

On the Capacity of Symmetric M-user Gaussian Interference Channels with Feedback

Lan V. Truong, *Member, IEEE*, Hirosuke Yamamoto, *Life Fellow, IEEE*

Abstract—A general time-varying feedback coding scheme is proposed for M -user fully connected symmetric Gaussian interference channels. Based on the analysis of the general coding scheme, we prove a theorem which gives a criterion for designing good time-varying feedback codes for Gaussian interference channels. The proposed scheme improves the Suh-Tse and Kramer inner bounds of the channel capacity for the cases of weak and not very strong interference when $M = 2$. This capacity improvement is more significant when the signal-to-noise ratio (SNR) is not very high. In addition, our coding scheme can be proved mathematically and numerically to outperform the Kramer code for $M \geq 2$ when the SNR is equal to the interference-to-noise ratio (INR). Besides, the generalized degrees-of-freedom (GDoF) of our proposed coding scheme can be proved to be optimal in the all network situations (very weak, weak, strong, very strong) for any M . The numerical results show that our coding scheme can attain better performance than the Suh-Tse coding scheme for $M = 2$ or the Mohajer-Tandon-Poor lattice coding scheme for $M > 2$. Furthermore, the simplicity of the encoding/decoding algorithms is another strong point of our proposed coding scheme compared with the Suh-Tse coding scheme when $M = 2$ and the Mohajer-Tandon-Poor lattice coding scheme when $M > 2$. More importantly, our results show that an optimal coding scheme for the symmetric Gaussian interference channels with feedback can be achieved by only using marginal posterior distributions under a better cooperation strategy between transmitters.

Index Terms—Gaussian Interference Channel with Feedback, Feedback, Posterior Matching, Iterated Function Systems.

I. INTRODUCTION

The interference channels (IC) were first studied by Ahlswede [1] in 1974, who established inner and outer bounds including the simultaneous decoding inner bound. Carleial [2] introduced the idea of rate splitting and established an inner bound using successive cancellation decoding and time-sharing. His inner bound was improved through simultaneous decoding and coded time sharing by Han and Kobayashi [3]. For the two-user Gaussian interference channel as a special case, there have been some significant progresses toward finding better inner and outer bounds although the capacity of this channel has been open for nearly 40 years. The approximation of the two-user Gaussian IC by a deterministic channel was first proposed by Bresler and Tse [4]. Furthermore, Etkin, Tse,

and Wang [5] proved that a very simple and explicit Han-Kobayashi type scheme can achieve the capacity for all values of the channel parameters within a single bit per second per hertz (bits/z/Hz).

Some other works have been done in the IC with feedback. Kramer developed a feedback strategy and derived an outer bound for the Gaussian channel. However, the gap between the outer bound and the inner bound becomes arbitrarily large with the increase of SNR (Signal-to-Noise Ratio) and INR (Interference-to-Noise Ratio) [6]. Jiang-Xin-Garg [7] found an achievable region in the discrete memoryless interference channel with feedback. However, their scheme employs three auxiliary random variables (requiring further optimization) and block Markov encoding (requiring a long block length). Suh and Tse [8], [9] characterized the capacity region within 2 bits/s/Hz and the symmetric capacity within 1 bit/s/Hz for the two-user Gaussian IC with feedback. They also indicated that feedback provides multiplicative gain at high SNR. However, their coding scheme does not work well when the SNR is close to the INR. It achieves even lower symmetric coding rate than the Kramer code when this condition happens. In addition, it has lower performance than the Kramer code when the $\alpha = \log \text{INR} / \log \text{SNR}$ is not very large and the SNR is low (c.f. Figs. 2-3 of this paper, or Fig. 14 in [9]). Recently, the Suh-Tse coding scheme has been extended to M -user Gaussian IC with feedback for $M \geq 3$ [10], [11] or the Gaussian IC with limited feedback [12]. The main ideas of these papers are to propose a method to manage the interference by turning the M -user Gaussian IC with feedback to an equivalent two-user one. Lattice codes, which are generally complicated in encoding and decoding, are used in these papers.

In this paper, we propose a new coding scheme based on the Kramer code [6] and the time-varying posterior matching code [13]–[16]. Our code can attain better coding rate by using a devised transmission cooperation strategy and decoding only their intended messages based on the fact that the posterior distributions can be measured online at all transmitters and their corresponding receivers. The proposed coding scheme has the following strong points.

- **Two-user case:** Our code can attain better symmetric coding rate than the Kramer code for all channel parameters since the Kramer code can be considered as a special case of our code and our code can be optimized more than the Kramer code for given channel parameters. Although the Kramer code cannot achieve the generalized degrees-of-freedom (GDoF) of this IC [8], our code can attain the same GDoF as the Suh-Tse code [8]. Furthermore, since

This work was supported in part by JSPS Grant-in-Aid for Scientific Research, No. 25289111. This paper was presented in part at the 2015 International Symposium on Information Theory.

L. Truong is with the Department of Computer Science, National University of Singapore, Singapore (e-mail: truongvl@comp.nus.edu.sg).

H. Yamamoto is with The University of Tokyo, Japan (e-mail: Hiro-suke@ieee.org).

Communicated by M. Costa, Associate Editor for Shannon Theory.

our code can achieve better performance than the Suh-Tse code [8] when $\alpha = \log INR / \log SNR$ is not very large (see Figs. 2 and 3 of this paper), our code overcomes all the weak-points of the Suh-Tse coding scheme and narrows the capacity gap to the Suh-Tse outer bound.

- **M -user case for $M \geq 3$:** Our code can achieve the GDoF of the M -user symmetric Gaussian IC with feedback. Some numerical results show that our coding scheme can attain better performance than the Mohajer-Tandon-Poor lattice coding scheme [11], which achieves the GDoF for very weak and strong interferences. The good performance of our code comes from the use of marginal posterior distributions under a devised cooperation strategy between transmitters. For the special case such that the SNR is equal to the INR, our code includes the Kramer code as a special case. But we note that the Kramer code cannot be constructed if the SNR is not equal to the INR for $M > 2$.

In Section II, we describe the notation used in this paper and the channel model of the Gaussian IC. We propose our coding scheme for the Gaussian IC and evaluate the decoding error probability of the proposed code in Section III. In Section IV, we evaluate the normalized covariance matrix of channel inputs generated by our coding scheme. Then, we derive the symmetric coding rate of our code and also treat several special cases in Section V. In Section VI, we show that our code can attain the GDoF of the Gaussian IC. Finally we show by numerical evaluations that our code can attain better symmetric coding rate than the Kramer code, the Suh-Tse code, and Mahajer-Tandon-Poor code.

II. CHANNEL MODEL AND PRELIMINARIES

A. Mathematical notations

Random variables and their realizations are denoted by upper-case letters and their corresponding lower-case letters, respectively. A real-valued random variable X is associated with a distribution $\mathbb{P}_X(\cdot)$ defined on the usual Borel σ -algebra over \mathbb{R} , and we write $X \sim \mathbb{P}_X$. The cumulative distribution function (c.d.f.) of X is given by $F_X(x) = \mathbb{P}_X((-\infty, x])$, and their inverse c.d.f is defined to be $F_X^{-1}(t) := \inf\{x : F_X(x) > t\}$. The uniform probability distribution over $(0, 1)$ is denoted through \mathcal{U} . Then, it is known that the following lemma holds.

Lemma 1 ([16, Lemma 1]). *Let X be a continuous random variable with $X \sim \mathbb{P}_X$ and Θ be a uniform distribution random variable which is statistically independent of X , i.e. $\Theta \sim \mathcal{U}$. Then $F_X^{-1}(\Theta) \sim \mathbb{P}_X$ and $F_X(X) \sim \mathcal{U}$.*

We also use the following notations: $\mathbf{Y}^{(n,m)} := (Y_1^{(m)}, Y_2^{(m)}, \dots, Y_n^{(m)})$, $\log x := \log_2(x)$, and $\exp_2(x) := 2^x$. Landau's symbols $O(\cdot)$ and $o(\cdot)$ are defined as follows.

$$f(n) = O(g(n))$$

if and only if there exists real positive constants N and C such that

$$|f(n)| \leq C|g(n)| \quad \text{for all } n > N.$$

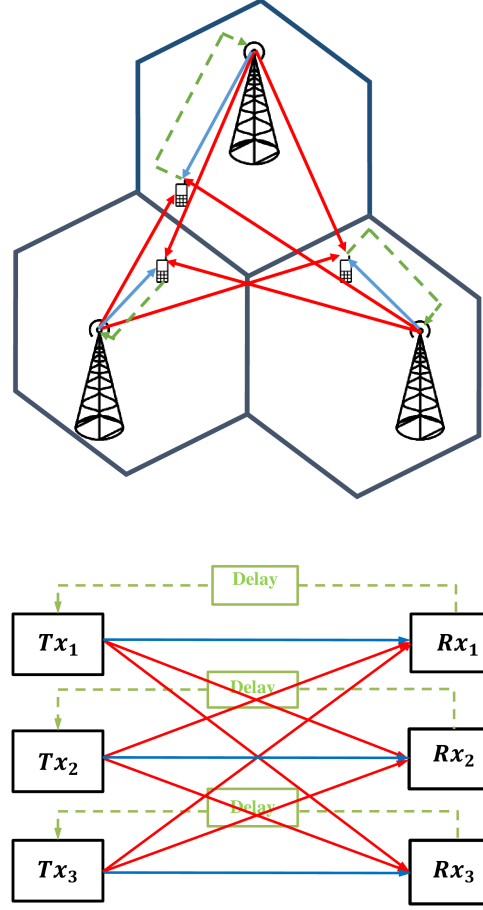


Fig. 1: Cellular network with base stations and three clients in (a), simplified and modeled as the network in (b).

Intuitively, this means that f does not grow faster than g .

$$f(n) = o(g(n))$$

if and only if there exists a real number N for any $C > 0$ such that $|f(n)| < C|g(n)|$ for all $n > N$. If $g(n) \neq 0$, this is equivalent to $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$.

A Hadamard matrix [17] of order M is an $(M \times M)$ matrix of $+1$ s and -1 s such that $\mathbf{H}\mathbf{H}^T = M\mathbf{I}_M$. In fact, it is not yet known for which values of M an \mathbf{H} exists. However, we know that if a Hadamard matrix of order M exists, then M is 1, 2, 4, or a multiple of 4. Moreover, for $M = 2^m$ where m a positive integer, we can construct \mathbf{H} by using the Sylvester method [19]. Besides, the Paley construction [17], which uses quadratic residues, can be used to construct Hadamard matrices of order M if $M = p + 1$ for a prime p and M is a multiple of 4.

Let α_n be the $[(n - 1 \bmod M) + 1]$ -th column of the Hadamard matrix \mathbf{H} . In this paper, we use column-permutation matrices of the Hadamard matrix \mathbf{H} , say $\mathbf{H}_n, n \in \mathbb{Z}^+$, which are defined as follows:

$$\mathbf{H}_n := \begin{bmatrix} \alpha_n & \alpha_{n+1} & \cdots & \alpha_{n+M-1} \end{bmatrix}. \quad (1)$$

B. Gaussian Interference Channel with Feedback

Consider a network with M pairs of transmitters/receivers shown in Fig. 1(b). Each transmitter $\mathbf{T}\mathbf{x}_m$ has a message $\Theta_m \sim \mathcal{U}(0, 1)$ that it wishes to send to its respective receiver $\mathbf{R}\mathbf{x}_m$. The signal transmitted by each transmitter is corrupted by the interfering signals sent by other transmitters, and received at the receiver. This can be mathematically modeled as

$$Y_n^{(m)} = X_n^{(m)} + a \sum_{k=1, k \neq m}^M X_n^{(k)} + Z_n^{(m)} \quad (2)$$

where $X_n^{(m)} \in \mathbb{R}$ is the transmitted symbol by sender m at time n ; $Y_n^{(m)} \in \mathbb{R}$ is the received signal by receiver m at the time n . We can assume without loss of generality $Z_n^{(m)} \sim \mathcal{N}(0, 1)$ and $a \geq 0$. We also assume that the output symbols are casually fed back to the corresponding senders and that the transmitted symbol $X_n^{(m)}$ at time n can depend on both the message Θ_m and the previous channel output sequences $\mathbf{Y}^{(n-1, m)} := (Y_1^{(m)}, Y_2^{(m)}, \dots, Y_{n-1}^{(m)})$, $\forall m \in \{1, 2, \dots, M\}$.

A *transmission scheme* for the M -user Gaussian interference channel with feedback is sequences of measurable functions $\{g_n^{(m)} : (0, 1) \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}\}_{n=1}^\infty, m \in \{1, 2, \dots, M\}$ so that the input to the channel generated by the transmitter is given by

$$X_n^{(m)} = g_n^{(m)}(\Theta_m, \mathbf{Y}^{(n-1, m)}). \quad (3)$$

A *decoding rule* for the M -user Gaussian interference channel with feedback are sequences of measurable mappings $\{\Delta_n^{(m)} : \mathbb{R}^n \rightarrow \mathcal{E}\}_{n=1}^\infty, m \in \{1, 2, \dots, M\}$ where \mathcal{E} is the set of all open intervals in $(0, 1)$ and $\Delta_n^{(m)}(\mathbf{y}^{(n, m)})$ is referred to as the decoded interval at receiver m . The error probabilities at time n associated with a transmission scheme and a decoding rule, is defined as

$$p_n^{(m)}(e) := \mathbb{P}(\Theta_m \notin \Delta_n^{(m)}(\mathbf{Y}^{(n, m)})), \quad \forall m = 1, 2, \dots, M, \quad (4)$$

and the corresponding coding rate vector $(R_n^{(1)}, R_n^{(2)}, \dots, R_n^{(M)})$ at time n is defined by

$$R_n^{(m)} := -\frac{1}{n} \log |\Delta_n^{(m)}(\mathbf{Y}^{(n, m)})|, \quad (5)$$

where $|\cdot|$ represents the length of an interval.

We say that a transmission scheme together with a decoding rule achieves a rate vector (R_1, R_2, \dots, R_M) over a Gaussian interference channel if for all $m \in \{1, 2, \dots, M\}$ we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(R_n^{(m)} < R_m) = 0, \quad (6)$$

$$\lim_{n \rightarrow \infty} p_n^{(m)}(e) = 0. \quad (7)$$

The rate vector is achieved within input power constraints $P^{(1)}, P^{(2)}, \dots, P^{(M)}$ if the following is satisfied:

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}[(X_k^{(m)})^2] \leq P^{(m)}, \quad \forall m = 1, 2, \dots, M. \quad (8)$$

We denote the set of all achievable rate tuples (R_1, R_2, \dots, R_M) by \mathcal{R} .

For the symmetric case [11], i.e. $P^{(1)} = P^{(2)} = \dots = P^{(M)} = P$ for some $P > 0$, let

$$SNR := P, \quad (9)$$

$$INR := a^2 P, \quad (10)$$

$$\alpha := \frac{\log INR}{\log SNR}, \quad (11)$$

and define the per-user generalized degrees of freedom for $R_m = R_m(SNR, \alpha)$ as

$$d(\alpha) = \frac{1}{M} \limsup_{SNR \rightarrow \infty} \frac{\max_{(R_1, R_2, \dots, R_M) \in \mathcal{R}} \sum_{m=1}^M R_m(SNR, \alpha)}{(1/2) \log(SNR)}. \quad (12)$$

If $R_1 = R_2 = \dots = R_M = R_{\text{sym}}$ we call R_{sym} a symmetric rate, and the symmetric capacity is defined by

$$C_{\text{sym}} := \sup\{R_{\text{sym}} : (R_{\text{sym}}, R_{\text{sym}}, \dots, R_{\text{sym}}) \in \mathcal{R}\}. \quad (13)$$

The per-user generalized degrees of freedom in (12) can be written as

$$d(\alpha) = \limsup_{SNR \rightarrow \infty} \frac{C_{\text{sym}}}{(1/2) \log(SNR)}. \quad (14)$$

An *optimal fixed rate* decoding rule for the M -user Gaussian interference channel with feedback for rate vector (R_1, R_2, \dots, R_M) is the one that decodes a vector of fixed length intervals $\{(J_1, J_2, \dots, J_M) : |J_m| = 2^{-nR_m} \text{ for } m \in \{1, 2, \dots, M\}\}$, which maximizes posteriori probabilities, i.e.,

$$\Delta_n^{(m)}(\mathbf{y}^{(n, m)}) = \underset{J_m \in \mathcal{E} : |J_m| = 2^{-nR_m}}{\operatorname{argmax}} \mathbb{P}_{\Theta_m | \mathbf{Y}^n}(J_m | \mathbf{y}^{(n, m)}). \quad (15)$$

It is easy to see that the optimal fixed rate decoding rule for the Gaussian interference channel with feedback is the traditional MAP, MMSE decoding rule.

An *optimal variable rate* decoding rule with target error probabilities $p_e^{(m)}(n) = \delta_n^{(m)}$ is the one that decodes a vector of minimal-length intervals (J_1, J_2, \dots, J_M) such that accumulated marginal posteriori probabilities exceeds corresponding targets, i.e.,

$$\Delta_n^{(m)}(\mathbf{y}^{(n, m)}) = \min_{J_m \in \mathcal{E} : \mathbb{P}_{\Theta_m | \mathbf{Y}^n}(J_m | \mathbf{y}^{(n, m)}) \geq 1 - \delta_n^{(m)}} |J_m|. \quad (16)$$

Both decoding rules use the marginal posterior distribution of the message point $\mathbb{P}_{\Theta_m | \mathbf{Y}^n}$ which can be calculated online at the transmitters and the receivers. Refer [13]–[15], and [16] for more details.

Lemma 2. *The achievability in the definition (6), (7), and (8) implies the achievability in the standard framework.*

Proof: See the detailed proof in papers [16], [13]. The idea is as follows. Assume that we use an encoding scheme $\{g_n^{(m)}\}_{m=1}^M$ and a decoding rule $\{\Delta_n^{(m)}\}_{m=1}^M$ to achieve a rate tuple $\{R_m^*\}_{m=1}^M$ under the assumption that all the transmitted messages Θ_m are mutually independent and uniformly distributed on $(0, 1)$ as in Section II-B. Then, we can show

the existence of M sequences $\Gamma_m = \{\theta_{i,n}^{(m)} \in (0, 1)\}_{i=1}^{\lfloor 2^{nR_m} \rfloor}$ of message point sets for $m = 1, 2, \dots, M$, where any two message points in Γ_m are separated from each other by at least $2^{-nR_m^*}$. In addition, a uniform achievability over $\{\Gamma_m\}_{m=1}^M$ is guaranteed, i.e., $\lim_{n \rightarrow \infty} \max_{\theta_n^{(m)} \in \Gamma_m} \mathbb{P}(\theta_n^{(m)} \notin \Delta_n^{(m)}(\mathbf{Y}^{(n,m)})) = 0$ for each $m \in \{1, 2, \dots, M\}$. By mapping message points $\{1, 2, \dots, \lfloor 2^{nR_m} \rfloor\}$ defined in the traditional way¹ to message points in Γ_m , each coding scheme in Section II-B is reduced to an equivalent coding scheme in the traditional setting as in [18]. The error probabilities of the associated scheme decay as $\sqrt{p_n^{(m)}(e)}$ [16, Proof of Lemma II.3]. ■

III. A GENERAL TIME-VARYING CODING SCHEME FOR THE SYMMETRIC GAUSSIAN INTERFERENCE CHANNEL WITH FEEDBACK

In this section, we propose a *time-varying encoding/decoding* scheme for the *symmetric* Gaussian interference channel with feedback. For this symmetric case, we assume that $P^{(1)} = P^{(2)} = \dots = P^{(M)} = P$ for some $P > 0$. The time-varying encoding scheme is as follows:

A. Encoding

- Step 1: Transmitter m sends $X_1^{(m)} = F_X^{-1}(\Theta_m)$, $m \in \{1, 2, \dots, M\}$, where $X \sim \mathcal{N}(0, P_1)$ for some $P_1 > 0$.
- Step $n + 1$, $n \geq 1$:

- All transmitters estimate

$$P_{n+1} = \frac{P_n}{\beta_n^2} \left([1 - b_n(1 - a)]^2 + ab_n\lambda_n[2(1 - a)b_n + Mab_n - 2] + \frac{b_n^2}{P_n} \right), \quad (17)$$

and each transmitter m sends

$$X_{n+1}^{(m)}\alpha_{n+1}^{(m)} := \frac{1}{\beta_n} (X_n^{(m)} - b_n\alpha_n^{(m)}Y_n^{(m)})\alpha_{n+1}^{(m)}, \quad (18)$$

where $\alpha_{n+1} := [\alpha_{n+1}^{(1)} \ \alpha_{n+1}^{(2)} \ \dots \ \alpha_{n+1}^{(M)}]^T$, $\alpha_{n+1}^{(m)} \in \{-1, 1\}$, is the $((n \bmod M) + 1)$ -th column of an $M \times M$ Hadamard matrix \mathbf{H} which is defined in Section II-A, and b_n and β_n are some real number sequences such that $0 < \limsup_{n \rightarrow \infty} \beta_n < 1$ and they are determined according to network situations.

- Receiver m receives

$$Y_{n+1}^{(m)} = X_{n+1}^{(m)}\alpha_{n+1}^{(m)} + a \sum_{k=1, k \neq m}^M \alpha_{n+1}^{(k)} X_{n+1}^{(k)} + Z_{n+1}^{(k)}, \quad (19)$$

and each receiver feedbacks the received signal to the corresponding transmitter.

Here, $\{P_1, \beta_n, b_n\}$ must be chosen to satisfy the following power constraint:

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbb{E}[(X_n^{(m)})^2] \leq P, \quad \forall m \in \{1, 2, \dots, M\}. \quad (20)$$

Remark 1. The outperform of time-varying codes over the Kramer codes [6] can be explained by our better choice of parameter triplet (P_1, β_n, b_n) for each network situation. We note that each triplet of parameter (P_1, β_n, b_n) represents a cooperation strategy. With better cooperation among all transmitters, we can achieve larger achievable rate region. Besides, the use of (real) Hadamard matrix as coefficient matrix allows us to analyze and optimize the performance easier than the use of (complex) DFT matrix for the same purpose as in [6]. (Refer to Corollary 3 in Section V-B for more details.)

B. Decoding

- At each time slot n , receiver $m \in \{1, 2, \dots, M\}$ selects a fixed interval $J_1^{(m)} = (s_m, t_m)$ as the decoded interval with respect to $X_n^{(m)}$. Here,

$$t_m := o(2^{n(\log(\beta + \varepsilon)^{-1} - R)}), \quad (21)$$

$$s_m := -t_m, \quad (22)$$

where $\beta := \limsup_{n \rightarrow \infty} \beta_n$, and R is any positive number such that $R < \log(\beta + \varepsilon)^{-1}$ for some $\varepsilon > 0$.

- Then, set $J_n^{(m)} = (T_{n-1}^{(m)}(s_m), T_{n-1}^{(m)}(t_m))$ as the decoded interval with respect to $X_1^{(m)}$, where

$$T_n^{(m)}(s) := w_1^{(m)} \circ w_2^{(m)} \circ \dots \circ w_n^{(m)}(s), \quad \forall s \in \mathbb{R} \quad (23)$$

and

$$w_n^{(m)}(s) := \beta_n s + b_n \alpha_n^{(m)} Y_n^{(m)}, \quad (24)$$

where \circ is the composition operation defined as $(f \circ g)(x) := f(g(x))$.

- The receiver m sets the decoded interval for the message Θ_m as

$$\Delta_n^{(m)}(\mathbf{Y}^{(n,m)}) = F_X(J_n^{(m)}), \quad (25)$$

where $X \sim \mathcal{N}(0, P_1)$.

We call this coding strategy the *Gaussian interference time-varying feedback coding strategy*, which is an *optimal variable rate* decoding rule with doubly exponential decay of targeted error probabilities (see the proof of the Theorem 1 in this paper).

¹In the traditional settings, we usually assume that message set at user m is $\{1, 2, \dots, 2^{nR_m}\}$, where R_m is an achievable rate of the code [18].

C. Analysis of decoding error probability

In this subsection, we evaluate the performance of the *Gaussian interference time-varying feedback coding strategy* defined in the above subsections.

Theorem 1. *Under the condition that $0 < \beta := \limsup_{n \rightarrow \infty} \beta_n < 1$ and $W := \sup_{n,m} \mathbb{E}[(X_n^{(m)})^2] < \infty$, the time-varying coding scheme for the symmetric Gaussian interference channel with feedback achieves the following symmetric rate:*

$$R_{\text{sym}} = \log \frac{1}{\beta} \quad (\text{bits/channel use}). \quad (26)$$

Proof: Define

$$f_X(x) := \frac{1}{\sqrt{2\pi P_1}} \exp\left(-\frac{x^2}{2P_1}\right), \quad (27)$$

$$K := \sup_{x \in \mathbb{R}} \{f_X(x)\} = \frac{1}{\sqrt{2\pi P_1}}. \quad (28)$$

Let $R_n^{(m)}$ be the instant rate to transmit the intended messages Θ_m to the receiver m . For any fixed symmetric rate R , we have

$$\begin{aligned} \mathbb{P}(R_n^{(m)} < R) &\stackrel{(a)}{=} \mathbb{P}\left(-\frac{1}{n} \log |\Delta_n^{(m)}(\mathbf{Y}^{(n,m)})| < R\right) \\ &= \mathbb{P}\left(|\Delta_n^{(m)}(\mathbf{Y}^{(n,m)})| > 2^{-nR}\right) \\ &\stackrel{(b)}{=} \mathbb{P}\left(\int_{J_n^{(m)}} f_X(x) dx > 2^{-nR}\right) \\ &\stackrel{(c)}{\leq} \mathbb{P}\left(|J_n^{(m)}| > 2^{-nR}/K\right), \end{aligned} \quad (29)$$

where (a) follows from (5), (b) follows from (25), and (c) follows from (28). In addition, it holds from (24) that for any $t, s \in \mathbb{R}$

$$|w_n^{(m)}(t) - w_n^{(m)}(s)| = \beta_n |t - s| \quad (30)$$

for all $m \in \{1, 2, \dots, M\}$ and $n = 1, 2, \dots$

Now, we will show that rate R is achievable if $R < R_{\text{sym}} = \log \beta^{-1}$. Since $R < R_{\text{sym}}$, we can find an $\epsilon > 0$ such that $R < \log(\beta + \epsilon)^{-1}$. Therefore, for n sufficiently large, we have

$$\begin{aligned} \mathbb{P}(R_n^{(m)} < R) &\stackrel{(a)}{\leq} K 2^{nR} \mathbb{E} \left[\left| w_1^{(m)} \circ w_2^{(m)} \cdots \circ w_{n-1}^{(m)}(t_m) \right. \right. \\ &\quad \left. \left. - w_1^{(m)} \circ w_2^{(m)} \cdots \circ w_{n-1}^{(m)}(s_m) \right| \right] \\ &\stackrel{(b)}{=} K 2^{nR} \mathbb{E} \left[\mathbb{E} \left(\left| w_1^{(m)} \circ w_2^{(m)} \cdots \circ w_{n-1}^{(m)}(t_m) \right. \right. \right. \\ &\quad \left. \left. - w_1^{(m)} \circ w_2^{(m)} \cdots \circ w_{n-1}^{(m)}(s_m) \right| \mathbf{Y}_2^{(n-1,m)} \right) \right] \\ &\stackrel{(c)}{=} K 2^{nR} \beta_1 \mathbb{E} \left[\mathbb{E} \left(\left| w_2^{(m)} \circ w_3^{(m)} \cdots \circ w_{n-1}^{(m)}(t_m) \right. \right. \right. \\ &\quad \left. \left. - w_2^{(m)} \circ w_3^{(m)} \cdots \circ w_{n-1}^{(m)}(s_m) \right| \mathbf{Y}_2^{(n-1,m)} \right) \right] \\ &\stackrel{(d)}{=} K 2^{nR} \beta_1 \mathbb{E} \left[\left| w_2^{(m)} \circ w_3^{(m)} \cdots \circ w_{n-1}^{(m)}(t_m) \right. \right. \end{aligned}$$

$$\begin{aligned} &\left. - w_2^{(m)} \circ w_3^{(m)} \cdots \circ w_{n-1}^{(m)}(s_m) \right| \Big] \\ &\vdots \\ &\stackrel{(e)}{=} K 2^{nR} \left(\prod_{i=1}^{n-1} \beta_i \right) \mathbb{E} \left[\left| w_n^{(m)}(t_m) - w_n^{(m)}(s_m) \right| \right] \\ &\stackrel{(f)}{=} K 2^{nR} \left(\prod_{i=1}^n \beta_i \right) |J_1^{(m)}| \\ &= K 2^{nR} (2^{\frac{1}{n} \sum_{i=1}^n \log \beta_i})^n |J_1^{(m)}| \\ &\stackrel{(g)}{\leq} K 2^{nR} 2^{n \log(\beta + \epsilon)} |J_1^{(m)}|, \end{aligned} \quad (31)$$

where (a) follows from the Markov's inequality and (29), (b) and (d) follow from the law of iterated expectations, (c) follows from (30) for each fixed $\mathbf{Y}_2^{(n,m)} = \mathbf{y}_2^{(n,m)}$, (e) follows from the recursive application of (b)-(d), (f) follows from (30), and (g) follows from $0 < \beta_n < \beta + \epsilon$ and Cesàro means for n sufficiently large.

From (21), (22), and (31), it is easy to see that $\mathbb{P}(R_n^{(m)} < R) \rightarrow 0$ since

$$|J_1^{(m)}| = 2t_m = o\left(2^{n(\log(\beta + \epsilon)^{-1} - R)}\right). \quad (32)$$

Now, from our encoding scheme, it is easy to see that

$$\mathbb{E}[X_n^{(m)}] = 0, \quad (33)$$

for all $m \in \{1, 2, \dots, M\}$ and $n = 1, 2, \dots$. On the other hand, observe that

$$\begin{aligned} X_n^{(m)} &\stackrel{(a)}{=} \beta_n X_{n+1}^{(m)} + b_n \alpha_n^{(m)} Y_n^{(m)} \\ &\stackrel{(b)}{=} w_n^{(m)}(X_{n+1}^{(m)}) \end{aligned} \quad (34)$$

for all $n \geq 1$ and $m \in \{1, 2, \dots, M\}$, where (a) and (b) follow from (18) and (24), respectively. Therefore, we have from (34) that

$$\begin{aligned} X_1^{(m)} &= w_1^{(m)}(X_2^{(m)}) \\ &= (w_1^{(m)} \circ w_2^{(m)})(X_3^{(m)}) \\ &\vdots \\ &= (w_1^{(m)} \circ w_2^{(m)} \circ w_3^{(m)} \cdots \circ w_{n-1}^{(m)})(X_n^{(m)}) \\ &= T_{n-1}^{(m)}(X_n^{(m)}), \end{aligned} \quad (35)$$

where the last equality in (35) follows from (23).

Let $W_n^{(m)} = \mathbb{E}[(X_n^{(m)})^2]$, and let $Q(x)$ as the well-known tail function of the standard normal distribution $\mathcal{N}(0, 1)$. Then,

from the Chernoff bound of this function, we obtain

$$\begin{aligned}
p_n^{(m)}(e) &= \mathbb{P}\left(\Theta_m \notin \Delta_n^{(m)}\left(\mathbf{Y}^{(n,m)}\right)\right) \\
&= \mathbb{P}\left(\Theta_m \notin F_X(J_n^{(m)})\right) \\
&= \mathbb{P}\left(X_1^{(m)} \notin J_n^{(m)}\right) \\
&\stackrel{(a)}{=} \mathbb{P}\left(T_{n-1}^{(m)}(X_n^{(m)}) \notin \left(T_{n-1}^{(m)}(s_m), T_{n-1}^{(m)}(t_m)\right)\right) \\
&= \mathbb{P}\left(X_n^{(m)} \notin J_1^{(m)}\right) \\
&\stackrel{(b)}{=} 2Q\left(\frac{|J_1^{(m)}|}{2\sqrt{W_n^{(m)}}}\right) \\
&\stackrel{(c)}{\leq} \exp\left(-\frac{|J_1^{(m)}|^2}{8W_n^{(m)}}\right), \tag{36}
\end{aligned}$$

where (a) follows from (35) and the definition of $J_n^{(m)}$, (b) follows from (33) and the fact that $J_1^{(m)}$ is symmetric and $s_m = -t_m$, and (c) follows from the Chernoff bound for the Q-function $0 < Q(x) \leq (1/2)\exp(-x^2/2), \forall x > 0$.

From $R < \log(\beta + \epsilon)^{-1} < R_{\text{sym}}$, we have that for each $m \in \{1, 2, \dots, M\}$, $|J_1^{(m)}| \rightarrow \infty$ as $n \rightarrow \infty$. Furthermore, from the assumption that $W_n^{(m)}$ is upper bounded by some W , we have

$$\frac{|J_1^{(m)}|^2}{8W_n^{(m)}} \geq \frac{|J_1^{(m)}|^2}{8W} \rightarrow \infty \tag{37}$$

as $n \rightarrow \infty$. Therefore, if $R < R_{\text{sym}}$, the error probabilities tend to zero as $-\log p_n^{(m)}(e) = o(2^{2n(\log(\beta + \epsilon)^{-1} - R)})$ from (32) and (36). Furthermore, since (P_1, b_n, β_n) is chosen to satisfy (20), the input power constraints are also satisfied. ■

Remark 2. Since we can estimate R_{sym} and know our desired rate R in advance, it is possible to choose ϵ appropriately. This means that the decoding algorithm can be implemented practically. However, there is a tradeoff between the transmission rate R (the possible values of ϵ) and the code length n . If we transmit at a rate R very close to R_{sym} , ϵ must be very small. As a result, the required N_ϵ may become very large. Furthermore, the fact that $\log(\beta + \epsilon)^{-1}$ is very close to R also implies that the error probabilities decay slowly to zero. Therefore, the code length n must be very large if we transmit nearly at R_{sym} . On the contrary, quite large ϵ makes the required N_ϵ smaller and the decay of error probabilities faster.

Remark 3. In the case of finite n , $\mathbb{P}(R_n^{(m)} < R)$ is not zero even if $J_1^{(m)}$ satisfies (32). But this does not worsen the error probability $p_n^{(m)}(e)$ if retransmission is allowed. Note that since each encoder m obtains $\mathbf{y}^{(n,m)}$ via the feedback channel, both encoder m and decoder m can know the value of $R_n^{(m)}$ for $\mathbf{y}^{(n,m)}$. Hence, they can know whether event $\{R_n^{(m)} < R\}$ occurred or not for received $\mathbf{y}^{(n,m)}$. If event $\{R_n^{(m)} < R\}$ occurs, they discard this transmission and resend the same message Θ_m . This retransmission decreases the coding rate of message Θ_m from $R_n^{(m)}$ to $R_n^{(m)}(1 - \mathbb{P}(R_n^{(m)} < R))$. But,

this degradation of coding rate is negligible if $\mathbb{P}(R_n^{(m)} < R)$ is sufficiently small.

Remark 4. If we cannot use the retransmission described in Remark 3, event $\{R_n^{(m)} < R\}$ makes a decoding error. In this case, we need to minimize the total decoding error probability given by $p_n^{(m)}(e) + \mathbb{P}(R_n^{(m)} < R)$, and hence we cannot attain double exponential order. By setting $|J_1^{(m)}|^2(\log e)/8W = n(\log(r^{(m)} + \epsilon)^{-1} - R)$ in (36), the error exponent of the total error probability is given by

$$\begin{aligned}
&\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \left[-\frac{1}{n} \log \left(p_n^{(m)}(e) + \mathbb{P}(R_n^{(m)} < R) \right) \right] \\
&\geq \lim_{\epsilon \rightarrow 0} \left(\log(r^{(m)} + \epsilon)^{-1} - R \right) \\
&= R_{\text{sym}} - R. \tag{38}
\end{aligned}$$

IV. NORMALIZED COVARIANCE MATRIX OF CHANNEL INPUTS FOR THE PROPOSED SCHEME

Firstly, we show the following propositions.

Proposition 1. For $\mathbb{E}[(X_n^{(1)})^2] = \mathbb{E}[(X_n^{(2)})^2] = \dots = \mathbb{E}[(X_n^{(M)})^2] := P_n$, define a normalized covariance matrix by

$$\begin{aligned}
\mathbf{R}_n &:= \frac{1}{P_n} \begin{bmatrix} \mathbb{E}[X_n^{(1)} X_n^{(1)}] & \dots & \mathbb{E}[X_n^{(1)} X_n^{(M)}] \\ \mathbb{E}[X_n^{(2)} X_n^{(1)}] & \dots & \mathbb{E}[X_n^{(2)} X_n^{(M)}] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[X_n^{(M)} X_n^{(1)}] & \dots & \mathbb{E}[X_n^{(M)} X_n^{(M)}] \end{bmatrix} \\
&= \begin{bmatrix} \rho_n^{(1,1)} & \dots & \rho_n^{(1,M)} \\ \rho_n^{(2,1)} & \dots & \rho_n^{(2,M)} \\ \vdots & \ddots & \vdots \\ \rho_n^{(M,1)} & \dots & \rho_n^{(M,M)} \end{bmatrix}, \tag{39}
\end{aligned}$$

where

$$\rho_n^{(m,k)} := \frac{\mathbb{E}[X_n^{(m)} X_n^{(k)}]}{P_n}, \tag{40}$$

$$\rho_n^{(m,k)} = \rho_n^{(k,m)}, \quad \rho_n^{(m,m)} = 1 \tag{41}$$

for all $m = 1, 2, \dots, M$ and $k = 1, 2, \dots, M$, then the following statement holds:

If the covariance matrix \mathbf{R}_n at time n has all the columns of the $M \times M$ Hadamard matrix as its eigenvectors, it follows that $\mathbb{E}[(X_{n+1}^{(1)})^2] = \mathbb{E}[(X_{n+1}^{(2)})^2] = \dots = \mathbb{E}[(X_{n+1}^{(M)})^2] := P_{n+1}$ and the covariance matrix \mathbf{R}_{n+1} at time $n+1$ also has all the columns of the $M \times M$ Hadamard matrix as its eigenvectors. In addition, (42) holds, where $\lambda_n^{(1)} = \lambda_n$, and $\lambda_{n+i}^{(k)}$ are the eigenvalue of the matrix \mathbf{R}_{n+i} associated with the eigenvector which is the $[(n+i+k-2 \bmod M) + 1]$ -th column of the Hadamard matrix \mathbf{H} for all $k = 1, 2, \dots, M$ and for all $i \in \{0, 1\}$.

Proof: The proof of Proposition 1 is given in Appendix A. ■

Remark 5. The use of modulated coefficients is a mathematical trick to force the covariance matrices \mathbf{R}_n among all

$$P_{n+1}\lambda_{n+1}^{(k)} = \begin{cases} \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \lambda_n^{(k+1)} + \frac{b_n^2}{P_n} \right), & k = 1, 2, \dots, M-1 \\ \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \lambda_n^{(1)} + \frac{b_n^2}{P_n} + ab_n \lambda_n^{(1)} [2(1-a)b_n + Mab_n - 2]M \right), & k = M \end{cases}, \quad (42)$$

transmitted signals to have a fixed set of eigenvectors at each time $n = 1, 2, \dots$. Thanks to this forcing mechanism, a relation between the eigenvalues of \mathbf{R}_n and \mathbf{R}_{n+1} (and/or \mathbf{R}_n and \mathbf{R}_1) can be established. This trick was first introduced by Ozarow and Leung for 2-user Gaussian MAC [19] and for 2-user Gaussian broadcast channel [20] in 1984. Kramer generalized this idea to design feedback codes based on Discrete Fourier Transform matrix (DFT) for M -user complex symbol IC channels [6] in 2002. Later, Truong and Yamamoto [14] combined these ideas with posterior matching idea [16] to design a feedback code for 2-user IC which outperforms Kramer code [6] and Suh-Tse code for 2-user IC real symbol channel [8], [9]. They also designed a feedback code which achieves the (optimal) linear feedback sum-capacity for the Gaussian broadcast channel [15].

Proposition 2. Every normalized covariance matrix \mathbf{R}_n has all the columns of the $M \times M$ Hadamard matrix as its eigenvectors.

Proof: Applying Lemma 1 with noting that $X \sim \mathcal{N}(0, P_1)$, we see that

$$\mathbb{E}[(X_1^{(1)})^2] = \mathbb{E}[F_X^{-1}(\Theta_1)F_X^{-1}(\Theta_1)] = P_1, \quad (43)$$

$$\mathbb{E}[(X_1^{(2)})^2] = \mathbb{E}[F_X^{-1}(\Theta_2)F_X^{-1}(\Theta_2)] = P_1, \quad (44)$$

⋮

$$\mathbb{E}[(X_1^{(M)})^2] = \mathbb{E}[F_X^{-1}(\Theta_M)F_X^{-1}(\Theta_M)] = P_1. \quad (45)$$

Hence, we have

$$\mathbb{E}[(X_1^{(1)})^2] = \mathbb{E}[(X_1^{(2)})^2] = \dots = \mathbb{E}[(X_1^{(M)})^2] = P_1. \quad (46)$$

Besides, since Θ_m and Θ_k are pairwise independent for $m \neq k$, we also have

$$\mathbb{E}[X_1^{(m)}X_1^{(k)}] = \mathbb{E}[F_X^{-1}(\Theta_m)F_X^{-1}(\Theta_k)] = 0, \quad \forall m \neq k. \quad (47)$$

It follows that

$$\mathbf{R}_1 = \frac{1}{P_1} \begin{bmatrix} \mathbb{E}[X_1^{(1)}X_1^{(1)}] & \dots & \mathbb{E}[X_1^{(1)}X_1^{(M)}] \\ \mathbb{E}[X_1^{(2)}X_1^{(1)}] & \dots & \mathbb{E}[X_1^{(2)}X_1^{(M)}] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[X_1^{(M)}X_1^{(1)}] & \dots & \mathbb{E}[X_1^{(M)}X_1^{(M)}] \end{bmatrix} \quad (48)$$

$$= \mathbf{I}_M, \quad (49)$$

where \mathbf{I}_M is the $M \times M$ identity matrix.

By using the induction arguments and the fact that the identity matrix \mathbf{I}_M has all the columns of the Hadamard matrix \mathbf{H} as its eigenvectors, together with the results of Proposition 1, we come to the conclusion. Note that we also have $\lambda_1^{(1)} = \lambda_1^{(2)} = \dots = \lambda_1^{(M)} = 1$. ■

Now, we show that some other well-known coding schemes are special variants of our coding strategy above.

A. Case of no interference ($a = 0$)

In this case, (163) and (165) in Appendix A, which is the proof of Proposition 1, become

$$P_{n+1} = \frac{P_n}{\beta_n^2} \left[(1 - b_n)^2 + \frac{b_n^2}{P_n} \right], \quad (50)$$

$$P_{n+1}\mathbf{R}_{n+1} = \frac{P_n}{\beta_n^2} \left[(1 - b_n)^2 \mathbf{R}_n + \frac{b_n^2}{P_n} \mathbf{I}_M \right]. \quad (51)$$

By setting the pair (P_n, b_n) as

$$P_n = P, \quad (52)$$

$$b_n = \frac{P}{P+1}, \quad (53)$$

we obtain from (50) and (51) that

$$\beta_n = \frac{1}{\sqrt{P+1}}, \quad (54)$$

$$\mathbf{R}_{n+1} = \frac{1}{1+P} \mathbf{R}_n + \frac{P}{1+P} \mathbf{I}_M. \quad (55)$$

Since $\mathbf{R}_1 = \mathbf{I}_M$, we have $\mathbf{R}_n = \mathbf{I}_M$ for $n = 1, 2, \dots$

In the non-interference case, the Gaussian interference channel with feedback becomes M separate point-to-point Gaussian channels with feedback. Our coding algorithm with the parameters given by (52) and (53) coincides with Shayevitz and Feder's posterior matching scheme [16], [21] (or a variant of Schalkwijk-Kailath's scheme [22], [23]). It is well-known that this coding scheme achieves the capacity of the channel.

B. Case of two transmitter and two receivers ($M = 2$)

In the special case $M = 2$, denote $\rho_n := \rho_n^{(1,2)}$ for simplicity. It is easy to show that $\lambda_n = 1 + |\rho_n|$ and $\alpha_n^{(1)}\alpha_n^{(2)} = \text{sgn}(\rho_n)$. By substituting these relations into (164) in Appendix A, we have

$$\begin{aligned} P_{n+1}\rho_{n+1} &= \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \rho_n + ab_n(1 + |\rho_n|) \right. \\ &\quad \times [2(1-a)b_n + 2ab_n - 2]\text{sgn}(\rho_n) \Big) \\ &= \frac{P_n \text{sgn}(\rho_n)}{\beta_n^2} \left(|\rho_n| - 2b_n(|\rho_n| + a) \right. \\ &\quad \left. + b_n^2[|\rho_n|(1+a^2) + 2a] \right). \end{aligned} \quad (56)$$

On the other hand, we can show from (163) that

$$\begin{aligned} P_{n+1} &= \frac{1}{\beta_n^2} \left(P_n - 2P_nb_n[1 + a|\rho_n| \right. \\ &\quad \left. + b_n^2[1 + P_n + a^2P_n + 2a|\rho_n|P_n] \right). \end{aligned} \quad (57)$$

(56) and (57) coincide with the equations (6) and (7) in [14]. By setting

$$P_n = P, \quad (58)$$

$$b_n = \frac{P(1 + a|\rho_n|)}{P(1 + a^2 + 2a|\rho_n|) + 1} \quad (59)$$

into (57), we obtain

$$\beta_n = \sqrt{\frac{a^2 P(1 - |\rho_n|^2) + 1}{P(1 + a^2 + 2a|\rho_n|) + 1}}. \quad (60)$$

With this choice of parameters, we obtain a new code which is an optimized version of Kramer code [6, Sec. VI-B]. In Sections V-B and VI, we will show that this variant code outperforms the Kramer code for all channel parameters.

C. Case of SNR = INR ($a = 1$)

This special case has been considered in [6]. From (42), we have

$$P_{n+1}\lambda_{n+1}^{(k)} = \begin{cases} \frac{P_n}{\beta_n^2} \left[\lambda_n^{(k+1)} + \frac{b_n^2}{P_n} \right], & k < M, \\ \frac{P_n}{\beta_n^2} \left[\lambda_n^{(1)} + Mb_n \lambda_n^{(1)} (Mb_n - 2) + \frac{b_n^2}{P_n} \right], & k = M. \end{cases} \quad (61)$$

Choose

$$\beta_n = \sqrt{\frac{P\lambda_n(M - \lambda_n) + 1}{PM\lambda_n + 1}}, \quad (62)$$

$$b_n = \frac{P_n \lambda_n}{P_n M \lambda_n + 1}. \quad (63)$$

Then, from (163), we have $P_{n+1} = P_n$. Therefore, if we set $P_1 = P$ as Kramer code [6], then we have $P_n = P$ for $\forall n \in \mathbb{N}$, i.e., the input power constraint is satisfied. Besides, from the relation (61) we also have

$$\lambda_{n+1}^{(k)} = \begin{cases} \frac{(PM\lambda_n^{(1)} + 1)\lambda_n^{(k+1)} - P(1 - a_n)(\lambda_n^{(1)})^2}{P\lambda_n^{(1)}(M - \lambda_n^{(1)}) + 1}, & k < M, \\ \frac{(PM\lambda_n^{(1)} + 1)\lambda_n^{(1)} - P[1 + (M - 1)a_n](\lambda_n^{(1)})^2}{P\lambda_n^{(1)}(M - \lambda_n^{(1)}) + 1}, & k = M, \end{cases} \quad (64)$$

where

$$a_n = 1 + \frac{1}{PM\lambda_n^{(1)} + 1}. \quad (65)$$

The equation (64) coincides with the one given by (76) in [6]. It is shown in [6] that for large M , the sum-rate is approximately $(\log M)/2 + \log \log M$. This rate is about $\log \log M$ larger than the sum-rate capacity without feedback, which is $\log(1 + PM)/2 \approx (\log M)/2$ (cf. [6]).

V. MAIN RESULTS

Theorem 2. *The symmetric rate*

$$R_{\text{sym}} = \frac{1}{2} \log^+ \frac{1}{\beta^2} \quad (66)$$

is achievable if the following relations hold for a triplet (b, β, λ) .

$$\beta^2 = [1 - b(1 - a)]^2 + ab\lambda[2(1 - a)b + Mab - 2] + \frac{b^2}{P}, \quad (67)$$

$$0 < \lambda < M, \quad (68)$$

$$(A \neq 0, A \neq C, \frac{\lambda^{(k)} - \lambda^{(k+1)}}{A - C} > 0, \forall k < M) \\ \text{or } (A = C \neq 1, \lambda^{(k)} = \lambda, \forall k \leq M), \quad (69)$$

$$(1 - CA^{M-1})\lambda = B(A^{M-1} + A^{M-2} + \dots + A + 1), \quad (70)$$

where

$$\log^+(x) := \max\{\log_2 x, 0\}, \quad (71)$$

$$\lambda^{(1)} := \lambda, \quad (72)$$

$$\lambda^{(k+1)} = \frac{\lambda^{(k)} - B}{A}, \quad k = 1, 2, \dots, M - 1, \quad (73)$$

$$A = \frac{[1 - b(1 - a)]^2}{\beta^2}, \quad (74)$$

$$B = \frac{b^2}{P\beta^2}, \quad (75)$$

$$C = \frac{[1 - b(1 - a)]^2 + Mab[2(1 - a)b + Mab - 2]}{\beta^2}. \quad (76)$$

Proof: The proof of Theorem 2 is given in Appendix B. ■

A. Case of no interference ($a = 0$)

Corollary 1. *Under no interference ($a = 0$), the time-varying coding scheme can achieve the following symmetric rate*

$$R_{\text{sym}} = \frac{1}{2} \log(P + 1). \quad (77)$$

Proof: In the case of $a = 0$, we see from (74)–(76) and (67) that

$$A = C = \frac{(1 - b)^2}{\beta^2}, \quad B = \frac{b^2}{P\beta^2}, \quad (78)$$

and

$$\beta^2 = (1 - b)^2 + \frac{b^2}{P}. \quad (79)$$

From (177) in Appendix B, (78), (79), λ satisfies

$$\lambda = \frac{b^2}{P\beta^2 - P(1 - b)^2} = 1 \in (0, M). \quad (80)$$

Furthermore, from (73), (78), (79) we also have that

$$\lambda^{(k)} = 1, \quad \forall k = 1, 2, \dots, M. \quad (81)$$

It is easy to see from (79) that the minimum of β^2 is equal to $1/(P + 1)$ which is obtained by $b = P/(P + 1)$. In this case, we have that

$$A = C = \frac{1}{P + 1} \neq 1. \quad (82)$$

Therefore, all the conditions in Theorem 2 are satisfied, and accordingly, the achievable symmetric rate is given by

$$R_{\text{sym}} = \frac{1}{2} \log^+ \frac{1}{\beta^2} = \frac{1}{2} \log(P+1). \quad (83)$$

It is well known that the above R_{sym} is the symmetric capacity of this channel. ■

B. Case of two transmitter and two receivers ($M = 2$)

Corollary 2. For a given a and P , let

$$\xi(\rho, b) := \frac{P}{P + b^2[1 + P + a^2P + 2aP\rho] - 2bP[1 + a\rho]}. \quad (84)$$

Then, the non-degraded symmetric Gaussian interference channel ($a > 0$) can achieve the following symmetric rate R_{sym} (bits/channel use):

$$R_{\text{sym}} = \frac{1}{2} \max_{\rho \in [0, \rho_0], b \in \{b_1^*, b_2^*\}} \log^+ \xi(\rho, b), \quad (85)$$

where

$$0 < \rho_0 := \sqrt{\frac{a^2P^2 + P - \sqrt{P[2a^2P^2 + P]}}{a^2P^2}} < 1, \quad (86)$$

and

$$b_{1,2}^* = \frac{2P\rho + aP + aP\rho^2}{2aP + 2P\rho + 2a^2P\rho + \rho + 2aP\rho^2} \pm \frac{\sqrt{P^2a^2\rho^4 - 2\rho^2(a^2P^2 + P) + a^2P^2}}{2aP + 2P\rho + 2a^2P\rho + \rho + 2aP\rho^2}. \quad (87)$$

Proof: In the case of $M = 2$, the normalized correlation matrix is

$$\mathbf{R}_n = \begin{bmatrix} 1 & \rho_n \\ \rho_n & 1 \end{bmatrix}, \quad (88)$$

and 2×2 Hadamard matrix is

$$\mathbf{H} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (89)$$

It is easy to see that the two eigenvalues of \mathbf{R}_n associated with two columns of \mathbf{H} are $1 \pm \rho_n$. Therefore, the assumption that $\lambda_n \rightarrow \lambda$ is equivalent to the assumption that $|\rho_n| \rightarrow \rho \in [0, 1]$ for $\rho = \lambda - 1$.

From (67), we have

$$P = \frac{1}{\beta^2} [P - 2bP(1 + a\rho) + b^2(1 + P + a^2P + 2aP\rho)]. \quad (90)$$

On the other hand, it also holds from (67), (74), and (75) that

$$\frac{ab[2(1-a)b + Mab - 2]}{\beta^2} = \frac{1 - A - B}{\lambda}. \quad (91)$$

Therefore, we obtain from (74), (76), (91) that

$$C = A + \frac{2}{\lambda} [1 - (A + B)]. \quad (92)$$

By substituting $M = 2$ into (70), we obtain

$$(1 - CA)\lambda = B(A + 1). \quad (93)$$

This leads to

$$\begin{aligned} \lambda &= CA\lambda + B(A + 1) \\ &= A\lambda \left(A + \frac{2}{\lambda} [1 - (A + B)] \right) + B(A + 1) \\ &= A^2\lambda + 2A[1 - (A + B)] + B(A + 1) \\ &= A^2\lambda + 2A(1 - A) + B(1 - A). \end{aligned} \quad (94)$$

Equivalently,

$$\lambda(1 - A^2) = (1 - A)(2A + B). \quad (95)$$

The relation (95) holds if λ satisfies $\lambda(1 + A) = 2A + B$, which means that

$$(1 + \rho)(1 + A) = 2A + B \quad (96)$$

or

$$\rho(1 + A) = 2A + B - 1 - A = A + B - 1. \quad (97)$$

Then

$$\begin{aligned} -\rho &= 1 - A - B + \rho A \\ &= 1 - \frac{[1 - b(1 - a)]^2}{\beta^2} - \frac{b^2}{P\beta^2} + \rho \frac{[1 - b(1 - a)]^2}{\beta^2} \\ &= \frac{1}{\beta^2} \left(\beta^2 - [1 - b(1 - a)]^2 - \frac{b^2}{P} + \rho[1 - b(1 - a)]^2 \right) \\ &= \frac{1}{\beta^2} \left[\rho - 2b(\rho + a) + b^2(\rho(1 + a^2) + 2a) \right], \end{aligned} \quad (98)$$

where the last equality holds from (67). Note that we assume that $a > 0$ in II-B.

Combining (90) with (98), we come to an equation system with three unknowns (b, ρ, β) as follows:

$$\begin{aligned} P &= \frac{1}{\beta^2} \left[P - 2bP(1 + a\rho) + b^2(1 + P + a^2P + 2aP\rho) \right], \\ -\rho &= \frac{1}{\beta^2} \left[\rho - 2b(\rho + a) + b^2(\rho(1 + a^2) + 2a) \right]. \end{aligned} \quad (99)$$

By considering ρ as a variable, we obtain the following quadratic equation in b for each fixed choice of ρ :

$$\begin{aligned} b^2[2aP + 2P\rho + 2a^2P\rho + \rho + 2aP\rho^2] \\ - 2b[2P\rho + aP + Pa\rho^2] + 2P\rho = 0. \end{aligned} \quad (101)$$

The discriminant of this quadratic equation is given by

$$\Delta = P^2a^2\rho^4 - 2\rho^2(a^2P^2 + P) + a^2P^2 := f(\rho). \quad (102)$$

Since $f(0) = a^2P^2 > 0$ and $f(1) = -2P < 0$, there exists the minimum value $\rho_0 \in (0, 1)$ such that $f(\rho_0) = 0$. Furthermore, it can be shown that the value of ρ_0 satisfies

$$0 < \rho_0 = \sqrt{\frac{a^2P^2 + P - \sqrt{P[2a^2P^2 + P]}}{a^2P^2}} < 1. \quad (103)$$

On the other hand, the first derivative satisfies

$$f'(\rho) = 4P^2a^2(\rho^3 - \rho) - 4P\rho \leq 0 \quad (104)$$

for all $\rho \in [0, 1]$. This means $\Delta = f(\rho) \geq 0$ for all $\rho \in [0, \rho_0]$. For all these values of ρ , it can easily be shown that (101) has two positive solutions b_1^*, b_2^* described in Corollary 2. In short, the equation system (90) and (98) has at least one solution (b, β) for each fixed $\rho \in [0, \rho_0]$. From Theorem 2, the symmetric Gaussian interference channel with feedback can achieve the following rate

$$\begin{aligned} R_{\text{sym}} &= \frac{1}{2} \log^+ \left(\frac{1}{\min_{\rho \in [0, \rho_0]} \beta^2} \right) \\ &= \frac{1}{2} \max_{\rho \in [0, \rho_0], b \in \{b_1^*, b_2^*\}} \log^+ \xi(\rho, b), \end{aligned} \quad (105)$$

where $\xi(\rho, b)$ is defined in (84). ■

Corollary 3. For $M = 2$, the proposed time-varying code outperforms the Kramer code in [6] for all channel parameters.

Proof: A variant of Kramer code is formed by setting the triple (b, β, ρ) , a solution of equation system (99) and (100), as follows:

$$b = \frac{P(1 + a\rho)}{P(1 + a^2 + 2a\rho) + 1}, \quad (106)$$

$$\beta = \sqrt{\frac{a^2 P(1 - \rho^2) + 1}{P(1 + a^2 + 2a\rho) + 1}}, \quad (107)$$

and ρ is the unique solution in $(0, 1)$ of the equation:

$$\begin{aligned} 2a^3 P^2 \rho^4 + a^2 P \rho^3 - 4aP(a^2 P + 1)\rho^2 \\ - (2a^2 P + P + 2)\rho + 2aP(a^2 P + 1) = 0. \end{aligned} \quad (108)$$

(See also in [6], [8], [24]). Our proposed code can optimize ρ as shown in (105) while the Kramer code must use the special ρ given by the solution of (108). Therefore, the proposed code can outperform the Kramer code. ■

Remark 6. Numerical evaluations shown in Figs. 2 and 3 in Section VI affirm that the achievable rate of our coding scheme in Corollary 2 is not worse than existing codes [6], [9] for all channel parameters. These figures show that our coding scheme achieves better performance than Suh-Tse code [9] when $\alpha = \log INR / \log SNR$ is not very large. In addition, our code can obtain better symmetric rate than (or at least equal to) the Kramer code for all channel parameters, and therefore it overcomes all the weak-points of the Suh-Tse coding scheme and narrows the capacity gap to the Suh-Tse outer bound [9].

C. A variant of Kramer's code for $a = 1$

For $a = 1$, by considering $\lambda \in [0, M]$ as a variable in (67), we obtain the following quadratic equation in b :

$$\beta^2 = 1 + b\lambda(Mb - 2) + \frac{b^2}{P} := f_\lambda(b). \quad (109)$$

Without considering other constraints, the function $f_\lambda(b)$ has the derivative

$$f'_\lambda(b) = 2Mb\lambda - 2\lambda + \frac{2b}{P}. \quad (110)$$

Note that $f'_\lambda(b) = 0$ if $b = b^* := P\lambda / (MP\lambda + 1)$. For this $b = b^*$, we obtain

$$\beta^2 = f_\lambda(b^*) = \frac{P\lambda(M - \lambda) + 1}{PM\lambda + 1}, \quad (111)$$

and hence

$$\beta = \sqrt{\frac{P\lambda(M - \lambda) + 1}{PM\lambda + 1}}. \quad (112)$$

In this case, we have from (74)–(76) that

$$A = \frac{1}{\beta^2} = \frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1}, \quad (113)$$

$$B = \frac{b^2}{P\beta^2} = \frac{P\lambda^2}{[P\lambda(M - \lambda) + 1][PM\lambda + 1]}, \quad (114)$$

$$C = \frac{1 + Mb(Mb - 2)}{\beta^2} = \frac{1}{[P\lambda(M - \lambda) + 1][PM\lambda + 1]}. \quad (115)$$

Since $A \neq 1$, substituting (113), (114), and (115) into (70) we obtain

$$\begin{aligned} \lambda &= CA^{M-1}\lambda + B \frac{A^M - 1}{A - 1} \\ &= \frac{\lambda}{[P\lambda(M - \lambda) + 1][PM\lambda + 1]} \left(\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1} \right)^{M-1} \\ &\quad + \frac{P\lambda^2}{[P\lambda(M - \lambda) + 1][PM\lambda + 1]} \left(\frac{\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1}}{\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1}} \right)^M - 1 \\ &= \frac{\lambda}{[P\lambda(M - \lambda) + 1][PM\lambda + 1]} \left(\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1} \right)^{M-1} \\ &\quad + \frac{1}{PM\lambda + 1} \left\{ \left(\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1} \right)^M - 1 \right\}. \end{aligned} \quad (116)$$

Hence,

$$\begin{aligned} \lambda &= -\frac{1}{PM\lambda + 1} \\ &\quad + \left(\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1} \right)^M \frac{1}{(PM\lambda + 1)^2} (\lambda + PM\lambda + 1). \end{aligned} \quad (117)$$

It follows that

$$\lambda + \frac{1}{PM\lambda + 1} = \left(\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1} \right)^M \frac{\lambda + PM\lambda + 1}{(PM\lambda + 1)^2}. \quad (118)$$

Hence,

$$\left(\lambda + \frac{1}{PM\lambda + 1} \right) \left(\frac{P\lambda(M - \lambda) + 1}{PM\lambda + 1} \right)^M = \frac{\lambda + PM\lambda + 1}{(PM\lambda + 1)^2}. \quad (119)$$

Then, we have

$$\begin{aligned} \left(\lambda + \frac{1}{PM\lambda + 1} \right) \left(\frac{P\lambda(M - \lambda) + 1}{PM\lambda + 1} \right)^M \\ = \frac{\lambda}{(PM\lambda + 1)^2} + \frac{1}{PM\lambda + 1}. \end{aligned} \quad (120)$$

Rewrite this equation as

$$\lambda = (MP\lambda + 1)^2 \left\{ \left(\lambda + \frac{1}{PM\lambda + 1} \right) \left(\frac{P\lambda(M - \lambda) + 1}{PM\lambda + 1} \right)^M - \frac{1}{PM\lambda + 1} \right\}. \quad (121)$$

Observe that the equation (121) coincides with the equation (92) in the paper [6]². Besides, we have from (119) that

$$\frac{PM\lambda + \lambda + 1}{PM\lambda^2 + \lambda + 1} = (PM\lambda + 1) \left(\frac{PM(M - \lambda) + 1}{PM\lambda + 1} \right)^M, \quad (122)$$

or

$$\frac{PM\lambda^2 + \lambda^2 + \lambda}{PM\lambda^2 + \lambda + 1} = \lambda(PM\lambda + 1) \exp_2 \left[M \log \left(\frac{PM(M - \lambda) + 1}{PM\lambda + 1} \right) \right]. \quad (123)$$

Note that the equation (123) coincides with the equation (93) in [6]. Moreover, observe that

$$A = \frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1} > 1, \quad (124)$$

$$B = \frac{P\lambda}{PM\lambda + 1} > 0. \quad (125)$$

Hence,

$$\lambda^{(k+1)} = \frac{\lambda^{(k)} - B}{A} < \lambda^{(k)}, k = 1, 2, \dots, M. \quad (126)$$

This means that the condition (69) in Theorem 2 is also satisfied, and therefore the achievable symmetric rate is given by

$$R_{\text{sym}} = \frac{1}{2} \log^+ \frac{1}{\beta^2} = \frac{1}{2} \log \left[\frac{PM\lambda + 1}{P\lambda(M - \lambda) + 1} \right]. \quad (127)$$

For M sufficiently large, it is also shown in [6] that the sum-rate is about $\lambda/2$, which is approximately $(\log M/2) + \log \log M$. This sum-rate is about $\log \log M$ larger than the sum-rate capacity without feedback, which is $\log(1 + PM)/2 \approx (\log M)/2$.

D. Extension of the Kramer's code for $a \neq 1$

In the same way as Section V-C, we consider the minimum value of the following equation for each fixed value of a :

$$\beta^2 = [1 - b(1 - a)]^2 + ab\lambda[2(1 - a)b + Mab - 2] + \frac{b^2}{P} := g_\lambda(b). \quad (128)$$

The first derivative of $g_\lambda(b)$ is

$$g'_\lambda(b) = \left(2(1 - a)^2 + \frac{2}{P} + [4a(1 - a) + 2Ma^2]\lambda \right) b - (2a\lambda + 2(1 - a)). \quad (129)$$

If we do not care about other restrictions, the function $g_\lambda(b)$ attains the minimum value at

$$b^* = \frac{a\lambda + (1 - a)}{(1 - a)^2 + 1/P + [2a(1 - a) + Ma^2]\lambda}. \quad (130)$$

²There is a typo in equation (92) in [6].

The minimum value of β^2 in this case is

$$\beta^2 = g_\lambda(b^*). \quad (131)$$

However, for the case $a \neq 1$, it is not easy to show the existence of a $\lambda > 0$ such that $(b^*, \sqrt{g_\lambda(b^*)}, \lambda)$ satisfies all the restrictions in Theorem 2. In [6], the coding scheme for $a \neq 1$ and $M > 2$ was also not proposed. The main difficulty is the overwhelming computation which happens when $a \neq 1$. It is also known that this method of choosing parameters is suboptimal at least for $M = 2$ as mentioned in Section V-B (cf. also [8], [9]). More specifically, it is shown in [8] that the Kramer code for $M = 2$ does not achieve the optimal generalized degree of freedom of the interference channel with feedback. In the following subsections, we show that a judicious choice of parameters of the time-varying code can achieve the generalized degree of this channel not only for $M = 2$ but also for $M > 2$. Our time-varying code, which achieves the optimal generalized degree of freedom, is proposed for any value of $a \notin \{0, 1\}$ and for any $M \in \mathbb{Z}^+$ where the Hadamard matrix exists.

E. Generalized Degree of Freedom of the Time-Varying Coding Scheme

In the following, we will characterize the achievable symmetric rate as the solution of a quartic equation.

Theorem 3. For $a \notin \{0, 1\}$, the following symmetric rate R_{sym} (bits/channel use) is achievable for M -user symmetric Gaussian channel with feedback:

$$R_{\text{sym}} = \frac{1}{2} \log^+ \left[\frac{1}{\inf_{A>1} \beta^2(A, a, P)} \right], \quad (132)$$

where $\beta = \beta(A, a, P)$ is the smallest positive real number satisfying the following constraints:

$$Z_4\beta^4 + Z_3\beta^3 + Z_2\beta^2 + Z_1\beta + Z_0 = 0, \quad (133)$$

and

$$0 < \frac{\beta^2(1 - A) - b^2/P}{ab[2(1 - a)b + Mab - 2]} < M, \quad (134)$$

where

$$b = \frac{1 - \beta\sqrt{A}}{1 - a}, \quad (135)$$

$$Z_4 = Y_0 + \frac{AY_2}{(1 - a)^2} + \frac{a[(M - 2)a + 2]}{P(1 - a)^4} A^2, \quad (136)$$

$$Z_3 = -4(\sqrt{A})^3 \frac{a[(M - 2)a + 2]}{P(1 - a)^4} + \frac{2a}{P(1 - a)^3} (\sqrt{A})^3 - \frac{2Y_2\sqrt{A}}{(1 - a)^2} - \frac{Y_1\sqrt{A}}{1 - a}, \quad (137)$$

$$Z_2 = \frac{a[(M - 2)a + 2]}{P(1 - a)^4} 6A - \frac{2a}{P(1 - a)^3} 3A + \frac{Y_2}{(1 - a)^2} + \frac{Y_1}{(1 - a)}, \quad (138)$$

$$Z_1 = -\frac{a[(M - 2)a + 2]}{P(1 - a)^4} 4\sqrt{A} + \frac{2a}{P(1 - a)^3} 3\sqrt{A}, \quad (139)$$

$$Z_0 = \frac{a[(M - 2)a + 2]}{P(1 - a)^4} - \frac{2a}{P(1 - a)^3}, \quad (140)$$

$$Y_0 = -\frac{(A^M - 1)(A - 1)}{\left(-MA^{M-1} + \frac{A^M - 1}{A-1}\right)}, \quad (141)$$

$$Y_1 = \frac{M(A - 1)A^{M-1}}{\left(-MA^{M-1} + \frac{A^M - 1}{A-1}\right)} 2a, \quad (142)$$

$$Y_2 = -\frac{(A^M - 1)}{P\left(-MA^{M-1} + \frac{A^M - 1}{A-1}\right)} - \frac{(M(A - 1)A^{M-1})}{\left(-MA^{M-1} + \frac{A^M - 1}{A-1}\right)} a[(M - 2)a + 2]. \quad (143)$$

Proof: The proof of Theorem 3 is given in Appendix C. ■

From Theorem 3, we obtain the following corollaries.

Corollary 4. For $\alpha = \log INR / \log SNR > 1$, the generalized degree of freedom of the proposed coding scheme is given by

$$d(\alpha) = \frac{\alpha}{2}. \quad (144)$$

Corollary 5. For $\alpha = \log INR / \log SNR < 1$, the generalized degree of freedom of the proposed coding scheme is given by

$$d(\alpha) = 1 - \frac{\alpha}{2}. \quad (145)$$

The proofs of Corollaries 4 and 5 are given in Appendices D and E, respectively.

VI. NUMERICAL RESULTS

We have performed some numerical evaluations and obtains some results which affirm our mathematical arguments in this paper (cf. Figs. 2–6). For the case $M = 2$, in Figs. 2 and 3, we show some numerical results of achievable symmetric rate for our proposed scheme in comparison with Suh-Tse scheme [9], Kramer scheme [6], and Suh-Tse outer bound [9]. These figures show that our coding scheme achieves better performance than Suh-Tse code when $\alpha = \log INR / \log SNR$ is not very large. In addition, our code can obtain better symmetric rate than (or at least equal to) the Kramer code for all channel parameters, and therefore it overcomes all the weak-points of the Suh-Tse coding scheme and narrows the capacity gap to the Suh-Tse outer bound.

Fig. 4 draws the curve of the generalized degree of freedom of the fully-connected M -user Gaussian interference channel with feedback as a function of $\alpha = \log INR / \log SNR$ for the case $\alpha \neq 1$ for any $M \geq 2$. This curve shows that the generalized degree of freedom $d(\alpha)$ is linearly decreasing and increasing in $\alpha < 1$ and $\alpha > 1$, respectively. For $\alpha = 1$, the generalized degree of freedom of this channel is not well-defined as shown in [11, Theorem 1]. This curve was shown to be optimal in [11, Theorem 1] for general M or for the case $M = 2$ [9]. Since other coding schemes [9], [11] which achieve the optimal generalized degree of freedom for the Gaussian interference channel with feedback are based on “cooperative interference alignment”, our results provide an important conclusion that the simple strategy “treating other users as noise” also works well if interference channels

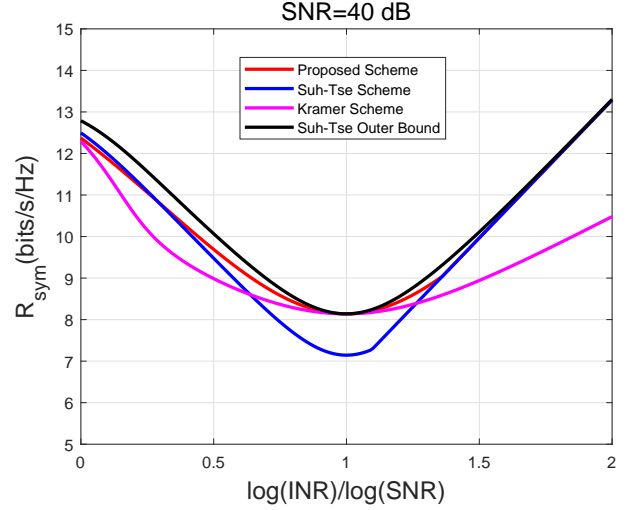


Fig. 2: Symmetric rate comparison at high SNR

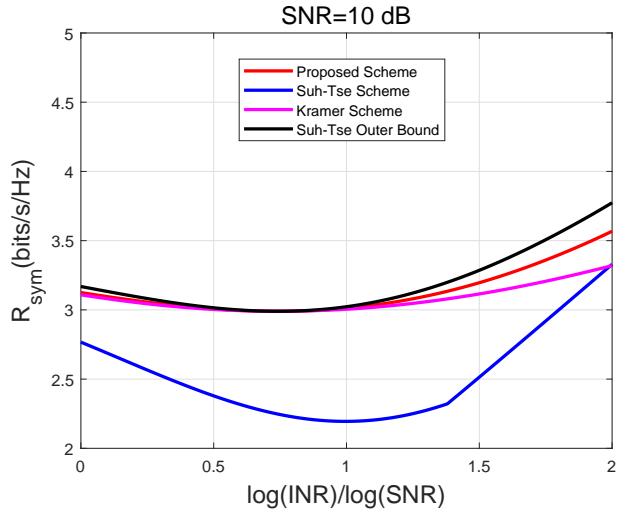


Fig. 3: Symmetric rate comparison at low SNR

allow feedback. Figs. 5 and 6 show that our coding scheme can even achieve better symmetric rate than the cooperative interference alignment strategy when numerically evaluated at some $M > 2$.

VII. CONCLUSION

A general time-varying posterior matching coding scheme for Gaussian interference channel with feedback was proposed. Based on the analysis of achievable symmetric rate of the channel, we proposed a coding scheme based on the ideas that a better cooperation strategy among transmitters will make the decoding process simpler and help to increase the achievable transmission symmetric rate. All receivers only need to decode their intended messages. Our proposed code has been shown to narrow down the gap to the Suh-Tse outer bound for the case $M = 2$. Besides, our code is optimal in the generalized degree of freedom sense for any $M \geq 2$. Our results show that the simple strategy “treating other users as noise” also works well if interference channels allow feedback. An interesting

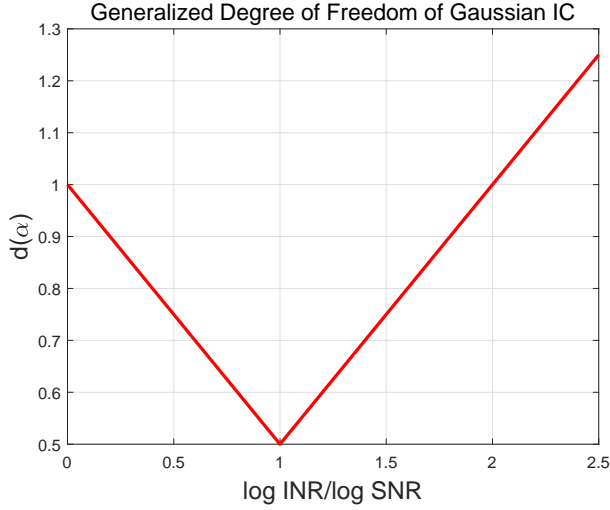


Fig. 4: Generalized degree of freedom for feedback Gaussian IC

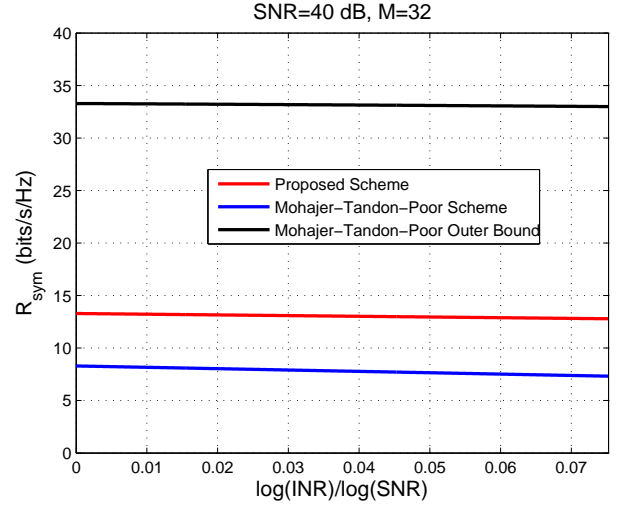


Fig. 6: Symmetric rate comparison for the very weak Gaussian IC at high SNR

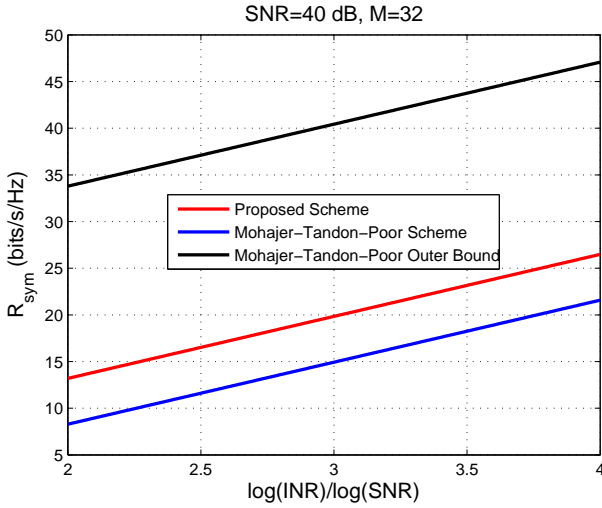


Fig. 5: Symmetric rate comparison for the very strong Gaussian IC at high SNR

future work is to find the way to characterize the achievable symmetric rate in a simpler way so that we can mathematically compare our code performance with other existing coding schemes.

APPENDIX A PROOF OF PROPOSITION 1

Define

$$\delta(n) = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases} \quad (146)$$

Firstly, observe that

$$Y_n^{(m)} = \alpha_n^{(m)} X_n^{(m)} + a \sum_{l=1, l \neq m}^M \alpha_n^{(l)} X_n^{(l)} + Z_n^{(m)} \quad (147)$$

$$= (1-a)\alpha_n^{(m)} X_n^{(m)} + a \sum_{l=1}^M \alpha_n^{(l)} X_n^{(l)} + Z_n^{(m)}. \quad (148)$$

By the transmission strategy, we have

$$X_{n+1}^{(m)} = \frac{1}{\beta_n} [X_n^{(m)} - b_n \alpha_n^{(m)} Y_n^{(m)}], \quad (149)$$

$$X_{n+1}^{(k)} = \frac{1}{\beta_n} [X_n^{(k)} - b_n \alpha_n^{(k)} Y_n^{(k)}]. \quad (150)$$

Hence, we obtain

$$\begin{aligned} \mathbb{E}[X_{n+1}^{(m)} X_{n+1}^{(k)}] &= \frac{1}{\beta_n^2} \left(\mathbb{E}[X_n^{(m)} X_n^{(k)}] - b_n \alpha_n^{(k)} \mathbb{E}[X_n^{(m)} Y_n^{(k)}] \right. \\ &\quad \left. - b_n \alpha_n^{(m)} \mathbb{E}[X_n^{(k)} Y_n^{(m)}] + b_n^2 \alpha_n^{(m)} \alpha_n^{(k)} \mathbb{E}[Y_n^{(m)} Y_n^{(k)}] \right). \end{aligned} \quad (151)$$

Observe that

$$\mathbb{E}[X_n^{(m)} Y_n^{(k)}] = (1-a)P_n \alpha_n^{(k)} \rho_n^{(m,k)} + P_n a \sum_{l=1}^M \alpha_n^{(l)} \rho_n^{(m,l)}, \quad (152)$$

$$\mathbb{E}[X_n^{(k)} Y_n^{(m)}] = (1-a)P_n \alpha_n^{(m)} \rho_n^{(k,m)} + P_n a \sum_{l=1}^M \alpha_n^{(l)} \rho_n^{(k,l)}, \quad (153)$$

$$\begin{aligned} \mathbb{E}[Y_n^{(m)} Y_n^{(k)}] &= \mathbb{E} \left([(1-a)\alpha_n^{(m)} X_n^{(m)} + a \sum_{l=1}^M \alpha_n^{(l)} X_n^{(l)} + Z_n^{(m)}] \right. \\ &\quad \left. \times [(1-a)\alpha_n^{(k)} X_n^{(k)} + a \sum_{t=1}^M \alpha_n^{(t)} X_n^{(t)} + Z_n^{(k)}] \right) \\ &= (1-a)^2 P_n \alpha_n^{(m)} \alpha_n^{(k)} \rho_n^{(m,k)} \\ &\quad + P_n a (1-a) \alpha_n^{(m)} \sum_{t=1}^M \alpha_n^{(t)} \rho_n^{(t,m)} \\ &\quad + P_n a (1-a) \alpha_n^{(k)} \sum_{l=1}^M \alpha_n^{(l)} \rho_n^{(l,k)} \\ &\quad + a^2 P_n \sum_{l=1}^M \sum_{t=1}^M \alpha_n^{(l)} \alpha_n^{(t)} \rho_n^{(l,t)} + \delta(m-k). \end{aligned} \quad (154)$$

Denote by λ_n the eigenvalue associated with the eigenvector α_n of the covariance matrix \mathbf{R}_n . Then, we have

$$\sum_{l=1}^M \sum_{t=1}^M \alpha_n^{(l)} \alpha_n^{(t)} \rho_n^{(l,t)} = \alpha_n^T \mathbf{R}_n \alpha_n = M \lambda_n, \quad (155)$$

$$\sum_{t=1}^M \alpha_n^{(t)} \rho_n^{(t,m)} = \alpha_n^T \rho_n^{(m)} = \alpha_n^{(m)} \lambda_n. \quad (156)$$

Therefore, using $(\alpha_n^{(m)})^2 = 1$,

$$\mathbb{E}[X_n^{(m)} Y_n^{(k)}] = (1-a) P_n \alpha_n^{(k)} \rho_n^{(m,k)} + P_n a \lambda_n \alpha_n^{(m)}, \quad (157)$$

$$\mathbb{E}[X_n^{(k)} Y_n^{(m)}] = (1-a) P_n \alpha_n^{(m)} \rho_n^{(k,m)} + P_n a \lambda_n \alpha_n^{(k)}, \quad (158)$$

$$\mathbb{E}[Y_n^{(m)} Y_n^{(k)}] = \delta(m-k) + (1-a)^2 P_n \alpha_n^{(m)} \alpha_n^{(k)} \rho_n^{(m,k)} + 2P_n a (1-a) \lambda_n + a^2 P M \lambda_n. \quad (159)$$

Substituting (157)–(159) into (151), we obtain

$$\begin{aligned} \mathbb{E}[X_{n+1}^{(m)} X_{n+1}^{(k)}] &= \frac{1}{\beta_n^2} \left(P_n \rho_n^{(m,k)} \right. \\ &\quad - b_n \alpha_n^{(k)} [(1-a) P_n \alpha_n^{(k)} \rho_n^{(m,k)} + P_n a \lambda_n \alpha_n^{(m)}] \\ &\quad - b_n \alpha_n^{(m)} [(1-a) P_n \alpha_n^{(m)} \rho_n^{(k,m)} + P_n a \lambda_n \alpha_n^{(k)}] \\ &\quad + b_n^2 \alpha_n^{(m)} \alpha_n^{(k)} [\delta(m-k) + 2P_n a (1-a) \lambda_n \\ &\quad \left. + a^2 P_n M \lambda_n + (1-a)^2 P_n \alpha_n^{(m)} \alpha_n^{(k)} \rho_n^{(m,k)}] \right) \\ &= \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \rho_n^{(m,k)} \right. \\ &\quad + ab_n \lambda_n [2(1-a)b_n + Mab_n - 2] \alpha_n^{(m)} \alpha_n^{(k)} \\ &\quad \left. + \frac{b_n^2}{P_n} \delta(m-k) \alpha_n^{(m)} \alpha_n^{(k)} \right). \end{aligned} \quad (160)$$

By setting $k = m$ in (160), we also obtain

$$\begin{aligned} \mathbb{E}[(X_{n+1}^{(m)})^2] &= \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \right. \\ &\quad \left. + ab_n \lambda_n [2(1-a)b_n + Mab_n - 2] + \frac{b_n^2}{P_n} \right), \end{aligned} \quad (161)$$

for all $k = 1, 2, \dots, M$. This means that

$$\mathbb{E}[(X_{n+1}^{(1)})^2] = \mathbb{E}[(X_{n+1}^{(2)})^2] = \dots = \mathbb{E}[(X_{n+1}^{(M)})^2] := P_{n+1}. \quad (162)$$

Therefore,

$$\begin{aligned} P_{n+1} &= \frac{P_n}{\beta_n^2} \left(\frac{b_n^2}{P_n} + [1 - b_n(1-a)]^2 \right. \\ &\quad \left. + ab_n \lambda_n [2(1-a)b_n + Mab_n - 2] \right) \end{aligned} \quad (163)$$

for all $m = 1, 2, \dots, M$.

On the other hand, from (40) and (160), we obtain

$$\begin{aligned} P_{n+1} \rho_{n+1}^{(m,k)} &= \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \rho_n^{(m,k)} \right. \\ &\quad + ab_n \lambda_n [2(1-a)b_n + Mab_n - 2] \alpha_n^{(m)} \alpha_n^{(k)} \\ &\quad \left. + \frac{b_n^2}{P_n} \delta(m-k) \alpha_n^{(m)} \alpha_n^{(k)} \right), \end{aligned} \quad (164)$$

which means from (39) that

$$\begin{aligned} P_{n+1} \mathbf{R}_{n+1} &= \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \mathbf{R}_n + \frac{b_n^2}{P_n} \mathbf{I}_M \right. \\ &\quad \left. + ab_n \lambda_n [2(1-a)b_n + Mab_n - 2] \alpha_n \alpha_n^T \right). \end{aligned} \quad (165)$$

Let α_n be the $[(n-1 \bmod M) + 1]$ -th column of the Hadamard matrix \mathbf{H} of order M , which is defined in Section II-A. Recall the definition of \mathbf{H}_n in (1), where $\mathbf{H}_n = [\alpha_n \ \alpha_{n+1} \ \dots \ \alpha_{n+M-1}]$.

Then from (165), we have

$$\begin{aligned} P_{n+1} \mathbf{H}_{n+1}^T \mathbf{R}_{n+1} \mathbf{H}_{n+1} &= \frac{P_n}{\beta_n^2} \left([1 - b_n(1-a)]^2 \mathbf{H}_{n+1}^T \mathbf{R}_n \mathbf{H}_{n+1} + \frac{b_n^2}{P_n} \mathbf{H}_{n+1}^T \mathbf{H}_{n+1} \right. \\ &\quad \left. + ab_n \lambda_n [2(1-a)b_n + Mab_n - 2] \mathbf{H}_{n+1}^T \alpha_n \alpha_n^T \mathbf{H}_{n+1} \right). \end{aligned} \quad (166)$$

Observe that

$$\begin{aligned} \alpha_n^T \mathbf{H}_{n+1} &= \alpha_n^T [\alpha_{n+1} \ \alpha_{n+2} \ \dots \ \alpha_{n+M-1} \ \alpha_{n+M}] \\ &= [0 \ 0 \ \dots \ 0 \ M], \end{aligned} \quad (167)$$

hence

$$\mathbf{H}_{n+1}^T \alpha_n \alpha_n^T \mathbf{H}_{n+1} = \text{diag}(0, 0, \dots, 0, M^2). \quad (168)$$

In addition, we also have

$$\mathbf{H}_{n+1}^T \mathbf{H}_{n+1} = M \mathbf{I}_M \quad (169)$$

by the definition of the Hadamard matrix. Now, since \mathbf{R}_n has all the columns of the Hadamard matrix as its eigenvectors, it also has all the columns of \mathbf{H}_{n+1} as its eigenvectors. Furthermore, since $\lambda_n^{(k)}$ is the eigenvalue of the matrix \mathbf{R}_n associated with the eigenvector which is the $[(n+k-2 \bmod M) + 1]$ -th column of the Hadamard matrix \mathbf{H} for all $k = 1, 2, \dots, M$, by the EVD (Eigenvalue Decomposition), we have

$$\begin{aligned} \mathbf{R}_n &= \mathbf{H}_{n+1} \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}) \mathbf{H}_{n+1}^{-1} \\ &= \frac{1}{M} \mathbf{H}_{n+1} \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}) \mathbf{H}_{n+1}^T, \end{aligned} \quad (170)$$

where (170) follows from (169). Furthermore, it follows from (170) that

$$\begin{aligned} \mathbf{H}_{n+1}^T \mathbf{R}_n \mathbf{H}_{n+1} &= \frac{1}{M} \mathbf{H}_{n+1}^T \mathbf{H}_{n+1} \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}) \mathbf{H}_{n+1}^T \mathbf{H}_{n+1} \\ &= \frac{1}{M} M \mathbf{I}_M \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}) M \mathbf{I}_M \\ &= M \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}). \end{aligned} \quad (171)$$

From (166), (168), (169), and (171), we have

$$\begin{aligned}
& P_{n+1} \mathbf{H}_{n+1}^T \mathbf{R}_{n+1} \mathbf{H}_{n+1} \\
&= \frac{P_n}{\beta_n^2} \left(\frac{b_n^2}{P_n} M \mathbf{I}_M \right. \\
&+ [1 - b_n(1 - a)]^2 M \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}) \\
&+ ab_n \lambda_n [2(1 - a)b_n + Mab_n - 2] \text{diag}(0, 0, \dots, 0, M^2) \Big). \tag{172}
\end{aligned}$$

As a result, the right hand side of the equation (172) is a diagonal matrix. In addition, from (172), we also have

$$\begin{aligned}
& P_{n+1} \mathbf{R}_{n+1} \mathbf{H}_{n+1} \\
&= \frac{1}{M} \mathbf{H}_{n+1} \frac{P_n}{\beta_n^2} \left([1 - b_n(1 - a)]^2 M \right. \\
&\quad \times \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}) + \frac{b_n^2}{P_n} M \mathbf{I}_M \\
&+ ab_n \lambda_n [2(1 - a)b_n + Mab_n - 2] \text{diag}(0, 0, \dots, 0, M^2) \Big) \\
&= \mathbf{H}_{n+1} \frac{P_n}{\beta_n^2} \left([1 - b_n(1 - a)]^2 \right. \\
&\quad \times \text{diag}(\lambda_n^{(2)}, \lambda_n^{(3)}, \dots, \lambda_n^{(M)}, \lambda_n^{(1)}) + \frac{b_n^2}{P_n} \mathbf{I}_M \\
&+ ab_n \lambda_n [2(1 - a)b_n + Mab_n - 2] \text{diag}(0, 0, \dots, 0, M) \Big). \tag{173}
\end{aligned}$$

From the relation (173), it is easy to see that all the columns of the matrix \mathbf{H}_{n+1} are eigenvectors of \mathbf{R}_{n+1} , so all the columns of the Hadamard matrix \mathbf{H} . Furthermore, since $\lambda_{n+1}^{(k)}$ is the eigenvalue of the matrix \mathbf{R}_{n+1} associated with the eigenvector which is the $[(n + k - 1 \bmod M) + 1]$ -th column of the Hadamard matrix \mathbf{H} for all $k = 1, 2, \dots, M$, it holds that $\mathbf{R}_{n+1} \mathbf{H}_{n+1} = \mathbf{H}_{n+1} \text{diag}(\lambda_{n+1}^{(1)}, \lambda_{n+1}^{(2)}, \dots, \lambda_{n+1}^{(M)})$. Hence, by (173), (42) of Proposition 1 holds.

APPENDIX B PROOF OF THEOREM 2

Firstly, it is easy to see that if we can force $P_n \rightarrow P$, $\lambda_n^{(k)} \rightarrow \lambda^{(k)}$, $b_n \rightarrow b$, $\beta_n \rightarrow \beta$ for some triplet (b, β, λ) , then from (42) we have

$$\lambda^{(k)} = \begin{cases} A\lambda^{(k+1)} + B, & k < M, \\ C\lambda^{(1)} + B, & k = M. \end{cases} \tag{174}$$

Therefore, we obtain

$$\begin{aligned}
\lambda &= \lambda^{(1)} \\
&= A\lambda^{(2)} + B \\
&= A[A\lambda^{(3)} + B] + B \\
&= A^2\lambda^{(3)} + AB + B \\
&\vdots \\
&= A^{M-1}\lambda^{(M)} + A^{M-2}B + A^{M-3}B + \dots + AB + B \\
&= CA^{M-1}\lambda + B(A^{M-1} + A^{M-2} + \dots + A + 1), \tag{175}
\end{aligned}$$

which means that

$$(1 - CA^{M-1})\lambda = B(A^{M-1} + A^{M-2} + \dots + A + 1), \tag{176}$$

and for the case $A \neq 1$,

$$\lambda = \frac{B(A^M - 1)}{(A - 1)(1 - CA^{M-1})}. \tag{177}$$

Moreover, the relation (67) holds from (163). We also note from (42) that all the other eigenvalues $\lambda^{(k)}$ satisfy the following relation:

$$\lambda^{(k+1)} = \frac{\lambda^{(k)} - B}{A}, \quad k = 1, 2, \dots, M - 1. \tag{178}$$

(Note that, $\lambda^{(1)} = \lambda$).

In the next part, we show a procedure to realize $P_M = P$ and $\lambda_M^{(k)} = \lambda^{(k)}$ for all $k = 1, 2, \dots, M$ by judiciously varying the values of P_k, b_k, β_k for all $k = 1, 2, \dots, M - 1$.

Define

$$A_n := \frac{P_n[1 - b_n(1 - a)]^2}{P_{n+1}\beta_n^2}, \tag{179}$$

$$B_n := \frac{b_n^2}{P_{n+1}\beta_n^2}, \tag{180}$$

$$C_n := \frac{P_n([1 - b_n(1 - a)]^2 + Mab_n[2(1 - a)b_n + Mab_n - 2])}{P_{n+1}\beta_n^2}. \tag{181}$$

From these definitions and (42) we have

$$\lambda_{n+1}^{(k)} = \begin{cases} A_n \lambda_n^{(k+1)} + B_n, & k < M, \\ C_n \lambda_n^{(1)} + B_n, & k = M. \end{cases} \tag{182}$$

Using the relation (182) recursively for $k = 1, 2, \dots, M - 1$, we obtain that

$$\begin{aligned}
\lambda_M^{(k)} &= A_{M-1} \lambda_{M-1}^{(k+1)} + B_{M-1} \\
&= A_{M-1} A_{M-2} \lambda_{M-2}^{(k+2)} \\
&\quad + A_{M-1} B_{M-2} + B_{M-1} \\
&\vdots \\
&= A_{M-1} A_{M-2} \dots A_k \lambda_k^{(M)} \\
&\quad + A_{M-1} A_{M-2} \dots A_{k+1} B_k \\
&\quad + A_{M-1} A_{M-2} \dots A_{k+2} B_{k+1} \\
&\quad + A_{M-1} B_{M-2} + B_{M-1} \\
&= A_{M-1} A_{M-2} \dots A_k C_{k-1} \lambda_{k-1}^{(1)} \\
&\quad + A_{M-1} A_{M-2} \dots A_k B_{k-1} \\
&\quad + A_{M-1} A_{M-2} \dots A_{k+1} B_k \\
&\quad + A_{M-1} A_{M-2} \dots A_{k+2} B_{k+1} \\
&\quad + A_{M-1} B_{M-2} + B_{M-1} \\
&= A_{M-1} A_{M-2} \dots A_k C_{k-1} A_{k-2} \lambda_{k-2}^{(2)} \\
&\quad + A_{M-1} A_{M-2} \dots A_k C_{k-1} B_{k-2} \\
&\quad + A_{M-1} A_{M-2} \dots A_k B_{k-1} \\
&\quad + A_{M-1} A_{M-2} \dots A_{k+1} B_k
\end{aligned}$$

$$\begin{aligned}
& + A_{M-1}A_{M-2} \dots A_{k+2}B_{k+1} \\
& + A_{M-1}B_{M-2} + B_{M-1} \\
& \vdots \\
& = A_{M-1}A_{M-2} \dots A_k C_{k-1} A_{k-2} \dots A_1 \lambda_1^{(k-1)} \\
& \quad + A_{M-1}A_{M-2} \dots A_k C_{k-1} A_{k-2} \dots A_2 B_1 \\
& \quad + A_{M-1}A_{M-2} \dots A_k C_{k-1} A_{k-2} \dots A_3 B_2 \\
& \quad + A_{M-1}A_{M-2} \dots A_k C_{k-1} B_{k-2} \\
& \quad + A_{M-1}A_{M-2} \dots A_k B_{k-1} \\
& \quad + A_{M-1}B_{M-2} + B_{M-1} \\
& \stackrel{(a)}{=} A_{M-1}A_{M-2} \dots A_k C_{k-1} A_{k-2} \dots A_1 \\
& \quad + A_{M-1}A_{M-2} \dots A_k C_{k-1} A_{k-2} \dots A_2 B_1 \\
& \quad + A_{M-1}A_{M-2} \dots A_k C_{k-1} A_{k-2} \dots A_3 B_2 \\
& \quad + A_{M-1}A_{M-2} \dots A_k C_{k-1} B_{k-2} \\
& \quad + A_{M-1}A_{M-2} \dots A_k B_{k-1} \\
& \quad + A_{M-1}B_{M-2} + B_{M-1} \tag{183}
\end{aligned}$$

for all $k = 1, 2, 3, \dots, M-1$. Here (a) follows from the fact that $\mathbf{R}_1 = \mathbf{I}_M$, so $\lambda_1^{(k)} = 1$, $k = 1, 2, \dots, M$.

We first fix A_n and C_n as $A_n = A$ and $C_n = C$. Then, we obtain the following relation from (183):

$$\begin{aligned}
\lambda_M^{(k)} &= A^{M-2}C + A^{M-3}CB_1 + A^{M-4}CB_2 + \dots \\
& \quad + A^{M-k}CB_{k-2} + A^{M-k}B_{k-1} \\
& \quad + A^{M-k-1}B_k + \dots + AB_{M-2} + B_{M-1} \\
&= [A^{M-2}C + A^{M-3}CB_1 + \dots \\
& \quad + A^{M-k}CB_{k-2} + A^{M-(k+1)}CB_{(k+1)-2}] \\
& \quad - A^{M-(k+1)}CB_{(k+1)-2} + A^{M-k}B_{k-1} \\
& \quad + [A^{M-(k+1)}B_k + \dots + AB_{M-2} + B_{M-1}] \\
&= \lambda_M^{(k+1)} - A^{M-k-1}CB_{k-1} + A^{M-k}B_{k-1}. \tag{184}
\end{aligned}$$

If $A = C \neq 1$, we have from (177) that

$$\lambda = \frac{B}{1-A}. \tag{185}$$

In addition, from (67), (75), (76), and $A = C$,

$$\begin{aligned}
A &= C \\
&= \frac{\beta^2 - (b^2/P)}{\beta^2} \\
&= 1 - \frac{b^2}{\beta^2 P} \\
&= 1 - B. \tag{186}
\end{aligned}$$

This leads to $\lambda = 1$. Note from (184) and $A = C$ that $\lambda_M^{(k)} = \lambda_M^{(k+1)}$, $k = 1, 2, \dots, M-1$. Hence, by setting $P_k = P$, $\beta_k = \beta$, $b_k = b$ for all $k = 1, 2, \dots, M-1$, we can achieve $\lambda_M^{(k)} = 1 = \lambda$ for all $k = 1, 2, \dots, M$ and $P_M = P$ (from the relations (163) and (67)).

If $A \neq 0$, $A \neq C$, we have from (180) and (184) that

$$\frac{b_{k-1}^2}{P_k \beta_{k-1}^2} = B_{k-1} = \frac{\lambda_M^{(k)} - \lambda_M^{(k+1)}}{(A-C)A^{M-k-1}}. \tag{187}$$

In order for (187) to have a solution pair $(\beta_{k-1}, P_k > 0)$, we need

$$\frac{\lambda^{(k)} - \lambda^{(k+1)}}{A-C} > 0. \tag{188}$$

Furthermore, from (179) and (181), the following relations must be satisfied

$$\frac{P_k[1 - b_k(1-a)]^2}{P_{k+1}\beta_k^2} = A, \tag{189}$$

and

$$\frac{P_k([1 - b_k(1-a)]^2 + Mab_k[2(1-a)b_k + Mab_k - 2])}{P_{k+1}\beta_k^2} = C. \tag{190}$$

From (187) and (189), we obtain

$$P_k = \frac{Ab_k^2}{B_k[1 - b_k(1-a)]^2}, \tag{191}$$

$$\beta_k^2 = \frac{b_k^2}{P_{k+1}B_k}, \tag{192}$$

for all $k = 1, 2, 3, \dots, M-1$. Since $P_k > 0$ obviously, we need $B_k > 0$ and $b_k \notin \{0, 1/(1-a)\}$ for all k . This condition is satisfied from (187) and (188).

Besides, we need to set $P_M = P$ and $\lambda_M^{(k)} = \lambda^{(k)}$ for all $k = 1, 2, \dots, M$.

The last thing we need to check is that there exists a $b_k \neq 0$ satisfying (190). From (74), (76), and (189), this condition is equivalent to that the following equation has at least a non-zero solution b_k :

$$\begin{aligned}
& \frac{P_k Mab_k[2(1-a)b_k + Mab_k - 2]^2}{P_{k+1}\beta_k^2} \\
&= \frac{PMab[2(1-a)b + Mab - 2]}{P\beta^2}. \tag{193}
\end{aligned}$$

Combing (193) with (189), the requirement becomes

$$\begin{aligned}
& \frac{A}{[1 - b_k(1-a)]^2} Mab_k[2(1-a)b_k + Mab_k - 2] \\
&= \frac{PMab[2(1-a)b + Mab - 2]}{P\beta^2}. \tag{194}
\end{aligned}$$

From (74), this relation is satisfied by choosing $b_k = b$ for all $k = 1, 2, 3, \dots, M-1$.

In short, for the case $A \neq 0$, $A \neq C$, $(\lambda^{(k)} - \lambda^{(k+1)})/(A-C) > 0$, $\forall k < M$, we can realize $P_M = P$ and $\lambda_M^{(k)} = \lambda^{(k)}$ for all $k = 1, 2, \dots, M$ by setting the parameters as follows:

$$b_k = b, \tag{195}$$

$$B_k = \frac{\lambda^{(k+1)} - \lambda^{(k+2)}}{(A-C)A^{M-k-2}}, \quad (\lambda^{(M+1)} := \lambda^{(1)}) \tag{196}$$

$$\beta_k^2 = \frac{b_k^2}{P_{k+1}B_k}, \tag{197}$$

$$P_k = \frac{AP_{k+1}\beta_k^2}{[1 - b_k(1-a)]^2}. \tag{198}$$

For $n \geq M$, we only need to set $b_n = b$, $\beta_n = \beta$, $P_n = P$ and obtain $\lambda_n^{(k)} = \lambda^{(k)}$ for all $k = 1, 2, \dots, M$ from the relation (42).

Observe that since $W_n^{(m)} = P$ for all $n \geq M$, we have $W = P < \infty$. Applying Theorem 1 to the above results, Theorem 2 is obtained. Note that, since Theorem 1 holds only for $0 < \beta < 1$, \log^+ must be used instead of \log .

APPENDIX C PROOF OF THEOREM 3

We will show that we can find a pair (b, β, λ) such that all the conditions in Theorem 2 are satisfied for each fixed $A > 1$. From (73), we have

$$\lambda^{(k+1)} - \lambda^{(k)} = \frac{1}{A} (\lambda^{(k)} - \lambda^{(k-1)}). \quad (199)$$

Therefore, the condition $\lambda^{(1)} > \lambda^{(2)} > \dots > \lambda^{(2)} > \lambda^{(M)}$ is satisfied if $\lambda^{(1)} > \lambda^{(2)}$. This condition is equivalent to

$$\lambda > \frac{\lambda - B}{A}. \quad (200)$$

Of course, this equation is satisfied if we choose $A > 1$. Moreover, we obtain from (74) that

$$b = \frac{1 \pm \beta\sqrt{A}}{1 - a}. \quad (201)$$

On the other hand, from (67), (74), (75), and (185) we obtain

$$C = A + \frac{M}{\lambda}(1 - A - B). \quad (202)$$

From (202), $A > 1$, $B \geq 0$, and $\lambda > 0$ we have that $A - C > 0$. This means that the condition (69) is satisfied.

Substituting (202) into (70) and noting that $A \neq 1$, we obtain

$$\left(1 - \left[A + \frac{M}{\lambda}(1 - A - B)\right] A^{M-1}\right) \lambda = B \frac{A^M - 1}{A - 1}. \quad (203)$$

This equation is equivalent to

$$\begin{aligned} & \left(1 - \left[A + \frac{M}{\lambda}(1 - A)\right] A^{M-1}\right) \lambda \\ &= B \left[-MA^{M-1} + \frac{A^M - 1}{A - 1}\right], \end{aligned} \quad (204)$$

or

$$B = \frac{(1 - [A + \frac{M}{\lambda}(1 - A)] A^{M-1}) \lambda}{-MA^{M-1} + \frac{A^M - 1}{A - 1}}. \quad (205)$$

Observe that $B = b^2/(P\beta^2)$, so we have

$$\frac{b^2}{P} = \beta^2 \frac{(1 - [A + \frac{M}{\lambda}(1 - A)] A^{M-1}) \lambda}{-MA^{M-1} + \frac{A^M - 1}{A - 1}}. \quad (206)$$

Rewrite (206) as

$$\frac{b^2}{P} = \beta^2 \frac{1 - A^M}{-MA^{M-1} + \frac{A^M - 1}{A - 1}} \lambda + \frac{M(A - 1)A^{M-1}}{-MA^{M-1} + \frac{A^M - 1}{A - 1}} \beta^2. \quad (207)$$

Therefore, from (67) and (68) we have

$$\lambda = \frac{\beta^2(1 - A) - \frac{b^2}{P}}{ab[2(1 - a)b + Mab - 2]}. \quad (208)$$

Replacing the relation (208) to (207), we obtain

$$\begin{aligned} \frac{b^2}{P} &= \frac{M(A - 1)A^{M-1}}{-MA^{M-1} + \frac{A^M - 1}{A - 1}} \beta^2 \\ &+ \beta^2 \left(\frac{1 - A^M}{-MA^{M-1} + \frac{A^M - 1}{A - 1}} \right) \left(\frac{\beta^2(1 - A) - \frac{b^2}{P}}{ab[2(1 - a)b + Mab - 2]} \right). \end{aligned} \quad (209)$$

Rearranging this relation, we have

$$\begin{aligned} & \frac{b^2}{P} ab[2(1 - a)b + Mab - 2] \\ &= \frac{(A^M - 1)(A - 1)}{\left(-MA^{M-1} + \frac{A^M - 1}{A - 1}\right)} \beta^4 - \frac{b^2(1 - A^M)}{P \left(-MA^{M-1} + \frac{A^M - 1}{A - 1}\right)} \beta^2 \\ &+ \frac{M(A - 1)A^{M-1}}{\left(-MA^{M-1} + \frac{A^M - 1}{A - 1}\right)} \beta^2 ab[2(1 - a)b + Mab - 2]. \end{aligned} \quad (210)$$

Then, we have an equation

$$\begin{aligned} & a[(M - 2)a + 2] \frac{b^4}{P} - 2a \frac{b^3}{P} - \frac{(A^M - 1)}{P \left(-MA^{M-1} + \frac{A^M - 1}{A - 1}\right)} \beta^2 b^2 \\ &- \frac{M(A - 1)A^{M-1}}{\left(-MA^{M-1} + \frac{A^M - 1}{A - 1}\right)} a[(M - 2)a + 2] \beta^2 b^2 \\ &+ \frac{M(A - 1)A^{M-1}}{\left(-MA^{M-1} + \frac{A^M - 1}{A - 1}\right)} 2a \beta^2 b \\ &- \frac{(A^M - 1)(A - 1)}{\left(-MA^{M-1} + \frac{A^M - 1}{A - 1}\right)} \beta^4 = 0. \end{aligned} \quad (211)$$

Using Y_0 , Y_1 , Y_2 defined by (141), (142), and (143), respectively, the equation (211) can be rewritten as

$$a[(M - 2)a + 2] \frac{b^4}{P} - 2a \frac{b^3}{P} + Y_2 \beta^2 b^2 + Y_1 \beta^2 b + Y_0 \beta^4 = 0. \quad (212)$$

Note that

$$b = \frac{1 \pm \sqrt{A}\beta}{1 - a}, \quad (213)$$

$$b^2 = \frac{1}{(1 - a)^2} (1 + A\beta^2 \pm 2\sqrt{A}\beta), \quad (214)$$

$$b^3 = \frac{1}{(1 - a)^3} (1 \pm (\sqrt{A})^3 \beta^3 \pm 3\sqrt{A}\beta + 3A\beta^2), \quad (215)$$

$$b^4 = \frac{1}{(1 - a)^4} (1 \pm 4\sqrt{A}\beta + 6A\beta^2 \pm 4(\sqrt{A})^3 \beta^3 + A^2 \beta^4). \quad (216)$$

Substituting these results into the equation (212), we attain

$$Z_4 \beta^4 + Z_3 \beta^3 + Z_2 \beta^2 + Z_1 \beta + Z_0 = 0. \quad (217)$$

Here, Z_4 , Z_2 , Z_0 are given by (136), (138), (140), respectively, and

$$Z_3 = \pm 4(\sqrt{A})^3 \frac{a[(M-2)a+2]}{P(1-a)^4} \mp \frac{2a}{P(1-a)^3} (\sqrt{A})^3 \pm \frac{2Y_2\sqrt{A}}{(1-a)^2} \pm \frac{Y_1\sqrt{A}}{1-a}, \quad (218)$$

$$Z_1 = \pm \frac{a[(M-2)a+2]}{P(1-a)^4} 4\sqrt{A} \mp \frac{2a}{P(1-a)^3} 3\sqrt{A}. \quad (219)$$

Last but not least, for $0 < \lambda < M$, (134) must hold. Now, since $A > 1$, (134) holds if and only if

$$b[2(1-a)b + Mab - 2] < 0. \quad (220)$$

This is equivalent to

$$0 < b < \frac{2}{(M-2)a+2}. \quad (221)$$

Since we assume that $\beta > 0$, it is easy to see that we must choose

$$b = \frac{1 - \beta\sqrt{A}}{1 - a}. \quad (222)$$

Therefore, Z_3 and Z_1 are given by (137) and (139), respectively. Finally, from (222), it is easy to see that (221) holds if and only if we choose β such that

$$\frac{Ma}{[(M-2)a+2]\sqrt{A}} < \beta < \frac{1}{\sqrt{A}} \quad \text{for } a < 1, \quad (223)$$

$$\frac{1}{\sqrt{A}} < \beta < \frac{Ma}{[(M-2)a+2]\sqrt{A}} \quad \text{for } a > 1. \quad (224)$$

This means that there exists a triplet (b, β, λ) which satisfies (133) and (134) and that two these conditions are sufficient conditions for (67)–(70) to hold for any $A > 1$. By Theorem 2, we conclude that Theorem 3 also holds.

APPENDIX D PROOF OF COROLLARY 4

Let $A = P^v$ for some $v > 0$ which will be determined later. On the other hand, since

$$\alpha = \frac{\log INR}{\log SNR} = \frac{\log(a^2 P)}{\log(P)} \quad (225)$$

then $a^2 P = P^\alpha$ and $a = P^{(\alpha-1)/2}$. For P sufficiently large, by keeping the dominant terms in the nominator and

denominator of fractions of polynomials in P , we have

$$Y_0 = -\frac{(A^M - 1)(A - 1)}{\left(-MA^{M-1} + \frac{A^{M-1}}{A-1}\right)} \approx \frac{1}{M-1} P^{2v}, \quad (226)$$

$$Y_1 = \frac{M(A-1)A^{M-1}}{\left(-MA^{M-1} + \frac{A^{M-1}}{A-1}\right)} 2a \approx \frac{MP^{vM}}{(-M+1)P^{v(M-1)}} 2P^{(\alpha-1)/2} \approx \frac{-2M}{M-1} P^{(v+\frac{\alpha-1}{2})}, \quad (227)$$

$$Y_2 = -\left(\frac{A^M - 1}{P\left(-MA^{M-1} + \frac{A^{M-1}}{A-1}\right)} - \frac{M(A-1)A^{M-1}}{-MA^{M-1} + \frac{A^{M-1}}{A-1}} a[(M-2)a+2]\right) \approx \frac{P^{v-1}}{M-1} + \frac{M(M-2)}{M-1} P^{v+\alpha-1} \approx \frac{M(M-2)}{M-1} P^{v+\alpha-1}. \quad (228)$$

Using the same arguments as above, it follows that

$$Z_4 \approx (M-1)P^{2v}, \quad (229)$$

$$Z_3 \approx -2MP^{\frac{3v}{2}}, \quad (230)$$

$$Z_2 \approx \frac{M^2}{(M-1)} P^v, \quad (231)$$

$$Z_1 \approx -(4M-2)P^{\frac{v}{2}-\alpha}, \quad (232)$$

$$Z_0 \approx MP^{-\alpha}. \quad (233)$$

The equation (133) will be satisfied if for P sufficiently large ($P \rightarrow \infty$) we can show that for some $\beta > 0$,

$$(M-1)P^{2v}\beta^4 - 2MP^{\frac{3v}{2}}\beta^3 + \frac{M^2}{M-1}P^v\beta^2 - (4M-2)P^{\frac{v}{2}-\alpha}\beta + MP^{-\alpha} \approx 0. \quad (234)$$

Moreover, β must satisfy (224), which becomes

$$P^{-v/2} < \beta < \frac{Ma}{[(M-2)a+2]} P^{-v/2} \approx \frac{M}{M-2} P^{-v/2}. \quad (235)$$

From [11, Theorem 1], we know that $d(\alpha) \leq \alpha/2$ for $\alpha > 1$ hence β cannot decay faster than $\alpha/4$ as P tends to infinity. To show that our coding scheme can achieve the optimal generalized degree of freedom $\alpha/2$, we set $\beta = P^{-\alpha/4}\gamma$ for some $\gamma > 0$ which does not depend on P and α and show that all the conditions in the Theorem 3 are satisfied. With this setting, it follows that

$$\frac{\alpha}{2} - 2\log_P(\gamma) < v < \frac{\alpha}{2} + 2\log_P\left(\frac{M}{M-2}\right) - 2\log_P(\gamma). \quad (236)$$

For $P \rightarrow \infty$, we have $v \rightarrow \alpha/2$. Hence, we must set $v = \alpha/2$. We will show that we can find such a γ to satisfy the equation (234). Indeed,

$$MP^{-\alpha} + (M-1)P^{2v}P^{-\alpha}\gamma^4 - 2MP^{\frac{3v}{2}}P^{-3\alpha/4}\gamma^3 + \frac{M^2}{M-1}P^vP^{-\alpha/2}\gamma^2 - (4M-2)P^{\frac{v}{2}-\alpha}P^{-\alpha/2}\gamma \approx 0. \quad (237)$$

The above equation is equivalent to for P sufficiently large

$$MP^{-\alpha} + (M-1)\gamma^4 - 2M\gamma^3 + \frac{M^2}{M-1}\gamma^2 + (4M-2)P^{-3\alpha/4}\gamma \approx 0. \quad (238)$$

By choosing $\gamma = M/(M-1)$, the equation (238) is satisfied. Next, we check that with our choices of γ and $v = \alpha/2$, $\beta = P^{-\alpha/4}\gamma$, we have $0 \leq \lambda \leq M$ for P sufficiently large.

Indeed, for P sufficiently large, we see that

$$b = \frac{1 - \beta\sqrt{A}}{1 - a} \approx (\gamma - 1)P^{(1-\alpha)/2}. \quad (239)$$

Hence,

$$b^2 \approx (\gamma - 1)^2 P^{1-\alpha}. \quad (240)$$

Then, from (208) we have

$$\begin{aligned} \lambda &= \frac{\beta^2(1-A) - b^2/P}{(M-2)a^2b^2 + 2ab^2 - 2ab} \\ &\approx \frac{P^{-\alpha/2}\gamma^2(1 - P^{\alpha/2}) - (\gamma - 1)^2P^{-\alpha}}{(M-2)(\gamma - 1)^2P + 2(\gamma - 1)^2P^{1-\alpha/2} - 2(\gamma - 1)\sqrt{P}} \\ &\approx 0. \end{aligned} \quad (241)$$

This means that $\lambda \searrow 0$ as $P \nearrow \infty$.

By the result of Theorem 1, for P sufficiently large, the achievable symmetric rate $R_{\text{sym}} = R_{\text{sym}}(\text{SNR}, \alpha)$ is approximate to

$$\begin{aligned} R_{\text{sym}}(\text{SNR}, \alpha) &\geq \frac{1}{2} \log^+ \frac{1}{\beta^2} \\ &\approx \frac{\alpha}{4} \log P. \end{aligned} \quad (242)$$

Therefore, the generalized degree of freedom of our code satisfies

$$d(\alpha) = \lim_{\text{SNR} \rightarrow \infty} \frac{R_{\text{sym}}(\text{SNR}, \alpha)}{(1/2) \log(\text{SNR})} \geq \frac{\alpha}{2}. \quad (243)$$

Since we know from [11, Theorem 1] that $d(\alpha) \leq \alpha/2$ we have

$$d(\alpha) = \frac{\alpha}{2}. \quad (244)$$

APPENDIX E PROOF OF COROLLARY 5

Similarly to the proof in Appendix D, we set $A = P^v$ where $v > 0$ will be determined later. Using the approximation ar-

guments, in which we keep the dominant terms in nominators and denominators of fractional expressions, we obtain

$$Y_0 \approx \frac{1}{M-1}P^{2v}, \quad (245)$$

$$Y_1 \approx -\frac{2M}{M-1}P^{v+(\alpha-1)/2}, \quad (246)$$

$$Y_2 \approx \frac{M(M-2)}{M-1}P^{v+\alpha-1}. \quad (247)$$

It follows that

$$Z_0 \approx 0, \quad (248)$$

$$Z_1 \approx -2P^{(\alpha-1)/2+v/2-1}, \quad (249)$$

$$Z_2 \approx -\frac{2M}{M-1}P^{v+(\alpha-1)/2}, \quad (250)$$

$$Z_3 \approx \frac{2M}{M-1}P^{3v/2+(\alpha-1)/2}, \quad (251)$$

$$Z_4 \approx \frac{1}{M-1}P^{2v}. \quad (252)$$

The equation (133) will be satisfied for P sufficiently large if

$$\begin{aligned} &\frac{1}{M-1}P^{2v}\beta^4 + \frac{2M}{M-1}P^{3v/2+(\alpha-1)/2}\beta^3 \\ &- \frac{2M}{M-1}P^{v+(\alpha-1)/2}\beta^2 - 2P^{(\alpha-1)/2+v/2-1}\beta \approx 0. \end{aligned} \quad (253)$$

From the paper [11], we know that $d(\alpha) \leq 1 - \alpha/2$ for $\alpha < 1$ hence β cannot decay faster than $\alpha/4 - 1/2$ as P tends to infinity. To show that our coding scheme can achieve the optimal generalized degree of freedom $1 - \alpha/2$, we set $\beta = P^{\alpha/4-1/2}\gamma$ where $\gamma \neq 0$ does not depend on P and α and show all the conditions in the Theorem 3 to be satisfied. With this setting, the equation (253) becomes

$$\begin{aligned} &\frac{1}{M-1}P^{2v+\alpha-2}\gamma^4 + \frac{2M}{M-1}P^{3v/2+5\alpha/4-2}\gamma^3 \\ &- \frac{2M}{M-1}P^{v+\alpha-3/2}\gamma^2 - 2P^{3\alpha/4+v/2-2}\gamma \approx 0. \end{aligned} \quad (254)$$

We will return to this equality later by judiciously choosing $v > 0$. Now, we need to make $0 < \lambda < M$.

In order to satisfy $\lambda > 0$, from previous arguments, we need to set

$$b = \frac{1 - \beta\sqrt{A}}{1 - a}, \quad (255)$$

and

$$\frac{Ma}{[(M-2)a+2]\sqrt{A}} < \beta < \frac{1}{\sqrt{A}}. \quad (256)$$

Moreover, from (255) and the choices of A and β , we have

$$b \approx 1 - P^{v/2+\alpha/4-1/2}\gamma, \quad (257)$$

$$\begin{aligned} b^2 &\approx 1 + P^{v+\alpha/2-1}\gamma^2 - 2P^{v/2+\alpha/4-1/2}\gamma \\ &\approx 1 - 2P^{v/2+\alpha/4-1/2}\gamma, \end{aligned} \quad (258)$$

if choosing $v < 1 - \alpha/2$. With the choice of $v < 1 - \alpha/2$, it also follows that

$$\begin{aligned} b^2 - b &\approx P^{v+\alpha/2-1}\gamma^2 - P^{v/2+\alpha/4-1/2}\gamma \\ &\approx -P^{v/2+\alpha/4-1/2}\gamma. \end{aligned} \quad (259)$$

From (208), we have

$$\begin{aligned}\lambda &\approx \frac{\gamma^2 P^{\alpha/2-1}(1-P^v) - P^{-1}(1-2P^{v/2+\alpha/4-1})\gamma}{-2P^{(\alpha-1)/2}P^{v/2+\alpha/4-1/2}\gamma} \\ &\approx \frac{-\gamma^2 P^{\alpha/2-1+v}}{-2\gamma P^{v/2+3\alpha/4-1}} \\ &= \frac{\gamma}{2} P^{v/2-\alpha/4}.\end{aligned}\quad (260)$$

To make λ bounded in $(0, M)$ when $P \rightarrow \infty$ and $\lambda > 0$, we should choose $v = \alpha/2$ and $0 < \gamma < 2M$ such as $\gamma = M$.

Now, we check the relation (254) when setting $v = \alpha/2$. It is easy to see that the left hand side of (254) will be

$$\begin{aligned}\frac{1}{M-1}P^{2(\alpha-1)}\gamma^4 + \frac{2M}{M-1}P^{2(\alpha-1)}\gamma^3 \\ - \frac{2M}{M-1}P^{3/2(\alpha-1)}\gamma^2 - 2P^{2(\alpha-1)}\gamma \rightarrow 0\end{aligned}\quad (261)$$

as $P \rightarrow \infty$ since $\alpha < 1$.

This means that for P sufficiently large and $\alpha < 1$, by Theorem 2, the achievable symmetric rate is approximate to

$$\begin{aligned}R_{\text{sym}}(\text{SNR}, \alpha) &\geq \frac{1}{2} \log^+ \frac{1}{\beta^2} \\ &\approx \frac{1}{2} \left(1 - \frac{\alpha}{2}\right) \log P.\end{aligned}\quad (262)$$

Therefore, the generalized degree of freedom of our code is greater than or equal to

$$d(\alpha) = \lim_{\text{SNR} \rightarrow \infty} \frac{R_{\text{sym}}(\text{SNR}, \alpha)}{(1/2) \log(\text{SNR})} \geq 1 - \frac{\alpha}{2}.\quad (263)$$

Since we know from [11, Theorem 1] that $d(\alpha) \leq 1 - \alpha/2$ for $\alpha < 1$, we have

$$d(\alpha) = 1 - \frac{\alpha}{2}, \quad \alpha < 1.\quad (264)$$

Acknowledgements

The authors would like to thank the associate editor Prof. Max Costa and the anonymous reviewers for their excellent and detailed comments that helped to improve the presentation of the paper.

REFERENCES

- [1] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probability*, vol. 2, no. 5, pp. 805–814, 1974.
- [2] A. B. Carleial, "Interference channels," *IEEE Trans. on Inform. Th.*, vol. 24, no. 1, pp. 60–70, 1978.
- [3] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. on Inform. Th.*, vol. 27, no. 1, pp. 49–60, 1981.
- [4] G. Bresler and D. Tse, "The two-user gaussian interference channel: a deterministic view," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 333–354, 2008.
- [5] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. on Inform. Th.*, vol. 54, no. 12, pp. 5534–5562, 2008.
- [6] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. on Inform. Th.*, vol. 48, no. 1, pp. 1423–1438, Jan. 2002.
- [7] J. Jiang, Y. Xin, and H. K. Garg, "Empirical processes and typical sequences," in *41st Annual Conference on Information Sciences and Systems*, Baltimore, Maryland, 2007.

- [8] C. Suh and D. Tse, "Symmetric feedback capacity of the Gaussian interference channel to within one bit," in *Proc. of Intl. Symp. on Inform. Th.*, Seoul, Korea, 2009.
- [9] —, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. on Inform. Th.*, vol. 57, no. 5, pp. 2667–2685, 2011.
- [10] R. Tandon, S. Mohajer, and H. V. Poor, "On the symmetric feedback capacity of the K-user cyclic Z-interference channel," *IEEE Trans. on Inform. Th.*, vol. 59, no. 5, pp. 2713–2733, 2013.
- [11] S. Mohajer, R. Tandon, and H. V. Poor, "On the feedback capacity of the fully connected K-user interference channel," *IEEE Trans. on Inform. Th.*, vol. 59, no. 5, pp. 2863–2881, 2013.
- [12] M. Ashraphijuo, V. Aggarwal, and X. Wang, "On the symmetric K-user linear deterministic interference channels with limited feedback," in *52nd Annual Allerton Conference*, 2014, pp. 366–371.
- [13] L. V. Truong, "Posterior matching scheme for Gaussian multiple access channel with feedback," in *Proc. of IEEE Information Theory Workshop*, Tasmania, Australia, 2014, pp. 476–480.
- [14] L. V. Truong and H. Yamamoto, "On the capacity of symmetric Gaussian interference channels with feedback," in *Proc. of Intl. Symp. on Inform. Th.*, Hong Kong, China, 2015, pp. 201–205.
- [15] —, "Posterior matching for Gaussian broadcast channels with feedback," *IEICE Transactions on Fundamentals*, vol. E100-A, no. 5, pp. 1165–1178, May 2017.
- [16] O. Shayevitz and M. Feder, "Optimal feedback communication via posterior matching," *IEEE Trans. on Inform. Th.*, vol. 57, no. 3, pp. 1186–1222, 2011.
- [17] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall, Inc., 1995.
- [18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.
- [19] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. on Inform. Th.*, vol. 30, no. 4, pp. 623–629, 1984.
- [20] L. Ozarow and S. Leung, "An achievable region and outer bound for the Gaussian broadcast channel with feedback (corresp.)," *IEEE Trans. on Inform. Th.*, vol. 30, no. 4, pp. 667–671, 1984.
- [21] O. Shayevitz and M. Feder, "Communication with feedback via posterior matching," in *Proc. of Intl. Symp. on Inform. Th.*, Nice, France, 2007.
- [22] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback—Part I: No bandwidth constraint," *IEEE Trans. on Inform. Th.*, vol. 12, no. 2, pp. 172–182, 1966.
- [23] J. Schalkwijk, "A coding scheme for additive noise channels with feedback—Part II: Band-limited constraint," *IEEE Trans. on Inform. Th.*, vol. 12, no. 4, pp. 183–189, 1966.
- [24] G. Kramer, "Correction to "feedback strategies for white Gaussian interference networks", and a capacity theorem for Gaussian interference channels with feedback," *IEEE Trans. on Inform. Th.*, vol. 50, no. 6, pp. 1373–1374, Jun. 2004.

Lan V. Truong (S'12–M'15) received the B.S.E. degree in electronics and telecommunications from the Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam, in 2003, and the M.S.E. degree from the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA, in 2011, and the Ph.D. degree from the Department of Electrical and Computer Engineering, National University of Singapore (NUS), Singapore, in 2018. He was an Operation and Maintenance Engineer with MobiFone Telecommunications Corporation, Hanoi, for several years. He spent one year as a Research Assistant with the NSF Center for Science of Information and Department of Computer Science, Purdue University, in 2012. From 2013 to 2015, he was an Academic Lecturer with the Department of Information Technology Specialization, FPT University, Hanoi, Vietnam. Since 2018, he has been a Research Assistant/Post-Doctoral Research Fellow with the Department of Computer Science, School of Computing, NUS. His research interests include information theory, machine learning, and communications.

Hirosuke Yamamoto (S'77–M'80–SM'03–F'11–LF'18) was born in Wakayama, Japan, in 1952. He received the B.E. degree from Shizuoka University, Shizuoka, Japan, in 1975 and the M.E. and Ph.D. degrees from the University of Tokyo, Tokyo, Japan, in 1977 and 1980, respectively, all in electrical engineering. In 1980, he joined Tokushima University. He was an Associate Professor at Tokushima University from 1983 to 1987, the University of Electro-Communications from 1987 to 1993, the University of Tokyo from 1993 to 1999, and a Professor at the University of Tokyo from 1999 to 2018. Since 2018, he has been a Professor Emeritus at the University of Tokyo and an Institute Professor at Research and Development Institute, Chuo University. In 1989–1990, he was a Visiting Scholar at the Information Systems Laboratory, Stanford University, Stanford, CA. His research interests are in Shannon theory, data compression algorithms, error correcting codes, and information theoretic cryptology.

Dr. Yamamoto served as the Chair of IEEE Information Theory Society Japan Chapter in 2002–2003, the TPC Co-Chair of the ISITA2004, the TPC Chair of the ISITA2008, the President of the SITA (Society of Information Theory and its Applications) in 2008–2009, the President of the ESS (Engineering Sciences Society) of IEICE in 2012–2013, an Auditor of the IEICE in 2016–2017, an Associate Editor for Shannon Theory, the IEEE TRANSACTIONS ON INFORMATION THEORY in 2007–2010, the Editor-in-Chief for the IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences in 2009–2011. He is a Fellow of the IEICE.