1, R_2) satisfying

$$R_1 \leq I(X_1; Y_1|W_2Q) \tag{60}$$
$$R_1 \leq I(X_1; Y_1|W_1W_2Q) + I(X_2W_1; Y_2|W_2Q) \tag{61}$$
$$R_2 \leq I(X_2; Y_2|W_1Q) \tag{62}$$
$$R_2 \leq I(X_2; Y_2|W_1W_2Q) + I(X_1W_2; Y_1|W_1Q) \tag{63}$$
$$R_1 + R_2 \leq I(X_1W_2; Y_1|Q) + I(X_2; Y_2|W_1W_2Q) \tag{64}$$
$$R_1 + R_2 \leq I(X_1; Y_1|W_1W_2Q) + I(X_2W_1; Y_2|Q) \tag{65}$$
$$R_1 + R_2 \leq I(X_1W_2; Y_1|W_1Q) + I(X_2W_1; Y_2|W_2Q) \tag{66}$$
$$2R_1 + R_2 \leq I(X_1W_2; Y_1|Q) + I(X_1; Y_1|W_1W_2Q) + I(X_2W_1; Y_2|W_2Q) \tag{67}$$
$$R_1 + 2R_2 \leq I(X_2; Y_2|W_1W_2Q) + I(X_2W_1; Y_2|Q) + I(X_1W_2; Y_1|W_1Q) \tag{68}$$
$$R_1, R_2 \geq 0. \tag{69}$$

Then we have

$$\mathcal{R}_{\mathrm{CMG}} = \bigcup_{P_1^* \in \mathcal{P}_1^*} \mathcal{R}_{\mathrm{CMG}}(P_1^*). \tag{70}$$

is an achievable rate region for the IC.

Proof: Refer to [17, Theorem D]. □

One can readily see that $\mathcal{R}_{\mathrm{CMG}} \subseteq \mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$ since the Chong–Motani–Garg rate region for the general IC has two additional constraints. Hence, we see that the two rate regions are equivalent, i.e., $\mathcal{R}_{\mathrm{CMG}} = \mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$.

Remark 3: We note that the only differences between $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*)$, $\mathcal{R}_{\mathrm{CMG}}(P_1^*)$, and $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*)$ lie only in the bounds for R_1 and R_2. This observation allows for answering the question posed by Kramer in [16] on the existence of $P^* \in \mathcal{P}^*$ such that $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \subsetneq \mathcal{R}_{\mathrm{CMG}}(P_1^*)$ for certain ICs where

$$P_1^*(q, w_1, w_2, x_1, x_2) = \sum_{u_1 \in \mathcal{U}_1, u_2 \in \mathcal{U}_2} P^*(q, u_1, u_2, w_1, w_2, x_1, x_2). \tag{71}$$

For the Gaussian IC, when we set $|\mathcal{Q}| = 1$, we can easily determine parameters where $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \subsetneq \mathcal{R}_{\mathrm{CMG}}(P_1^*)$. We assume the following customary restriction on the input signals where W_1, W_2, X_1, and X_2 are Gaussian random variables:

$$\frac{\mathbb{E}\left[W_1^2\right]}{\mathbb{E}\left[X_1^2\right]} = \alpha, \quad \frac{\mathbb{E}\left[W_2^2\right]}{\mathbb{E}\left[X_2^2\right]} = \beta \tag{72}$$

such that $\alpha \in [0,1]$, $\beta \in [0,1]$, $\mathbb{E}\left[X_1^2\right] = P_1$, and $\mathbb{E}\left[X_2^2\right] = P_2$. From Fig. 2, when we set $P_1 = P_2 = 1$, $c_{12}^2 = c_{21}^2 = 0.4$, and $\alpha = 0.5$ and $\beta = 0.85$, $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \subsetneq \mathcal{R}_{\mathrm{CMG}}(P_1^*) \subsetneq \mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*)$.

It is interesting to note that there exist fixed distributions satisfying (71) where $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \subsetneq \mathcal{R}_{\mathrm{CMG}}(P_1^*) \subsetneq \mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*)$. However, when maximized over all possible distributions, all three descriptions are equivalent, i.e., they describe the same rate region.

Remark 4: This work started with the effort of the first three authors in describing a new coding scheme for the IC [15] (see also Remark 2). Kramer then applied Fourier–Motzkin elimination to the Chong–Motani–Garg rate region [16] and showed that the sum-rate bound is the same as the Han–Kobayashi rate region. Kobayashi and Han also applied Fourier–Motzkin elimination to both their original Han–Kobayashi rate region and the Chong–Motani–Garg rate region [17]. In fact, we are adopting the description in [17] of both the Han–Kobayashi and the Chong–Motani–Garg rate regions in this correspondence. Finally, the first three authors and the fourth author independently proved that two of the inequalities in both the Chong–Motani–Garg and the Han–Kobayashi rate regions are unnecessary. A shortened version of the fourth author's proof was given during the presentation at the 2006 Zurich Seminar [16]; we adopt this in the correspondence. This proves the equivalence between $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}$ and $\mathcal{R}_{\mathrm{CMG}}$ and establishes the simplified form $\mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}$.

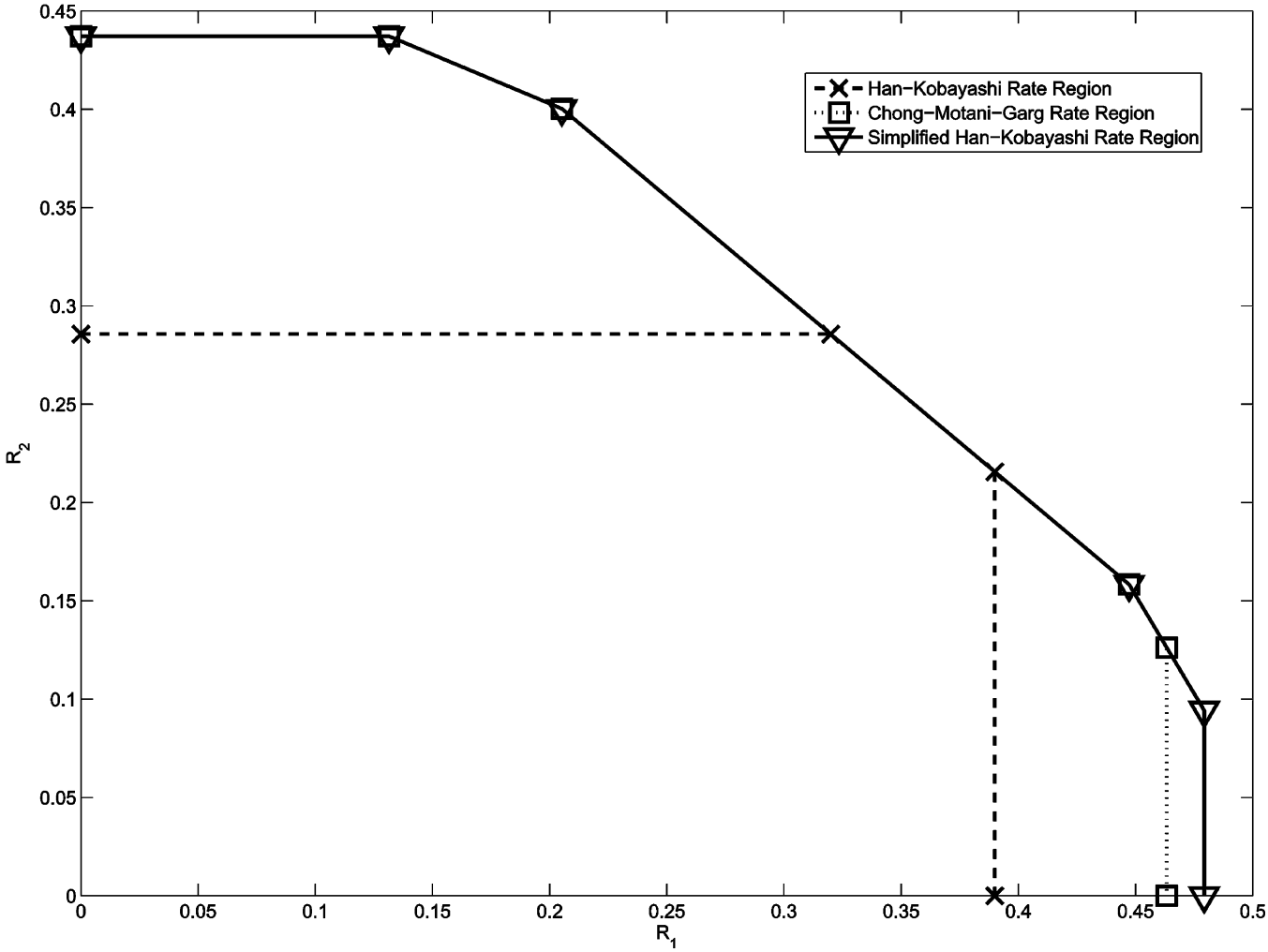

Fig. 2. An example where $\mathcal{R}_{\mathrm{HK}}^{\mathrm{o}}(P^*) \subsetneq \mathcal{R}_{\mathrm{CMG}}(P_1^*) \subsetneq \mathcal{R}_{\mathrm{HK}}^{\mathrm{c}}(P_1^*)$.

IV. Conclusion

Our main contribution is a simplified description of the celebrated Han–Kobayashi inner bound on the capacity region of the IC. This description sheds more light on the role of the auxiliary random variables and the corresponding coding/decoding strategy.

Appendix A
Proof of Lemma 3

Codebook Generation: Generate a codeword $\boldsymbol{q}^n$ of length n, generating each element independent and identically distributed (i.i.d.) according to $\prod_{i=1}^{n} p_Q(q_i)$. For the codeword $\boldsymbol{q}^n$, generate $\lfloor 2^{nT_1} \rfloor$ conditionally independent codewords $\boldsymbol{w}_1^n(j)$, $j \in \{1,2,\ldots,\lfloor 2^{nT_1} \rfloor\}$, generating each element i.i.d. according to $\prod_{i=1}^{n} p_{W_1|Q}(w_{1i}|q_i)$. For the codeword $\boldsymbol{q}^n$, and each of the codewords $\boldsymbol{w}_1^n(j)$, generate $\lfloor 2^{nS_1} \rfloor$ conditionally independent codewords $\boldsymbol{x}_1^n(j,k)$, $k \in \{1,2,\ldots,\lfloor 2^{nS_1} \rfloor\}$, generating each element i.i.d. according to $\prod_{i=1}^{n} p_{X_1|W_1Q}(x_{1i}|w_{1i}(j), q_i)$. For the codeword $\boldsymbol{q}^n$, generate $\lfloor 2^{nT_2} \rfloor$ conditionally independent codewords $\boldsymbol{w}_2^n(l)$, $l \in \{1,2,\ldots,\lfloor 2^{nT_2} \rfloor\}$, generating each element i.i.d. according to $\prod_{i=1}^{n} p_{W_2|Q}(w_{2i}|q_i)$. For the codeword $\boldsymbol{q}^n$, and each of the codewords $\boldsymbol{w}_2^n(l)$, generate $\lfloor 2^{nS_2} \rfloor$ conditionally independent codewords $\boldsymbol{x}_2^n(l,m)$, $m \in \{1,2,\ldots,\lfloor 2^{nS_2} \rfloor\}$, generating each element i.i.d. according to $\prod_{i=1}^{n} p_{X_2|W_2Q}(x_{2i}|w_{2i}(l), q_i)$.

Encoding: For encoder 1, to send the codeword pair (j,k), send the corresponding codeword $\boldsymbol{x}_1^n(j,k)$. For encoder 2, to send the codeword pair (l,m), send the corresponding codeword $\boldsymbol{x}_2^n(l,m)$.

Decoding: Receiver 1 determines the unique $\left(\hat{j},\hat{k}\right)$ and a $\hat{l}$ such that

$$\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n\left(\hat{j}\right), \boldsymbol{x}_1^n\left(\hat{j},\hat{k}\right), \boldsymbol{w}_2^n\left(\hat{l}\right), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}(Q, W_1, X_1, W_2, Y_1). \quad (73)$$

Receiver 2 determines the unique $\left(\hat{l},\hat{m}\right)$ and a $\hat{\hat{j}}$ such that

$$\left(\boldsymbol{q}^n, \boldsymbol{w}_2^n\left(\hat{l}\right), \boldsymbol{x}_2^n\left(\hat{l},\hat{m}\right), \boldsymbol{w}_1^n\left(\hat{\hat{j}}\right), \boldsymbol{y}_2^n\right) \in A_\epsilon^{(n)}(Q, W_2, X_2, W_1, Y_2). \quad (74)$$

Analysis of the Probability of Error: We consider only the decoding error of probability for receiver RX_1. The same analysis applies for receiver RX_2. By the symmetry of the random code construction, the conditional probability of error does not depend on which pair of indices is sent. Thus, the conditional probability of error is the same as the unconditional probability of error. So, without loss of generality, we assume that $(j,k) = (1,1)$ and $(l,m) = (1,1)$ was sent.

We have an error if the correct codewords, $\{\boldsymbol{w}_1^n(1), \boldsymbol{x}_1^n(1,1), \boldsymbol{w}_2^n(1)\}$ are not jointly typical with the received

sequence. If incorrect codewords $\left\{\boldsymbol{w}_1^n\left(\hat{j}\right), \boldsymbol{x}_1^n\left(\hat{j},\hat{k}\right), \boldsymbol{w}_2^n\left(\hat{l}\right)\right\}$ are jointly typical with the received codeword, i.e., $\hat{j} \neq 1$ or $\hat{k} \neq 1$, an error is also declared. However, no error is declared if $\left\{\boldsymbol{w}_1^n(1), \boldsymbol{x}_1^n(1,1), \boldsymbol{w}_2\left(\hat{l} \neq 1\right)\right\}$ are jointly typical with the received sequence. Define the following event:

$$E_{jkl} = \left\{\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n(j), \boldsymbol{x}_1^n(j,k), \boldsymbol{w}_2^n(l), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}\right\}. \tag{75}$$

Then by the union of events bound

$$\begin{aligned}
P_e^{(n)} &= P\left(E_{111}^c \bigcup \cup_{(j,k)\neq(1,1)} E_{jkl}\right) \\
&\leq P\left(E_{111}^c\right) + \sum_{j\neq 1,k=1,l=1} P\left(E_{j11}\right) \\
&\quad + \sum_{j=1,k\neq 1,l=1} P\left(E_{1k1}\right) + \sum_{j\neq 1,k\neq 1,l=1} P\left(E_{jk1}\right) \\
&\quad + \sum_{j\neq 1,k=1,l\neq 1} P\left(E_{j1l}\right) + \sum_{j=1,k\neq 1,l\neq 1} P\left(E_{1kl}\right) \\
&\quad + \sum_{j\neq 1,k\neq 1,l\neq 1} P\left(E_{jkl}\right) \\
&\leq P\left(E_{111}^c\right) + 2^{nT_1} 2^{-n(I(X_1;Y_1|W_2Q)-4\epsilon)} \\
&\quad + 2^{nS_1} 2^{-n(I(X_1;Y_1|W_1W_2Q)-4\epsilon)} \\
&\quad + 2^{n(S_1+T_1)} 2^{-n(I(X_1;Y_1|W_2Q)-4\epsilon)} \\
&\quad + 2^{n(T_1+T_2)} 2^{-n(I(W_2X_1;Y_1|Q)-4\epsilon)} \\
&\quad + 2^{n(S_1+T_2)} 2^{-n(I(W_2X_1;Y_1|W_1Q)-4\epsilon)} \\
&\quad + 2^{n(S_1+T_1+T_2)} 2^{-n(I(W_2X_1;Y_1|Q)-4\epsilon)}.
\end{aligned} \tag{76}$$

Since $\epsilon > 0$ is arbitrary, the conditions of Lemma 3 imply that each term tends to 0 as $n \to \infty$. Refer to Appendix B for a detailed analysis of the error probabilities. The above bound shows that the average probability of error, averaged over all choices of codebooks in the random code construction, is arbitrarily small. Hence, there exists at least one code $\mathcal{C}^*$ with arbitrarily small probability of error.

Appendix B
Computation of the Error Probability

We only consider the error probability of receiver RX_1. The typicality inequalities follow from [13, Theorems 14.2.1–14.2.3]. For $j \neq 1$, we have

$$\begin{aligned}
&P\left(E_{j11}\right) \\
&\quad = P\left(\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n(j), \boldsymbol{x}_1^n(j,1), \boldsymbol{w}_2^n(1), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}\right) \\
&\quad = \sum_{\substack{(\boldsymbol{q}^n, \boldsymbol{w}_1^n, \boldsymbol{x}_1^n, \\ \boldsymbol{w}_2^n, \boldsymbol{y}_1^n) \in A_\epsilon^{(n)}}} p\left(\boldsymbol{w}_1^n, \boldsymbol{x}_1^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{w}_2^n, \boldsymbol{y}_1^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{q}^n\right) \\
&\quad \leq \left|A_\epsilon^{(n)}\right| 2^{-n(H(W_1X_1|Q)-\epsilon)} 2^{-n(H(W_2Y_1|Q)-\epsilon)} 2^{-n(H(Q)-\epsilon)} \\
&\quad \leq 2^{-n(H(W_1X_1|Q)+H(W_2Y_1|Q)+H(Q)-H(QW_1X_1W_2Y_1)-4\epsilon)} \\
&\quad = 2^{-n(I(X_1;Y_1|W_2Q)-4\epsilon)}.
\end{aligned}$$

For $k \neq 1$ we have

$$\begin{aligned}
&P\left(E_{1k1}\right) \\
&\quad = P\left(\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n(1), \boldsymbol{x}_1^n(1,k), \boldsymbol{w}_2^n(1), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}\right) \\
&\quad = \sum_{\substack{(\boldsymbol{q}^n, \boldsymbol{w}_1^n, \boldsymbol{x}_1^n, \\ \boldsymbol{w}_2^n, \boldsymbol{y}_1^n) \in A_\epsilon^{(n)}}} p\left(\boldsymbol{x}_1^n | \boldsymbol{q}^n \boldsymbol{w}_1^n\right) p\left(\boldsymbol{y}_1^n | \boldsymbol{q}^n \boldsymbol{w}_1^n \boldsymbol{w}_2^n\right) p\left(\boldsymbol{q}^n \boldsymbol{w}_1^n \boldsymbol{w}_2^n\right) \\
&\quad \leq \left|A_\epsilon^{(n)}\right| 2^{-n(H(X_1|QW_1)-\epsilon)} 2^{-n(H(Y_1|QW_1W_2)-\epsilon)} \\
&\qquad \times 2^{-n(H(QW_1W_2)-\epsilon)} \\
&\quad \leq 2^{-n(H(X_1|QW_1)+H(Y_1|QW_1W_2)+H(QW_1W_2)-4\epsilon)} \\
&\qquad \times 2^{nH(QW_1X_1W_2Y_1)} \\
&\quad = 2^{-n(I(X_1;Y_1|W_1W_2Q)-4\epsilon)}.
\end{aligned}$$

For $j \neq 1, k \neq 1$ we have

$$\begin{aligned}
&P\left(E_{jk1}\right) \\
&\quad = P\left(\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n(j), \boldsymbol{x}_1^n(j,k), \boldsymbol{w}_2^n(1), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}\right) \\
&\quad = \sum_{\substack{(\boldsymbol{q}^n, \boldsymbol{w}_1^n, \boldsymbol{x}_1^n, \\ \boldsymbol{w}_2^n, \boldsymbol{y}_1^n) \in A_\epsilon^{(n)}}} p\left(\boldsymbol{w}_1^n, \boldsymbol{x}_1^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{w}_2^n, \boldsymbol{y}_1^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{q}^n\right) \\
&\quad \leq \left|A_\epsilon^{(n)}\right| 2^{-n(H(W_1X_1|Q)-\epsilon)} 2^{-n(H(W_2Y_1|Q)-\epsilon)} 2^{-n(H(Q)-\epsilon)} \\
&\quad \leq 2^{-n(H(W_1X_1|Q)+H(W_2Y_1|Q)+H(Q)-H(QW_1X_1W_2Y_1)-4\epsilon)} \\
&\quad = 2^{-n(I(X_1;Y_1|W_2Q)-4\epsilon)}.
\end{aligned}$$

For $j \neq 1, l \neq 1$ we have

$$\begin{aligned}
&P\left(E_{j1l}\right) \\
&\quad = P\left(\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n(j), \boldsymbol{x}_1^n(j,1), \boldsymbol{w}_2^n(l), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}\right) \\
&\quad = \sum_{\substack{(\boldsymbol{q}^n, \boldsymbol{w}_1^n, \boldsymbol{x}_1^n, \\ \boldsymbol{w}_2^n, \boldsymbol{y}_1^n) \in A_\epsilon^{(n)}}} p\left(\boldsymbol{w}_1^n \boldsymbol{x}_1^n \boldsymbol{w}_2^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{y}_1^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{q}^n\right) \\
&\quad \leq \left|A_\epsilon^{(n)}\right| 2^{-n(H(W_1X_1W_2|Q)-\epsilon)} \\
&\qquad \times 2^{-n(H(Y_1|Q)-\epsilon)} 2^{-n(H(Q)-\epsilon)} \\
&\quad \leq 2^{-n(H(W_1X_1W_2|Q)+H(Y_1|Q)+H(Q)-H(QW_1X_1W_2Y_1)-4\epsilon)} \\
&\quad \leq 2^{-n(I(X_1W_2;Y_1|Q)-4\epsilon)}.
\end{aligned}$$

For $k \neq 1, l \neq 1$ we have

$$\begin{aligned}
&P\left(E_{1kl}\right) \\
&\quad = P\left(\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n(1), \boldsymbol{x}_1^n(1,k), \boldsymbol{w}_2^n(l), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}\right) \\
&\quad = \sum_{\substack{(\boldsymbol{q}^n, \boldsymbol{w}_1^n, \boldsymbol{x}_1^n, \\ \boldsymbol{w}_2^n, \boldsymbol{y}_1^n) \in A_\epsilon^{(n)}}} p\left(\boldsymbol{x}_1^n \boldsymbol{w}_2^n | \boldsymbol{q}^n \boldsymbol{w}_1^n\right) p\left(\boldsymbol{y}_1^n | \boldsymbol{q}^n \boldsymbol{w}_1^n\right) p\left(\boldsymbol{q}^n \boldsymbol{w}_1^n\right) \\
&\quad \leq \left|A_\epsilon^{(n)}\right| 2^{-n(H(X_1W_2|QW_1)-\epsilon)} 2^{-n(H(Y_1|QW_1)-\epsilon)} \\
&\qquad \times 2^{-n(H(QW_1)-\epsilon)} \\
&\quad = 2^{-n(H(X_1W_2|QW_1)+H(Y_1|QW_1)+H(QW_1)-4\epsilon)} \\
&\qquad \times 2^{nH(QW_1X_1W_2Y_1)} \\
&\quad \leq 2^{-n(I(X_1W_2;Y_1|W_1Q)-4\epsilon)}.
\end{aligned}$$

For $j \neq 1, k \neq 1, l \neq 1$ we have

$$\begin{aligned}
&P\left(E_{jkl}\right) \\
&\quad = P\left(\left(\boldsymbol{q}^n, \boldsymbol{w}_1^n(j), \boldsymbol{x}_1^n(j,k), \boldsymbol{w}_2^n(l), \boldsymbol{y}_1^n\right) \in A_\epsilon^{(n)}\right) \\
&\quad = \sum_{\substack{(\boldsymbol{q}^n, \boldsymbol{w}_1^n, \boldsymbol{x}_1^n, \\ \boldsymbol{w}_2^n, \boldsymbol{y}_1^n) \in A_\epsilon^{(n)}}} p\left(\boldsymbol{w}_1^n \boldsymbol{x}_1^n \boldsymbol{w}_2^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{y}_1^n | \boldsymbol{q}^n\right) p\left(\boldsymbol{q}^n\right) \\
&\quad \leq \left|A_\epsilon^{(n)}\right| 2^{-n(H(W_1X_1W_2|Q)-\epsilon)} \\
&\qquad \times 2^{-n(H(Y_1|Q)-\epsilon)} 2^{-n(H(Q)-\epsilon)} \\
&\quad \leq 2^{-n(H(W_1X_1W_2|Q)+H(Y_1|Q)+H(Q)-H(QW_1X_1W_2Y_1)-4\epsilon)} \\
&\quad = 2^{-n(I(X_1W_2;Y_1|Q)-4\epsilon)}.
\end{aligned}$$

ACKNOWLEDGMENT

We are grateful to the Associate Editor, Gerhard Kramer, for his detailed comments on earlier versions of this manuscript, for his insightful suggestions, and for providing us with a simplified proof of the existence of deterministic encoding functions achieving the same marginal probability distributions. Comments made by the anonymous reviewers proved most helpful as well. We would like to thank one of the anonymous reviewers for pointing out the fact that two of the inequalities (41) and (42) were redundant, hence, simplifying our proof of the main result.

REFERENCES

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, Berkeley, CA, 1961, vol. 1, pp. 611–644.
[2] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, no. 5, pp. 805–814, 1974.
[3] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.
[4] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.
[5] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 5, pp. 569–570, Sep. 1975.
[6] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 6, pp. 786–788, Nov. 1981.
[7] S. T. Chung and J. M. Cioffi, "The capacity region of frequency-selective Gaussian interference channels under strong interference," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1812–1821, Sep. 2007.
[8] N. Liu and S. Ulukus, "The capacity region of a class of discrete degraded interference channels," in *Proc. 44th Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sep. 2006.
[9] R. Benzel, "The capacity region of a class of discrete additive degraded interference channels," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 2, pp. 228–231, Mar. 1979.
[10] H. F. Chong, M. Motani, and H. K. Garg, "The capacity region of a class of interference channels," in *Proc. 2007 IEEE Int. Symp. Information Theory.*, Nice, France, Jun. 2007, pp. 2856–2860.
[11] A. A. El Gamal and M. H. M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 343–346, Mar. 1982.
[12] M. H. M. Costa and A. A. El Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 5, pp. 710–711, Sep. 1987.
[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[14] T. M. Cover, "An achievable rate region for the broadcasting channel," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 4, pp. 399–404, Jul. 1975.
[15] H. F. Chong, M. Motani, and H. K. Garg, "A comparison of two achievable rate regions for the interference channel," *UCSD-ITA*, Feb. 2006 [Online]. Available: http://ita.ucsd.edu/workshop/06/talks/papers/276.pdf
[16] G. Kramer, "Review of rate regions for interference channels," in *Proc. Int. Zurich Seminar*, Zurich, Switzerland, Feb. 2006, pp. 162–165.
[17] K. Kobayashi and T. S. Han, "A further consideration of the HK and CMG regions for the interference channel," *UCSD-ITA*, Jan. 2007 [Online]. Available: http://ita.ucsd.edu/workshop/07/files/paper/paper%5f133.pdf

Constructions of Difference Systems of Sets and Disjoint Difference Families

Cui-Ling Fan, Jian-Guo Lei, and Yan-Xun Chang

***Abstract*—Difference systems of sets (DSSs) are combinatorial structures that are a generalization of cyclic difference sets and arise in connection with code synchronization. In this correspondence, we give some constructions of DSS from cyclic designs and get some infinite classes of optimal difference systems of sets.**

***Index Terms*—Code synchronization, cyclic design, cyclic difference family, difference system of sets (DSS), optical orthogonal code.**

I. INTRODUCTION

A difference system of sets (DSS) with parameters $(n, \tau_0, \tau_1, \ldots, \tau_{q-1}, \rho)$ is a collection of q disjoint subsets $Q_i \subseteq \{1, 2, \ldots, n\}$, $|Q_i| = \tau_i$, $0 \le i \le q-1$, such that the multiset

$$\{a - b \pmod n : a \in Q_i, b \in Q_j, 0 \le i, j \le q-1,\ i \ne j\} \quad (1)$$

contains every number i, $1 \le i \le n-1$, at least ρ times. A DSS is *perfect* if every number i, $1 \le i \le n-1$, is contained exactly ρ times in the multiset (1). A DSS is *regular* if all subsets Q_i are of the same size: $\tau_0 = \tau_1 = \cdots = \tau_{q-1} = m$. We use the notation (n, m, q, ρ) for a regular DSS on n points with q subsets of size m.

Difference systems of sets were introduced by Levenshtein [17] and were used for the construction of codes that allow for synchronization in the presence of errors. A q-ary code of length n is a subset of the set F_q^n of all vectors of length n over $F_q = \{0, 1, \ldots, q-1\}$. If q is a prime power, we usually identify F_q with a finite field of order q, in which case i $(1 \le i \le q-1)$ stands for the ith power of a primitive element. A *linear* q-ary code (q a prime power) is a linear subspace of F_q^n. If $x = x_1 \cdots x_n$, $y = y_1 \cdots y_n \in F_q^n$, and $0 \le i \le n-1$, the ith *joint* of x and y is defined as $T_i(x, y) = x_{i+1} \cdots x_n y_1 \cdots y_i$. In particular, $T_i(x, x)$ is a cyclic shift of x. The *comma-free index* $\rho = \rho(C)$ of a code $C \subseteq F_q^n$ is defined as

$$\rho = \min d(z, T_i(x, y))$$

where the minimum is taken over all $x, y, z \in C$ and all $i = 1, \ldots, n-1$, and d is the Hamming distance between vectors in F_q^n. The comma-free index $\rho(C)$ allows one to distinguish a codeword from a joint of two codewords (and hence provides for synchronization of codewords), provided that at most $\lfloor \rho(C)/2 \rfloor$ errors have occurred in the given codeword [13].

Manuscript received September 8, 2007; revised January 30, 2008. The work of J.-G. Lei was supported by the National Natural Science Foundation of China (NSFC) under Grant 10571043. The work of Y.-X. Chang was supported by the National Natural Science Foundation of China (NSFC) under Grant 10771013.

C.-L. Fan is with the College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang, Hebei 050016, China (e-mail: clfan@yahoo.cn).

J.-G. Lei is with the College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang, Hebei 050016, China and also with the Institute of Mathematics, Beijing Jiaotong University, Beijing 100044, China (e-mail: leijg1964@yahoo.com.cn).

Y.-X. Chang is with the Institute of Mathematics, Beijing Jiaotong University, Beijing 100044, China (e-mail: yxchang@bjtu.edu.cn).

Communicated by I. Dumer, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2008.924673