LQG Control Approach to Gaussian Broadcast Channels with Feedback

Ehsan Ardestanizadeh, Paolo Minero, and Massimo Franceschetti

Abstract—**A code for communication over the k-receiver additive white Gaussian noise broadcast channel with feedback is presented and analyzed using tools from the theory of linear quadratic Gaussian optimal control. It is shown that the performance of this code depends on the noise correlation at the receivers and it is related to the solution of a discrete algebraic Riccati equation. For the case of independent noises, the sum rate achieved by the proposed code, satisfying average power constraint P, is characterized as $1/2 \log(1 + P\phi)$, where the coefficient $\phi \in [1, k]$ quantifies the power gain due to the presence of feedback. When specialized to the case of two receivers, this includes a previous result by Elia and strictly improves upon the code of Ozarow and Leung. When the noises are correlated, the pre-log of the sum-capacity of the broadcast channel with feedback can be strictly greater than one. It is established that for all noise covariance matrices of rank r the pre-log of the sum capacity is at most $k - r + 1$ and, conversely, there exists a noise covariance matrix of rank r for which the proposed code achieves this upper bound. This generalizes a previous result by Gastpar and Wigger for the two-receiver broadcast channel.**

I. INTRODUCTION

Consider the communication problem over the k-receiver additive white Gaussian noise (AWGN) broadcast channel (BC) with feedback depicted in Fig. 1. Here, a sender wishes to communicate k independent messages to k distinct receivers that observe the sequence of transmitted signals corrupted by k, possibly correlated, AWGN sequences. It is known that the presence of feedback from receivers to sender improves communication performance over broadcast channels. Specifically, Dueck [1] showed, by providing a specific example, that feedback can enlarge the capacity region of additive memoryless broadcast channels where the noises at the receivers are correlated, while Ozarow and Leung [2] proved that feedback can be beneficial by providing a means of cooperation between receivers and sender even when the noises at the receivers are independent. However, a computable characterization of the capacity region of this channel remains a long-standing open problem.

In this paper, we construct a code, which we refer to as *LQG code*, for communicating over the AWGN-BC with feedback and we characterize its performance using tools from optimal control. We show that the minimum power needed by the LQG code for reliable transmission of messages encoded at

Ehsan Ardestanizadeh and Massimo Franceschetti are with the Information Theory and Applications Center (ITA) of the California Institute of Telecommunications and Information Technologies (CALIT2), Department of Electrical and Computer Engineering, University of California, San Diego CA, 92093, USA. Email: ehsan@ucsd.edu, massimo@ece.ucsd.edu. Paolo Minero is with the Department of Electrical Engineering of the University of Notre Dame, Notre Dame, IN, 46556, USA. Email: pminero@nd.edu

a given set of rates depends on the correlation between the noises at the receivers and is determined by the solution of a discrete algebraic Riccati equation (DARE) which arises in the analysis of the linear quadratic Gaussian (LQG) optimal control problem. The LQG code is then used to investigate some properties of the capacity region of the AWGN-BC with feedback.

First, we consider the case of independent noises at the receivers, which is the most interesting in practice. By solving the corresponding DARE, it is shown that the LQG code achieves sum rate equal to $1/2 \log(1 + P\phi(k, P))$ under average power constraint P when messages are encoded at the same rate. Here, the real coefficient $\phi(k, P) \in [1, k]$ represents the power gain compared to the no-feedback sum capacity $1/2 \log(1 + P)$, and for fixed k it is increasing in P, that is, more power allows more gain. In particular, as $P \to \infty$, $\phi \to k$, and the sum rate tends to $1/2 \log(1 + kP)$, which is the same as the sum capacity of the single-input multiple-output (SIMO) channel. Note that in the SIMO channel the receivers are co-located, whereas in the AWGN-BC the receivers are separately located but feedback links to the transmitter are available. Hence, the power gain due to feedback can be interpreted as the amount of *cooperation* among the receivers established through feedback, which allows the transmitter to align the signals intended for different receivers coherently and use power more efficiently.

Next, we investigate how the sum capacity, the supremum of achievable sum rates, scales as the power P at the transmitter increases. If the sum capacity scales as $(\gamma/2) \log(1 + P)$ as $P \to \infty$, then we refer to γ as the pre-log[1] of the channel. We show that the pre-log of the AWGN-BC with feedback depends on the rank of the correlation matrix of the noises at the receivers. Specifically, if the rank is r, then the pre-log can be at most $k - r + 1$. Conversely, for any $r \in \{1, \ldots, k\}$, there exists a noise covariance matrix of rank r for which the upper bound on the pre-log is tight and is achieved by the LQG code. In particular, the pre-log is equal to k for some rank-one covariance matrix. This generalizes a previous result by Gastpar and Wigger [3] for the two-receiver AWGN-BC to the case of k receivers.

Finally, we wish to mention some additional related works. Elia [4] followed a control-theoretic approach and presented a linear code for the two-receiver AWGN-BC with independent noises, which outperforms the Ozarow–Leung (OL) code [2]. Our code, when specialized to the case of two receivers and independent noises, provides the same performance as

---

[1]The pre-log is also known as the number of degrees of freedom as it is equal to the number of orthogonal point-to-point channels with the same sum capacity.

Elia's code [4]. Wu et al. [5] applied the LQG theory to study Gaussian networks with feedback, where the noises at the receivers are independent, but did not provide explicit solutions. Along the same lines, it has been shown in [6] that the linear code proposed by Kramer [7] for the k-sender multiple access channel with feedback can be obtained by solving an LQG control problem.

It is worth noting that the LQG code is derived from an optimal control for a linear system and hence is optimal among the subclass of linear codes. For the AWGN-BC with feedback, we show that the LQG code provides better performance compared to the OL code for $k = 2$ and hence outperforms the Kramer code [7] for $k \geq 3$, which is an extension of the OL code. However, it remains to be proven whether the LQG code achieves the feedback sum capacity.

The rest of the paper is organized as follows. Section II presents the problem definition. Section III discusses the point-to-point communication problem over the AWGN channel with feedback from a control theoretic perspective. This viewpoint is then generalized in Section IV, which presents the LQG code for communicating over the k-user AWGN-BC with feedback. The following two sections are devoted to studying the performance of the LQG code: in Section V we provide the analysis for the case of independent noises at the receivers, while in Section VI we characterize the pre-log gain when the noises are correlated. Finally, Section VII concludes the paper.

II. DEFINITIONS

Consider the communication problem where a sender wishes to communicate k independent messages $M_1, \ldots, M_k$ to k distinct receivers by n transmissions over the AWGN-BC channel with feedback depicted in Fig. 1. At each time $i \in \{1, \ldots, n\}$, the channel outputs are given by

$$\mathbf{Y}_i = \mathbf{1} X_i + \mathbf{Z}_i \tag{1}$$

where X_i is the transmitted symbol by the sender and $\mathbf{1}_{k \times 1} = (1, \ldots, 1)^T$ is the column vector of ones of length k. The vector $\mathbf{Y}_i := (Y_{1i}, \ldots, Y_{ki})^T$ contains the k channel outputs at time i, that is, Y_{ji} is the channel output for the receiver j at time i. Similarly Z_{ji} denotes the noise for the receiver j at time i. The noise vector

$$\mathbf{Z}_i := (Z_{1i}, \ldots, Z_{ki})^T \sim \mathrm{N}(0, K_z) \text{ i.i.d.}$$

is assumed to be independent of the transmitted messages, and independently and identically distributed (i.i.d.) from a Gaussian distribution with zero mean and covariance matrix K_z. Without loss of generality, we assume that $Z_{ji} \sim \mathrm{N}(0, 1)$ has unit variance, so the diagonal elements of K_z are equal to one.

We assume that the output symbols are causally and noiselessly fed back to the sender so that the transmitted symbol X_i at time i can depend on the message vector $\mathbf{M} := (M_1, \ldots, M_k)^T$, and the previous channel output vectors $\mathbf{Y}^{i-1} := (\mathbf{Y}_1, \ldots, \mathbf{Y}_{i-1})$.

Definition 1: A $(2^{nR_1}, \ldots, 2^{nR_k}, n)$ code for the AWGN-BC with feedback consists of

1) k discrete messages $M_j \sim \mathrm{Unif}\{1, \ldots, 2^{nR_j}\}$,

2) an encoder that assigns a symbol $X_i(\mathbf{M}, \mathbf{Y}^{i-1})$ to the message vector $\mathbf{M}$ and the previous channel output vectors $\mathbf{Y}^{i-1}$ for each $i \in \{1, \ldots, n\}$, and

3) k decoders, where decoder j assigns an estimate $\hat{M}_j$ to each sequence $Y_j^n := (Y_{j1}, \ldots, Y_{jn})$.

Let $M_1, \ldots, M_k$ be independent and the probability of error be defined as

$$P_e^{(n)} = \mathsf{P}(M_j \neq \hat{M}_j \text{ for some } j).$$

Then, we say that $(R_1, \ldots, R_k)$ is an achievable rate vector under (asymptotic block) power constraint P if there exists a sequence of $(2^{nR_1}, \ldots, 2^{nR_k}, n)$ codes such that

$$\lim_{n \to \infty} P_e^{(n)} = 0$$

and

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}(X_i^2) \leq P. \tag{2}$$

We refer to $R = \sum_{j=1}^{k} R_j$ as the sum rate of an achievable rate vector.

Definition 2: The closure of the set of achievable rate vectors $(R_1, \ldots, R_k)$ under power constraints P is called the capacity region $\mathscr{C}(P, K_z)$. The sum capacity $C(P, K_z)$ is defined as

$$C(P, K_z) := \max \left\{ \sum_{j=1}^{k} R_j : (R_1, ..., R_k) \in \mathscr{C}(P, K_z) \right\}$$

and the pre-log $\gamma(K_z)$ is defined as

$$\gamma(K_z) = \limsup_{P \to \infty} \frac{C(P, K_z)}{\frac{1}{2} \log(1 + P)}.$$

Definition 3: An n-code for the AWGN-BC with feedback consists of

1) k continuous messages $\Theta_j \sim \mathrm{Unif}(0, 1)$

2) an encoder that assigns a symbol $X_i(\mathbf{\Theta}, \mathbf{Y}^{i-1})$ to the message vector $\mathbf{\Theta} = (\Theta_1, \ldots, \Theta_k)^T$ and the previous channel output vectors $\mathbf{Y}^{i-1}$ for each $i \in \{1, \ldots, n\}$, and

3) k decoders, where decoder j assigns an estimate $\hat{\Theta}_j$ to each sequence $Y_j^n = (Y_{j1}, \ldots, Y_{jn})$.

Let $\Theta_1, \ldots, \Theta_k$ be independent and the mean square errors $D_1^{(n)}, \ldots, D_k^{(n)}$ at time n be defined as

$$D_j^{(n)} = \mathsf{E}(\Theta_j - \hat{\Theta}_j)^2, \quad j = 1, \ldots, k.$$

Then, we say that $(E_1, \ldots, E_k)$ is an achievable mean square error (MSE) exponent vector under (asymptotic block) power constraint P if there exists a sequence of n-codes such that

$$E_j = \lim_{n \to \infty} -\frac{1}{2n} \log D_j^{(n)}, \quad j = 1, \ldots, k$$

and (2) holds.

The definitions of achievable MSE exponent and rate vectors are closely related, as established by the following lemma.
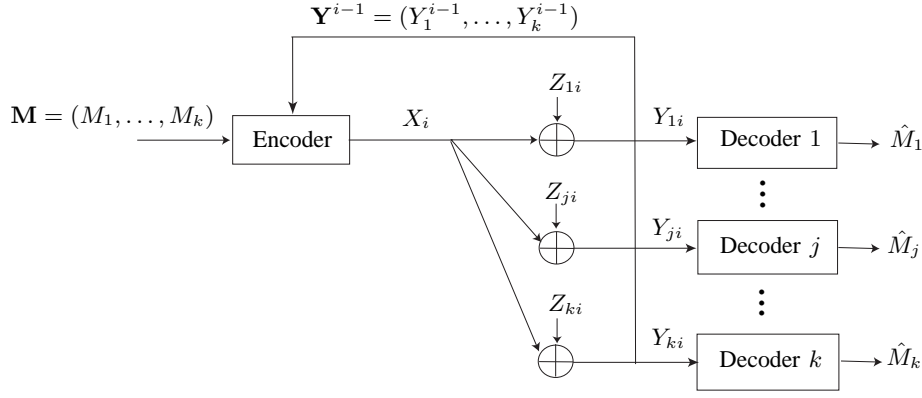
Fig. 1. The k-receiver AWGN broadcast channel with feedback.

Lemma 1: Let $(E_1, \ldots, E_k)$ be an achievable MSE exponent vector under power constraint P, and $(R_1, \ldots, R_k)$ be such that

$$R_j < E_j, \quad j = 1, \ldots, k.$$

Then, the rate vector $(R_1, \ldots, R_k)$ is achievable under power constraint P.

Proof: See Appendix A.

III. LQG APPROACH: POINT-TO-POINT CHANNELS

Before presenting the LQG code for the AWGN-BC, we first revisit the communication problem over the point-to-point AWGN channel with feedback in Fig. 2, and demonstrate how the theory of LQG control can be used to design codes for communication over such channel. It is well known [4] that the capacity-achieving code by Schalkwijk and Kailath (SK) [8], [9] can be derived as the solution of an optimal control problem. However, here we provide a derivation of this result that naturally generalizes to the case of multiple receivers.

Let $S_1 \in \mathbb{R}$ be the initial state of an unstable linear system with open-loop dynamics

$$S_i = aS_{i-1}, \quad i = 2, 3, \ldots$$

for some coefficient $a > 1$, that is stabilized by a controller having full state information and where the control signal is corrupted by AWGN (see Fig. 3). The closed-loop dynamic of this system is given by

$$S_i = aS_{i-1} + Y_{i-1}, \quad i = 2, 3, \ldots \tag{3}$$

where

$$Y_i = X_i + Z_i, \quad Z_i \sim \mathrm{N}(0, 1), \tag{4}$$

and

$$X_i = \pi_i(S_i), \quad i = 1, 2, \ldots \tag{5}$$

The mappings $\{\pi_i\}_{i=1}^{\infty}$ are referred to as the *control*, and the linear dynamical system in (3), which is characterized by the unstable mode a, is simply referred to as the *system*. We say that a control is stabilizing if $\limsup_{i \to \infty} \mathsf{E}(S_i^2) < \infty$.

Given a control (5) for the system (3), we construct a sequence of n-codes for the point-to-point AWGN channel (4)

with feedback (cf. Definition 3 in the special case of $k = 1$) as follows.

1) Encoder: Given a continuous message $\Theta \sim \mathrm{Unif}(0, 1)$, the encoder recursively forms

$$\begin{aligned} S_1 &= \Theta \\ S_i &= aS_{i-1} + Y_{i-1}, \quad i = 2, 3, \ldots, n \end{aligned} \tag{6}$$

and transmits $X_i(\Theta, Y^{i-1}) = \pi_i(S_i)$ for each $i \in \{1, \ldots, n\}$.

2) Decoder: the decoder sets $\hat{\Theta}_i = -a^{-i}\hat{S}_{i+1}$ as an estimate of the message Θ, where $\hat{S}_i$ is recursively formed as

$$\begin{aligned} \hat{S}_1 &= 0 \\ \hat{S}_i &= a\hat{S}_{i-1} + Y_{i-1}, \quad i = 2, 3, \ldots, n. \end{aligned} \tag{7}$$

Lemma 2: If the control $\{\pi_i\}$ stabilizes the system with open-loop mode $a > 1$, then the n-code described by (6) and (7) achieves MSE exponent $E = \log a$ under power constraint

$$P = \limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}(\pi_i^2(S_i)). \tag{8}$$

Proof: Combining (6) and (7) we have

$$\begin{aligned} S_{i+1} &= a^i \Theta + \hat{S}_{i+1} \\ &= a^i(\Theta - \hat{\Theta}_i) \end{aligned} \tag{9}$$

where the last equality follows from the fact that $\hat{\Theta}_i = -a^{-i}\hat{S}_{i+1}$. From (9), the MSE of the estimate $\hat{\Theta}_n$ is $D^{(n)} = \mathsf{E}(\Theta - \hat{\Theta}_n)^2 = a^{-2n}\mathsf{E}(S_{n+1}^2)$. Since the control is stabilizing we know $\limsup_{n \to \infty} \mathsf{E}(S_n^2) < \infty$ and hence

$$E = \lim_{n \to \infty} -\frac{1}{2n} \log D^{(n)} = \log a.$$

Noting that $X_i = \pi_i(S_i)$, we conclude that MSE exponent $\log a$ is achievable under asymptotic power equal to $\limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}(\pi_i^2(S_i))$. ∎

Lemma 2 shows that for a fixed $a > 1$ any stabilizing control yields a code achieving MSE exponent $E = \log a$. In general, though, different codes require a different asymptotic power constraint (8). Thus, we are interested in finding the stabilizing control whose associate asymptotic power is
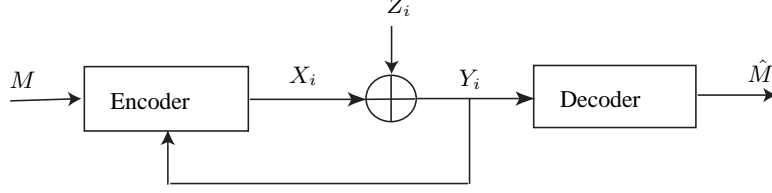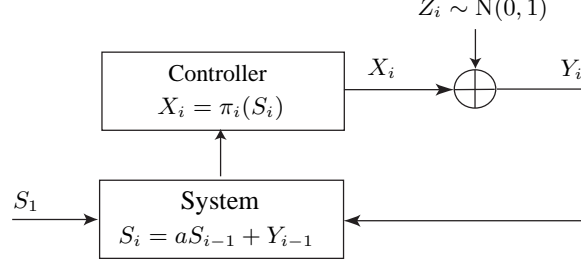
Fig. 2. The AWGN channel with feedback.

Fig. 3. Stabilization over the AWGN channel.

minimal. It is known from the theory of LQG control [10] that the linear stationary control

$$X_i = -cS_i, \quad i = 1, 2, \ldots \tag{10}$$

where $c = (a^2 - 1)/a$ attains the minimum asymptotic power (8), which is given by

$$P^* = a^2 - 1.$$

Hence, from Lemma 2, the linear code corresponding to this optimal LQG control, which we refer to as the LQG code, achieves the MSE exponent

$$\log a = \frac{1}{2}\log(1 + P^*)$$

under power constraint P^*. By Lemma 1, when specialized to $k = 1$, we conclude that the LQG code achieves any rate $R < \log a = 1/2 \log(1 + P^*)$ under power constraint P^*, and hence is capacity achieving.

A natural question to ask is what is the connection between the LQG code and the SK code, where the sender recursively transmits the estimation error at the receiver? To answer this question, note that combining (6) and (10) we can write a recursion for the channel input X_i as

$$X_{i+1} = a\left(X_i - \frac{c}{a}Y_i\right) \tag{11}$$

with $X_1 = -c\Theta$. The recursion converges, irrespectively of the initial value, since $a - c = 1/a < 1$. On the other hand, the channel input update in the SK can be represented by the following recursion [11],

$$X_{i+1} = a\left(X_i - \frac{\mathsf{E}(X_iY_i)}{\mathsf{E}(Y_i^2)}Y_i\right) \tag{12}$$

with $X_1 = \Theta$. Comparing (11) and (12), we can see that the LQG code is asymptotically equivalent to the SK code if

$$\frac{\mathsf{E}(X_iY_i)}{\mathsf{E}(Y_i^2)} \to \frac{c}{a} \quad \text{as } i \to \infty \tag{13}$$

such that, in steady state, $(c/a)Y_i$ tends to the minimum mean squared error estimate of X_i given Y_i.

By plugging (10) into (6), we have the closed-loop dynamics for S_i as

$$S_i = a^{-1}S_{i-1} + Z_{i-1}.$$

As $i \to \infty$ the second moment of the state converges to $a^2/(a^2 - 1)$ and since $X_i = -cS_i$,

$$\mathsf{E}(X_i^2) \to a^2 - 1$$
$$\mathsf{E}(Y_i^2) \to a^2.$$

Therefore, as $i \to \infty$,

$$\frac{\mathsf{E}(X_iY_i)}{\mathsf{E}(Y_i^2)} = \frac{\mathsf{E}(X_i^2)}{\mathsf{E}(Y_i^2)} \to \frac{a^2 - 1}{a^2} = \frac{c}{a}$$

where the last equality follows from the fact that the optimal control is given by $c = (a^2 - 1)/a$. Hence, the LQG code and the SK code are asymptotically equivalent.

IV. LQG CODE: AWGN BROADCAST CHANNEL WITH FEEDBACK

In this section we extend the control theoretic approach to the case of a k-receiver AWGN-BC with feedback. We do so by considering the control problem depicted in Fig. 4, in which a k-dimensional unstable dynamical system is stabilized by a scalar controller having full state observation and where the scalar controller is perturbed by k, possibly correlated, AWGN
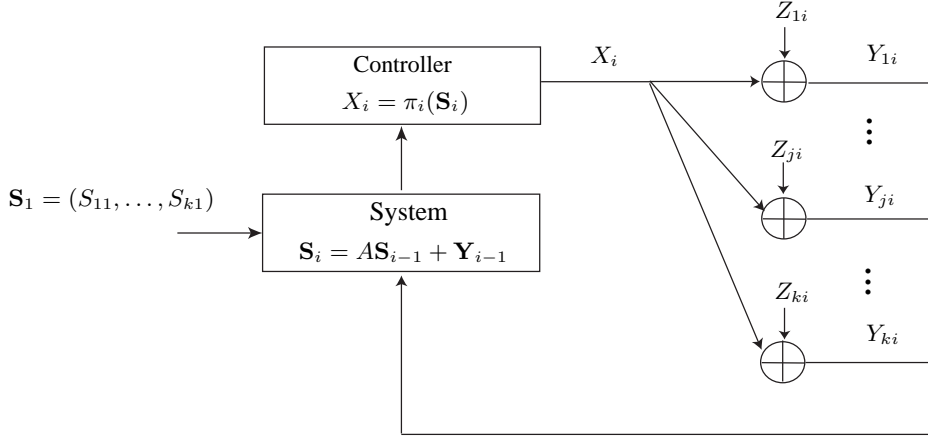
$$M_j$$
$$\hat{M}_k$$

Fig. 4. Control over the AWGN broadcast channel

noises, each affecting a different component of the state vector. We show that any controller stabilizing the system in mean-square sense yields a code for the k-receiver AWGN-BC with feedback. In particular, the LQG code is obtained from the minimum variance stabilizing control which can be computed using the LQG control theory.

A. Code Design Based on a Control Approach

Assume for simplicity that channel input and outputs in (1) are complex numbers and that the additive noise vector is drawn i.i.d. from a circular symmetric complex Gaussian distribution with covariance K_z. It is easy to see that if $(R_1, \ldots, R_k)$ is achievable under power constraint P per each real dimension of the complex channel, then $(R_1, \ldots, R_k)$ is achievable under the same power constraint over the original real channel. In fact, one transmission over the complex channel can be reproduced by two consecutive transmission (of real and imaginary part, respectively) over the real channel.

Let

$$A = \operatorname{diag}(a_1, \ldots, a_k) \in \mathbb{C}^{k \times k} \tag{14}$$

where $a_j \in \mathbb{C}$, $j = 1 \ldots, k$, are distinct points outside the unit circle, i.e., $|a_j| > 1$, and A' denote the *conjugate transpose* of the matrix A. Consider the linear dynamical system with open-loop matrix A shown in Fig. 4,

$$\mathbf{S}_1 = \boldsymbol{\Theta}$$
$$\mathbf{S}_i = A\mathbf{S}_{i-1} + \mathbf{Y}_{i-1}, \quad i = 2, 3, \ldots \tag{15}$$

where $\mathbf{S}_i = (S_{1i}, \ldots, S_{ki})^T \in \mathbb{C}^k$ represents the state of the system at time i, $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_k)^T$ is a vector of complex random variables such that $\Theta_1, \ldots, \Theta_k$ are drawn i.i.d. from a uniform distribution over $(0, 1) \times (0, 1) \subset \mathbb{C}$, and $\mathbf{Y}_i \in \mathbb{C}^k$ denotes the vector of complex channel outputs, i.e.,

$$\mathbf{Y}_i = BX_i + \mathbf{Z}_i, \quad \mathbf{Z}_i \sim \mathcal{CN}(0, K_z) \tag{16}$$

where $X_i \in \mathbb{C}$ here represents the scalar complex control signal, $\mathbf{Z}_i \in \mathbb{C}^k$ denote the noise vector at time i, and

$$B = [1, \ldots, 1]'_{1 \times k}. \tag{17}$$

At every discrete time i, the control input can depend only on the state of the system at time i, so

$$X_i = \pi_i(\mathbf{S}_i), \quad i = 1, 2, \ldots \tag{18}$$

for some function $\pi_i : \mathbb{C}^k \to \mathbb{C}$. We refer to the sequence $\{\pi_i\}$ as the control. Since $|a_j| > 1$, the eigenvalues of A are outside the unit circle and the open-loop system (15) is unstable. We say that the control $\{\pi_i\}$ stabilizes the closed-loop system if

$$\limsup_{n \to \infty} \mathsf{E}(\|\mathbf{S}_n\|^2) < \infty$$

where $\|\mathbf{S}_n\|^2$ denotes the 2-norm of the vector $\mathbf{S}_n$.

Given the system (15) and the control (18), we present the following sequence of n-codes for the k-receiver AWGN-BC with feedback (16).

1) Encoder: At each time $i \in \{1, \ldots, n\}$ the encoder recursively forms $\mathbf{S}_i$ as in (15) and transmits $X_i = \pi_i(\mathbf{S}_i)$.
2) Decoders: At each time $i \in \{1, \ldots, n\}$ decoder j forms an estimate $\hat{\Theta}_{ji} = -a_j^{-i} \hat{S}_{j(i+1)}$ for the desired message Θ_j, where $\hat{S}_{ji}$ is recursively formed as

$$\hat{S}_{j1} = 0$$
$$\hat{S}_{ji} = a_j \hat{S}_{j(i-1)} + Y_{j(i-1)}, \quad i = 2, 3, \ldots, n. \tag{19}$$

The following lemma characterizes the set of MSE exponent vectors that can be achieved by the sequence of n-codes so generated.

Lemma 3: Let $\{\pi_i\}$ be a stabilizing control for (15). Then, the MSE exponent vector $(\log|a_1|, \ldots, \log|a_k|)$ is achievable under power constraint

$$P = \limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}(\pi_i^2(\mathbf{S}_i)).$$

Proof: See Appendix B.

B. LQG Code based on Optimal LQG Control

According to Lemma 3, for a fixed open-loop matrix A, *any* stabilizing control yields a sequence of n-codes for the AWGN-BC with feedback that achieves the same MSE exponent vector $(\log|a_1|, \ldots, \log|a_k|)$ under a power constraint

determined by the asymptotic control variance. The following theorem characterizes the performance of the LQG code, which corresponds to the minimum variance control that can be computed using the LQG control theory.

Theorem 1: Let A and B be given as in (14) and (17), and K_z be the covariance matrix of the noise vector in the AWGN-BC (16). Then, the rate vector $(\log|a_1|, \ldots, \log|a_k|)$ is achievable under power constraint

$$P(A, K_z) = CK_sC' = \operatorname{tr}(GK_z) \qquad (20)$$

where

$$C = (B'GB + 1)^{-1}B'GA \qquad (21)$$

and G is the unique positive definite solutions to the discrete algebraic Riccati equation (DARE)

$$G = A'GA - A'GB(B'GB + 1)^{-1}B'GA \qquad (22)$$

such that $A - BC$ is stable, that is, every eigenvalue of $A - BC$ lies inside the unit circle, and K_s is the unique solution to the discrete algebraic Lyapunov equation (DALE)

$$K_s = (A - BC)K_s(A - BC)' + K_z. \qquad (23)$$

The proof of the theorem makes use of the following two lemmas.

Lemma 4: Given the unstable open-loop matrix (14), the stationary and linear control

$$X_i = -C\mathbf{S}_i \qquad (24)$$

where C and G are given in (21) and (22), respectively, stabilizes the closed-loop system (15) and minimizes the asymptotic average control power

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}(\pi_i^2(\mathbf{S}_i)). \qquad (25)$$

The minimum stationary power is given by

$$\operatorname{tr}(GK_z)$$

where K_z is the covariance matrix of the noise vector in (16).

Proof: Plugging (16) into (15), we have

$$\mathbf{S}_i = A\mathbf{S}_{i-1} + BX_{i-1} + \mathbf{Z}_{i-1}. \qquad (26)$$

Consider the problem of finding the stabilizing control that minimizes the asymptotic average control power

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}(\pi_i^2(\mathbf{S}_i)).$$

This problem is similar to the standard LQG problem [10] in the special case where the cost function does not depend on the state. For this problem, we can derive the Riccati equation (22) and the stationary linear control (21), similar to the solution to the LQG problem, to establish a sufficient condition for optimality in terms of the asymptotic power.

Unlike the standard LQG problem, though, here we require the control to stabilize the system (see Lemma 3). Next, we show that there exists a unique solution to (22) such that the control C in (21) is stabilizing, that is, $A - BC$ is stable. Since the eigenvalues of A are all outside of the unit circle and the

elements of B are nonzero we know (A, B) is detectable, that is, there exists a $C \in \mathbb{C}^{1 \times k}$ such that $A - BC$ is stable. Then, by [12, Lemma 2.4] there exists a *unique* positive definite solution to (22) for which $A - BC$ is stable. ∎

The following lemma characterizes the asymptotic performance of the minimum variance control.

Lemma 5: Let the linear control in (24) be stabilizing, that is, all eigenvalues of $A - BC$ lies inside the unit circle. Then, the asymptotic average control power (25) is given by

$$CK_sC' \qquad (27)$$

where K_s is the unique solution to the following DALE

$$K_s = (A - BC)K_s(A - BC)' + K_z.$$

Remark 1: From Lemma 4 and Lemma 5 it is clear that the performance of the described feedback code depends on the correlation among the noises at the receivers.

Proof: Plugging (24) into the closed-loop system dynamics (26) we have

$$\mathbf{S}_i = (A - BC)\mathbf{S}_{i-1} + \mathbf{Z}_{i-1}.$$

Let $K_{s,i}$ denote the covariance matrix of the state $\mathbf{S}_i$, then we have the following discrete algebraic Lyapunov recursion

$$K_{s,i+1} = (A - BC)K_{s,i}(A - BC)' + K_z.$$

By assumption $(A - BC)$ is stable and $K_{s,0} \succ 0$, hence the above recursion converges to the unique positive-definite solution to the following discrete algebraic Lyapunov equation (DALE)

$$K_s = (A - BC)K_s(A - BC)' + K_z.$$

Note that $X_i = -C\mathbf{S}_i$ and hence $\mathsf{E}(X_i^2) = CK_{s,i}C'$, which completes the proof. ∎

Proof of Theorem 1: According to Lemma 4 and Lemma 5, for $A = \operatorname{diag}(a_1, \ldots, a_k)$ with $|a_j| > 1$, there exists a stabilizing control with asymptotic power equal to (20). Furthermore, according to Lemma 3, we can construct a sequence of $n$codes corresponding to this control, which achieves the MSE exponent vector $(\log|a_1|, \ldots, \log|a_k|)$ with the same asymptotic power as the control. Finally, by Lemma 1, we conclude that rate vector $(\log|a_1|, \ldots, \log|a_k|)$ is achievable under asymptotic power constraint (20). ∎

Example 1: Consider the special case of a two-receiver AWGN-BC, and let

$$A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \quad K_z = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

for $|a_1| > 1$, $|a_2| > 1$, and $-1 < \rho < 1$. By solving (22) and plugging the solution into (20) we obtain that $(\log|a_1|, \log|a_2|)$ is an achievable rate pair under power constraint

$$\frac{1}{|a_1 - a_2|^2}\Big(|a_1a_2 - 1|^2(|a_1|^2 + |a_2|^2 - 2)$$
$$- \rho(|a_1|^2 - 1)(|a_2|^2 - 1)(\operatorname{Re}(a_1a_2') - 2)\Big).$$

In the special case where the noises at the receivers are independent ($\rho = 0$), the code in [4] has the same performance as the LQG code.

V. INDEPENDENT NOISES: POWER GAIN

In this section, we analyze the performance of the LQG code in the special case of independent noises, i.e., assuming that $K_z = I_{k \times k}$, which is the case for most practical scenarios. We characterize the sum rate $R(k, P)$ achievable by the LQG code under power constraint P for the symmetric case where diagonal elements of A in (14) are

$$a_j = a e^{\frac{2\pi\sqrt{-1}}{k}(j-1)}, \quad j = 1, \ldots, k. \tag{28}$$

and $a > 1$ is real.

Theorem 2: Given A as in (28), the LQG code achieves symmetric sum rate $R(k, P)$, i.e., $R_j = R(k, P)/k$, $j = 1, \ldots, k$, under power constraint P where

$$R(k, P) = \frac{1}{2} \log(1 + P\phi)$$

and $\phi(k, P)$ is the unique solution in the interval $[1, k]$ to

$$(1 + P\phi)^{k-1} = \left(1 + \frac{P}{k}\phi(k - \phi)\right)^k.$$

Remark 2: The quantity $\phi(k, P)$ is the power gain compared to the no feedback sum capacity $1/2 \log(1 + P)$. This power gain can be interpreted as the amount of cooperation among the receivers established through feedback, which allows the transmitter to align signals intended for different receivers coherently and use power more efficiently.

Proof: For A defined as (28), we know by Theorem 1 that the sum rate $R = k \log a$ is achievable under power constraint $P = \text{tr}(G)$ where G is the unique solution to the DARE (22). The following lemma characterizes the solution to the DARE (22) for the symmetric choice A.

Lemma 6: [13, Lemma 12] Suppose that the open-loop matrix A is of the form (28). Then the unique positive-definite solution G to (22) is circulant with real eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_k > 0$ satisfying

$$\lambda_i = \frac{1}{a^2}\lambda_{i-1}$$

for $i = 2, \ldots, k$. The largest eigenvalue λ_1 satisfies

$$1 + k\lambda_1 = a^{2k} \tag{29}$$

$$\left(1 + \lambda_1 \left(k - \frac{\lambda_1}{G_{jj}}\right)\right) = a^{2(k-1)}. \tag{30}$$

From (29) and (30) we have

$$(1 + k\lambda_1)^{k-1} = \left(1 + \lambda_1(k - \frac{\lambda_1}{G_{jj}})\right)^k. \tag{31}$$

The solution to (22) is unique and we conclude that (31) has a unique solution for G_{jj} given λ_1 and vice-versa. Consider $\lambda_1(k, P)$ corresponding to the case where

$$G_{jj} = \frac{P}{k}, \quad j = 1, \ldots, k.$$

Note that the solution G to (22) is circulant and has equal elements on the diagonal. From (29) we know the LQG code corresponding to $\lambda_1(k, P)$ achieves sum rate

$$k \log a = \frac{1}{2} \log(1 + k\lambda_1(k, P)).$$

under power constraint $\text{tr}(G) = P$. The following change of variable

$$\phi = \frac{k}{P}\lambda_1$$

completes the proof. ∎

A. Comparison with the AWGN Multiple Access Channel

The LQG approach can be also applied to the AWGN-MAC with feedback. It is known [6] that the LQG code for AWGN-MAC has the same performance as the Kramer code [7], which achieves the linear sum capacity [13], the supremum sum rate achievable by linear codes. Let $R_{\text{MAC}}(k, P)$ denotes the symmetric sum rate achievable by the LQG code for the k-sender AWGN multiple access channel (MAC) with feedback where each sender has power constraint P. Then, we have [6, Theorem 4]:

$$R_{\text{MAC}}(k, P) = \frac{1}{2} \log(1 + kP\phi)$$

where $\phi(k, P)$ is the unique solution to

$$(1 + kP\phi)^{k-1} = (1 + P\phi(k - \phi))^k.$$

Comparing with Theorem 2, it is not hard to see that

$$R_{\text{BC}}(k, P) = R_{\text{MAC}}(k, P/k).$$

This shows that under the same *sum power* constraint P, the sum rate achievable by the LQG code over MAC and BC are equal. This connection between the MAC and the BC is also considered in [5].

B. Ozarow-Leung (OL) code for $k = 2$

We compare the LQG code with the OL code for $k = 2$ and $K_z = I$. The OL code can be represented as follows [11]:

$$X_i = S_{1i} + S_{2i}$$

where

$$\begin{bmatrix} S_1 \\ S_2 \end{bmatrix}_{i+1} = \begin{bmatrix} a & 0 \\ 0 & -a \end{bmatrix}\left(\begin{bmatrix} S_1 \\ S_2 \end{bmatrix}_i - \begin{bmatrix} \frac{\mathsf{E}[S_{1i}Y_{1i}]}{\mathsf{E}[Y_{1i}^2]} & 0 \\ 0 & \frac{\mathsf{E}[S_{2i}Y_{2i}]}{\mathsf{E}[Y_{2i}^2]} \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}_i\right) \tag{32}$$

In the sequel, we present the LQG code in a similar form as (32) for comparison. Let the system in (15) be re-written as

$$\mathbf{S}_i = A\mathbf{S}_{i-1} + \text{diag}(\tilde{B})\mathbf{Y}_{i-1}.$$

where

$$A = \begin{bmatrix} a & 0 \\ 0 & -a \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} -b \\ b \end{bmatrix} \tag{33}$$

where $b \neq 0$ is any constant and $a > 1$. The choice of b does not affect the performance of the optimal control as one can cancel out the effect of b by properly scaling the control signal. Thus, without loss of generality, in (15) we picked $b = 1$. According to the channel (16) the closed-loop system is

$$\mathbf{S}_i = A\mathbf{S}_{i-1} + \tilde{B}X_{i-1} + \text{diag}(\tilde{B})\mathbf{Z}_{i-1}.$$

By substituting B with $\tilde{B}$, the optimal control can be characterized by Lemma 4 as follows. The solution to the DARE (22) is

$$G = -\frac{(a^4 - 1)}{4a^4 b^2} \begin{bmatrix} 1 + a^2 & 1 - a^2 \\ 1 - a^2 & 1 + a^2 \end{bmatrix}$$

which yields the optimal control

$$C = [c_1 \ c_2] = -\frac{(a^4 - 1)}{2a^3 b} \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

To obtain the power we need to substitute $K_z = I$ with

$$Q = \text{diag}(\tilde{B}) \, \text{diag}(\tilde{B})' = \begin{bmatrix} b^2 & 0 \\ 0 & b^2 \end{bmatrix}.$$

and the asymptotic variance of the channel input X is given by

$$P = \text{tr}(GQ) = b^2 \, \text{tr}(G) = \frac{1}{2a^2}(a^2 - 1)(a^2 + 1)^2. \quad (34)$$

Notice that (34) does not depend on the parameter b. Thus, we can choose b arbitrarily without affecting the overall performance of the code. In particular, by choosing

$$b = \frac{a^4 - 1}{2a^3}$$

we have that $X_i = S_{1i} + S_{2i}$ as in the OL code. However, the state in OL is updated as in (32) while in the LQG code,

$$\begin{bmatrix} S_1 \\ S_2 \end{bmatrix}_{i+1} = \begin{bmatrix} a & 0 \\ 0 & -a \end{bmatrix} \left(\begin{bmatrix} S_1 \\ S_2 \end{bmatrix}_i - \begin{bmatrix} b/a & 0 \\ 0 & b/a \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}_i \right). \quad (35)$$

To compare with the OL code, we need to find the asymptotic covariance matrix of the state. By substituting K_z with Q, the asymptotic covariance matrix K_s is given by the DALE (23),

$$K_s = \frac{b^2}{2a^2} \begin{bmatrix} (a^2 + 1)^2 + \frac{2}{a^2 - 1} & a^4 + 1 \\ a^4 + 1 & (a^2 + 1)^2 + \frac{2}{a^2 - 1} \end{bmatrix}$$

and

$$\lim_{i \to \infty} \frac{\mathsf{E}[S_{2i} Y_{2i}]}{\mathsf{E}[Y_{2i}^2]} = \lim_{i \to \infty} \frac{\mathsf{E}[S_{1i} Y_{1i}]}{\mathsf{E}[Y_{1i}^2]}$$
$$= \frac{-c_1 (K_s)_{11} - c_2 (K_s)_{12}}{\text{tr}(GQ) + 1}$$
$$= b \cdot \frac{a^3(a^2 - 1)}{a^6 + a^4 + a^2 - 1}.$$

Notice that

$$b \cdot \frac{a^3(a^2 - 1)}{a^6 + a^4 + a^2 - 1} < \frac{b}{a}.$$

Therefore, unlike in the point-to-point setting discussed in Section III, here the OL code and the LQG code are not asymptotically equivalent. Although both codes achieve rate pair $(\log a, \log a)$, by Lemma 4, the OL code requires more asymptotic power than the LQG code and hence it is suboptimal.

VI. CORRELATED NOISES: PRE-LOG GAIN

In this section, we show that structured correlation among the noises at the receivers can increase the capacity significantly. We consider the high power regime and study the pre-log $\gamma(K_z)$ as a function of covariance matrix K_z, which represents the number of orthogonal point-to-point channels with the same sum capacity.

Theorem 3: For all K_z of rank r

$$\gamma(K_z) \le k - r + 1.$$

Conversely, for any $r = 1, \ldots, k$, there exists K_z such that $\text{rank}(K_z) = r$ and

$$\gamma(K_z) = k - r + 1.$$

Proof: First, we prove the upper bound by induction. By assumption K_z contains r linearly independent rows, let us assume, without loss of generality, the last r rows. Assume that receivers $k - r + 1, \cdots, k$ share their received signals and form a single receiver equipped with r receive antennas and let $\mathbf{Y}_{k-r+1} := (Y_{k-r+1}, \ldots, Y_k)^T$ denote the vector of received signals by this multiple antenna receiver. The corresponding AWGN vector BC with feedback is specified by

$$Y_j = X + Z_j, \qquad j = 1, \cdots, k - r,$$
$$\mathbf{Y}_{k-r+1} = \mathbf{1}_{r \times 1} X + \mathbf{Z}_{k-r+1}$$

where $(Z_1, \cdots, Z_{k-r}, \mathbf{Z}_{k-r+1}) \sim \text{N}(0, K_z)$, $\mathbf{Z}_{k-r+1} \sim \text{N}(0, \tilde{K}_z)$, and by assumption $\tilde{K}_z$ is full rank and invertible.

Now suppose that the sender of this channel wishes to send message M_j to receiver j, $j = 1, \cdots, k - r + 1$, under power constraint P. Since we made the optimistic assumption that a subset of receiver can cooperate, the sum capacity of this channel is an outer bound on the sum capacity of the original AWGN-BC. Note that for every $j = 1, \cdots, k - r$, the rate

$$R_j < \frac{1}{2} \log(1 + P)$$

is upper bounded by the capacity of the point-to-point AWGN channel. The rate R_{k-r+1} for the $(k-r+1)$-th receiver with r multiple antenna is upper bounded by the capacity of a single input multiple output (SIMO) [14] channel:

$$R_{k-r+1} \le \frac{1}{2} \log(1 + P|\tilde{K}^{-\frac{1}{2}} \mathbf{1}_{r \times 1}|^2)$$

where by assumption $\tilde{K}$ is invertible. Thus, the sum capacity of this channel is upper bounded by $(k - r)/2 \log(1 + P) + 1/2 \log(1 + P|\tilde{K}^{-\frac{1}{2}} \mathbf{1}_{r \times 1}|^2)$, and therefore the pre-log $\gamma(K_z)$ can be at most $k - r + 1$.

Next, we show $\gamma(K_z) = k$ is achievable by the LQG code for some K_z of rank one, i.e., $r = 1$. For $r = 2, \ldots, k$ similar argument holds. Suppose that the open-loop matrix A is as in (28). By Theorem 1, the symmetric rate vector $(\log a, \ldots, \log a)$ is achievable under the power constraint $\text{tr}(GK_z)$, where G is the circulant matrix in Lemma 6. Note that any circulant matrix can be written as $F\Lambda F'$, where F is the k point discrete Fourier transform matrix with

$$F_{jl} = \frac{1}{\sqrt{k}} e^{-2\pi \sqrt{-1}(j-1)(l-1)/k},$$

for $j, k \in \{1, \cdots, k\}$, and $\Lambda = \text{diag}([\lambda_1, \ldots, \lambda_k])$ is the matrix with eigenvalues on its diagonal. Suppose that the noise covariance matrix is also a circulant matrix with diagonal entries equal to 1. In particular, let

$$\tilde{K}_z = F\tilde{\Lambda}F', \quad \tilde{\Lambda} = \text{diag}([0, \ldots 0, k]). \quad (36)$$

Then, we have

$$\text{tr}(G\tilde{K}_z) = \text{tr}(F\Lambda\tilde{\Lambda}F') = k\lambda_k$$

where by Lemma 6,

$$k\lambda_k = \frac{k\lambda_1}{a^{2(k-1)}} = \frac{a^{2k} - 1}{a^{2(k-1)}}.$$

Therefore, for the symmetric choice of A in (28), the LQG code achieves sum rate

$$R = k \log a \quad (37)$$

under power constraint

$$P = \text{tr}(G\tilde{K}_z) = \frac{a^{2k} - 1}{a^{2(k-1)}} \quad (38)$$

and we have

$$\begin{aligned}
\gamma(\tilde{K}_z) &= \lim_{P \to \infty} \frac{C(P, \tilde{K}_z)}{\frac{1}{2}\log(1 + P)} \\
&\geq \lim_{P \to \infty} \frac{R}{\frac{1}{2}\log(1 + P)} \\
&= \lim_{a \to \infty} \frac{k \log a}{\frac{1}{2}\log a^2} \\
&= k.
\end{aligned} \quad (39)$$

where (39) follows by plugging sum rate R and power P from (37) and (38). Moreover, we know $\gamma(\tilde{K}_z) \leq k$ since $\text{rank}(K_z) = 1$. Hence, $\gamma(\tilde{K}_z) = k$ for the covariance matrix $\tilde{K}_z$ in (36).

To complete the proof, we show that for every $r \in \{2, \ldots, k-1\}$ a pre-log equal to $k - r + 1$ is achievable for some K_z such that $\text{rank}\, K_z = r$. Consider

$$K_z = \begin{bmatrix} M_{k-r+1, k-r+1} & 0_{k-r+1, r-1} \\ 0_{r-1, k-r+1} & I_{r-1, r-1} \end{bmatrix}$$

where $0_{i,j}$ denotes the zero matrix of dimension $i \times j$, $I_{i,i}$ is the identity matrix of dimension i, and $M_{k-r+1, k-r+1}$ is the $(k - r + 1) \times (k - r + 1)$ circulant matrix having first row equal to the last column of the discrete Fourier transform matrix of dimension $(k - r + 1) \times (k - r + 1)$. Clearly, $\text{rank}(K_z) = \text{rank}(M_{k-r+1, k-r+1}) + \text{rank}(I_{r-1, r-1}) = r$. On the other hand, suppose that the transmitter communicates only to users $1, 2, \cdots, k-r+1$, while the transmission rate for the remaining users is set to zero. We can use a similar argument as above and show that the LQG code for the corresponding $(k - r + 1)$-receiver AWGN-BC with feedback and noise covariance matrix $M_{k-r+1, k-r+1}$ achieves a pre-log equal to $k - r + 1$. ∎

Remark 3: To achieve $\gamma = k$, we used the LQG code. However, the same pre-log can be achieved even with codes which are less power efficient since we are considering only the pre-log of the sum rate in the high power regime. For instance, for the special case of $k = 2$, Gastpar and Wigger [3] showed that the OL code, which is suboptimal, achieves pre-log two for anti-correlated noises.

VII. CONCLUSION

Using tools from control theory we have presented a code for the k-user AWGN-BC with feedback, called the LQG code, which we have then used to investigate some properties of the capacity region of this channel. When the noises at the receivers are independent the pre-log of the sum capacity is at most one, so feedback can yield at most a power gain over the case without feedback. We have quantified the power gain achieved by the LQG code and shown that in the case where $k = 2$, the LQG code recovers a previous result of Elia which strictly improves upon the OL code. In the case where the noises at the receivers are correlated, instead, the pre-log of the sum capacity can be strictly greater than one. We established that for all noise covariance matrixes of rank r the pre-log is at most $k - r + 1$ and, conversely, there exists a covariance matrix for which this upper bound is achieved by the LQG code. In particular, a pre-log equal to k is achievable for some circulant noise covariance matrix of rank one. This generalizes previous results obtained by Gastpar and Wigger for the case $k = 2$.

The LQG approach exploited here could be in principle useful for other multi-user communication channels with feedback, when the subclass of linear codes can lead to optimal or close to optimal solutions.

APPENDIX A
PROOF OF LEMMA 1

The proof is closely related to the proof of an analogous statement for the communication problem over the multiple access channel with feedback [6]. By assumption, there exists a sequence of n-codes for $\Theta_j \sim \text{Unif}(0, 1)$, such that

$$E_j = \lim_{n \to \infty} -\frac{1}{2n} \log D_j^{(n)}, \quad j = 1, \ldots, k \quad (40)$$

and (2) holds. Given the sequence of n-codes and $R_j < E_j$, $j = 1 \ldots, k$, we construct a sequence of $(2^{nR_1}, \ldots, 2^{nR_k}, n)$ codes such that $\lim_{n \to \infty} P_e^{(n)} = 0$.

First, we map the discrete message $m_j \in \mathcal{M}_j = [1 : 2^{nR_j}]$ to a message point $\theta_j(m_j) \in \Theta_j$, where Θ_j is a set of 2^{nR_j} message points in the unit interval such that the distance between any two message points is greater than or equal to 2^{-nR_j}. To send $m_j \in \mathcal{M}_j$, we use the given n-code and the corresponding message point $\theta_j(m_j)$. The decoder first forms the estimate of the message point $\hat{\theta}_j(y^n)$ according to the given n-code, and then chooses $\hat{m}_j$ such that $\theta_j(\hat{m}_j)$ is the closest message point to $\hat{\theta}_j(y^n)$. As the distance between any two message points is greater than or equal to 2^{-nR_j}, the average probability of error is bounded as follows,

$$P_e^{(n)} \leq \max_j \max_{\theta \in \Theta_j} \mathsf{P}\left\{|\Theta_j - \hat{\Theta}_j| > \frac{1}{2} \cdot 2^{-nR_j}\Big|\Theta_j = \theta\right\}. \quad (41)$$

To show $\lim_{n \to \infty} P_e^{(n)} = 0$, consider

$$p_{j,n} := \mathsf{P}\left\{|\Theta_j - \hat{\Theta}_j| > \frac{1}{2} \cdot 2^{-nR_j}\right\} \qquad (42)$$

$$\leq 4 \cdot 2^{2nR_j} \cdot D_j^{(n)} \qquad (43)$$

$$\leq 4 \cdot 2^{2nR_j} \cdot 2^{-2n(E_j - \epsilon_n)} \qquad (44)$$

$$= 4 \cdot 2^{-2n(E_j - R_j - \epsilon_n)} \qquad (45)$$

where $\epsilon_n \to 0$ as $n \to \infty$. The inequalities (43) and (44) follow from the Chebyshev inequality and (40), respectively. From (45) and the assumption $R_j < E_j$, we have

$$p_{j,n} \to 0 \text{ as } n \to 0 \quad j = 1, \ldots, k. \qquad (46)$$

Next, by the similar argument as in [15, Lemma II.3] we show that condition (46) is sufficient to prove that there exists a set of message points in the unit interval such that the distance between any two message points is greater than or equal to 2^{-nR_j} and

$$\lim_{n \to \infty} \max_{\theta_j \in \Theta_j} \mathsf{P}\left\{|\Theta_j - \hat{\Theta}_j| > \frac{1}{2} \cdot 2^{-nR_j}\Big|\Theta_j = \theta_j\right\} = 0 \quad (47)$$

for $j = 1, \ldots, k$.

Define the event

$$T_{j,n} = \Big\{\theta \in (0,1):$$
$$\mathsf{P}\left\{|\Theta_j - \hat{\Theta}_j| > \frac{1}{2} \cdot 2^{-nR_j}\Big|\Theta_j = \theta\right\} > \sqrt{p_{j,n}}\Big\}.$$

Then we have $p_{j,n} > \sqrt{p_{j,n}}\,\mathsf{P}(T_{j,n})$ and hence

$$\mathsf{P}(T_{j,n}) < \sqrt{p_{j,n}}.$$

To choose $\Theta_{\mathbf{j}}$ such that $\Theta_j \cap T_{j,n} = \emptyset$ and also the distance between any two message points is greater than or equal to 2^{-nR_j}, it is sufficient that $|\Theta_j|2^{-nR_j} \leq 1 - \sqrt{p_{j,n}}$ or considering (46),

$$|\Theta_j| \leq (1 - \epsilon_n) \cdot 2^{nR_j}$$

where $\epsilon_n \to 0$ as $n \to \infty$. Moreover, by the definition of $T_{j,n}$ and the fact that $\Theta_j \cap T_n = \emptyset$ we have

$$\max_{\theta \in \Theta_j} \mathsf{P}\left\{|\Theta - \hat{\Theta}| > \frac{1}{2} \cdot 2^{-nR_j}\Big|\Theta_j = \theta\right\} \leq \sqrt{p_{j,n}} \qquad (48)$$

and considering (46), the condition (47) holds.

Combining (41) and (47), we have $\lim_{n \to \infty} P_e^{(n)} = 0$. Moreover, since the given n-code satisfies the power constraint P, the constructed $(2^{nR_1}, \ldots, 2^{nR_k}, n)$ code also satisfies the same power constraint. Hence, we conclude that the rate vector $(R_1, \ldots, R_k)$ is achievable under power constraint P.

APPENDIX B
PROOF OF LEMMA 3

Let $\hat{\mathbf{S}}_i = (\hat{S}_{1i}, \ldots, \hat{S}_{ki})^T$ where $\hat{S}_{ji}$ is given in (19). Then, we have

$$\hat{\mathbf{S}}_1 = 0$$
$$\hat{\mathbf{S}}_i = A\hat{\mathbf{S}}_{i-1} + \mathbf{Y}_{i-1} \quad i = 2, 3, \ldots$$

where $A = \mathrm{diag}(a_1, \ldots, a_k)$ is the same as in (14). Considering the recursion for $\hat{\mathbf{S}}_i$, we can rewrite the system dynamics given in (15) as

$$\mathbf{S}_{i+1} = A^i\mathbf{\Theta} + \hat{\mathbf{S}}_{i+1}$$
$$= A^i(\mathbf{\Theta} - \hat{\mathbf{\Theta}}_i) \qquad (49)$$

where $\hat{\mathbf{\Theta}}_i = (\hat{\Theta}_{1i}, \ldots, \hat{\Theta}_{ki})^T$ and the last equality follows from the decoder rule by which $\hat{\Theta}_{ji} = -a^i\hat{S}_{j(i+1)}$. From (49), the MSE for the message Θ_j at time n is given by

$$D_j^{(n)} = \mathsf{E}(\Theta_j - \hat{\Theta}_{jn})^2 = |a_j|^{-2n}(K_{n+1})_{jj}. \qquad (50)$$

where $K_n := \mathrm{Cov}(\mathbf{S}_n)$ is the covariance matrix of $\mathbf{S}_n$. The achievability of MSE exponent $E_j = \log(|a_j|)$ follows from (50) and the assumption of stability $\limsup_{n \to \infty}(K_n)_{jj} < \infty$. The asymptotic power follows from the fact that $X_i = \pi_i(\mathbf{S}_i)$.

REFERENCES

[1] G. Dueck, "Partial feedback for two-way and broadcast channels," *Inf. Control*, vol. 46, pp. 1–15, 1980.

[2] L. H. Ozarow and C. S. K. Leung, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, pp. 667–671, 1984.

[3] M. A. Wigger and M. Gastpar, "The pre-log of Gaussian broadcast with feedback can be two," in *Proc. IEEE International Symposium on Information Theory*, 2008.

[4] N. Elia, "When Bode meets Shannon: Control oriented feedback communication schemes," *IEEE Trans. Automatic Control*, vol. 47, pp. 1477–1488, 2004.

[5] W. Wu, S. Vishwanath, and A. Arapostathis, "Gaussian interference networks with feedback: Duality, sum capacity and dynamic team problem," in *Proc. 44th Allerton Conference on Communication, Control, and Computation*, 2005.

[6] E. Ardestanizadeh and M. Franceschetti, "Control-theoretic approach to communication with feedback: Fundamental limits and code design," 2010, preprint, http://arxiv.org/abs/1006.5265.

[7] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1423–1438, 2002.

[8] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–I: No bandlimited constraint," *IEEE Trans. Inf. Theory*, vol. 12, pp. 172–182, 1966.

[9] J. P. M. Schalkwijk, "A coding scheme for additive noise channels with feedback–II: Band-limited signals," *IEEE Trans. Inf. Theory*, vol. 12, pp. 183–189, 1966.

[10] G. Chen and S. Hsu, *Linear Stochastic Control Systems*. 2nd ed. CRC Press, 1995.

[11] A. El Gamal and Y.-H. Kim, *Lecture notes on network information theory*. Stanford University and University of California San Diego, 2009.

[12] Y.-H. Kim, "Feedback capacity of stationary Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 56, pp. 57–85, 2010.

[13] E. Ardestanizadeh, M. A. Wigger, T. Javidi, and Y.-H. Kim, "Linear sum capacity for Gaussian multiple access channel with feddback," 2010, preprint, http://arxiv.org/abs/1002.1781.

[14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[15] O. Shayevitz and M. Feder, "Optimal feedback communication via posterior matching," preprint, http://arxiv.org/abs/0909.4828.