# On the Transmission of a Bivariate Gaussian Source Over the Gaussian Broadcast Channel With Feedback

Yonathan Murin, Ben-Gurion University, Israel
Yonatan Kaspi, UC San Diego, USA
Ron Dabora, Ben-Gurion University, Israel
Deniz Gündüz, Imperial College London, UK

***Abstract*—We study the uncoded transmission of a bivariate Gaussian source over a two-user symmetric Gaussian broadcast channel with a unit-delay noiseless feedback (GBCF), assuming that each (uncoded) source sample is transmitted using a *finite* number of channel uses, and that the transmission scheme is linear. We consider three transmission schemes: The scheme of Ardestanizadeh et al., which is based on linear quadratic Gaussian (LQG) control theory, the scheme of Ozarow and Leung (OL), and a novel scheme derived in this work designed using a dynamic programing (DP) approach. For the LQG scheme we characterize the minimal number of channel uses needed to achieve a specified mean-square error (MSE). For the OL scheme we present lower and upper bounds on the minimal number of channel uses needed to achieve a specified MSE, which become tight when the signal-to-noise ratio approaches zero. Finally, we show that for any fixed and finite number of channel uses, the proposed DP scheme achieves MSE lower than the MSE achieved by either the LQG or the OL schemes.**

## I. Introduction

We study the transmission of a bivariate Gaussian source over a two-user Gaussian broadcast channel (GBC) with noiseless causal feedback (NCF), referred to as the GBCF. Motivated by applications with strict delay constraints, we focus on linear uncoded transmission schemes, namely, schemes that do not encode over sequences of source symbols. We further assume that a *finite* number of channel symbols is used for the transmission of each source symbol, and aim at characterizing the minimal number of channel uses required to achieve a target mean-square error (MSE).

The capacity region of the GBCF is not known; however, by extending the Schalkwijk-Kailath (SK) scheme of [1] to two-user GBCFs, the work [2] showed that NCF can enlarge the capacity region of the GBC. In this work we refer to the linear transmission scheme developed in [2] as the *OL scheme*. This scheme was later extended to GBCFs with more than two users and to Gaussian interference channels with NCF (GICFs) in [3]. Recently, in [4], we extended the OL scheme by using estimators with memory instead of the memoryless estimators used in the original OL scheme of [2].

A different approach for channel coding over the GBCF was developed based on control theory. Such a transmission scheme for the two-user GBCF with independent noises was presented in [5], where it was shown to achieve rate pairs outside the achievable rate region of the OL scheme. The work [6] used linear quadratic Gaussian (LQG) control to develop a scheme without requiring the independent noises assumption in [5], and also presented a linear transmission scheme for GBCFs with more than two users. We refer to this scheme as the *LQG scheme*. Recently, [7] showed that for the two-user GBCF with independent noises that have the same variance, the LQG scheme achieves the maximal sum-rate among all possible linear-feedback schemes.

GBCFs and GICFs were also studied in [8] which presented a transmission scheme whose sum-rate approaches the full-cooperation bound as the signal-to-noise ratio (SNR) increases to infinity. Very recently, [9] showed that the capacity region of the GBCF with independent noises and only a common message cannot be achieved by linear feedback schemes.

While [2]–[9] studied channel coding for the GBCF, in [10] we considered joint source-channel coding, i.e., the transmission of correlated sources over the GBCF using a finite number of channel uses for sending each source pair sample. In particular, in [10] we applied both the OL and LQG schemes to the transmission of a pair of correlated Gaussian sources over the GBCF and derived bounds on the number of channel uses needed to achieve a target MSE pair. One of the contributions of the present work is improving upon the bounds presented in [10]. Lastly, we note that [11] studied the transmission of correlated Gaussian sources over the two-user multiple-access channel (MAC) with NCF, and established an upper bound on the energy-distortion tradeoff for the symmetric scenario.

***Main Contributions*:** We study the transmission of a bivariate Gaussian source over the symmetric GBCF with correlated noises, focusing on *linear and memoryless transmission schemes*. In such schemes the transmitted signal at any time index is restricted to be a linear combination of the channel outputs and the encoder state (parameters) at the previous time index. First, we consider the LQG scheme and provide an *explicit and exact characterization* of the minimal number of channel uses required to achieve a target MSE at each receiver (note that the bounds derived in [10] were not exact). We then proceed to study the OL scheme and derive upper and lower bounds on the minimal number of channel uses required to achieve a target MSE at each receiver. *These bounds become tight as the SNR approaches zero*, and improve upon the bounds presented in [10]. Finally, we present a new linear and memoryless transmission scheme designed using a dynamic programing (DP) approach [12], to which we refer as the *DP scheme*. For a finite number of channel uses, the newly derived DP scheme is shown to achieve an MSE lower than both OL and LQG. Yet, the analysis of this scheme is highly complicated and finding the coefficients of this scheme becomes computationally infeasible as the number of channel uses becomes large.

This work was supported by Israel Science Foundation under grant 396/11.

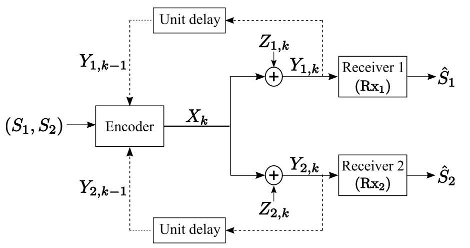


Fig. 1: The two-user GBCF with correlated sources $(S_1, S_2)$. $\hat{S}_1$ and $\hat{S}_2$ are the reconstructions of $S_1$ and $S_2$, respectively, after $K$ transmissions.

The rest of this paper is organized as follows: The problem formulation is introduced in Section II. The LQG and OL schemes are studied in Sections III and IV, respectively. The DP scheme is derived in Section V, and discussion along with a numerical example are given in Section VI.

***Notations***: We use capital letters to denote random variables, e.g., $X$, boldface letters to denote column vectors, e.g., $\mathbf{X}$, and sans-serif font to denote matrices, e.g., $\mathsf{M}$. We use $\mathbb{E}\{\cdot\}$, $(\cdot)^T$, $\log(\cdot)$ and $\Re$ to denote expectation, transpose, natural basis logarithm, and the set of real numbers, respectively. Lastly, we let $\text{sgn}(x)$ denote the sign of $x$, where $\text{sgn}(0) \triangleq 1$, and denote the ceiling function of $x$ by $\lceil x \rceil$.

## II. Problem Formulation

The two-user GBCF is depicted in Fig. 1. All the signals are real. The encoder observes a realization of a bivariate Gaussian source denoted by $\mathbf{S} = [S_1, S_2]^T$, and is required to send $S_i, i = 1, 2$, to the $i$'th receiver, denoted by $\text{Rx}_i$. Let $S_i \sim \mathcal{N}(0, \sigma_s^2)$, $\mathbb{E}\{S_1 S_2\} = \rho_s \sigma_s^2$, $|\rho_s| < 1$, and $\mathsf{Q}_s \triangleq \mathbb{E}\{\mathbf{S}\mathbf{S}^T\}$. Each bivariate source symbol is transmitted using $K$ channel uses. The channel outputs at each receiver at time $k, k=1, 2, \ldots, K$, are given by $Y_{i,k} = X_k + Z_{i,k}, i = 1, 2$, where $[Z_{1,k}, Z_{2,k}]^T$ are jointly Gaussian i.i.d over time $k$, independent of $\mathbf{S}$, and $\mathbb{E}\{Z_i^2\} = \sigma_z^2, \mathbb{E}\{Z_{1,k} Z_{2,k}\} = \rho_z \sigma_z^2, |\rho_z| < 1$. Letting $\mathbf{B} \triangleq [1, 1]^T, \mathbf{Y}_k \triangleq [Y_{1,k}, Y_{2,k}]^T$, and $\mathbf{Z}_k \triangleq [Z_{1,k}, Z_{2,k}]^T$, the received signal can be written as:

$$\mathbf{Y}_k = \mathbf{B} X_k + \mathbf{Z}_k, \quad k = 1, 2, \ldots, K. \tag{1}$$

$\text{Rx}_i, i=1, 2$, uses its $K$ channel outputs, $\mathbf{Y}_{i,1}^K$, to estimate $S_i$: $\hat{S}_i = \phi_i(\mathbf{Y}_{i,1}^K), \phi_i : \Re^K \rightarrow \Re$. The encoder maps the observed source and the received NCF into a channel input at time $k$ via $X_k = f_k(S_1, S_2, \mathbf{Y}_{1,1}^{k-1}, \mathbf{Y}_{2,1}^{k-1}), f_k : \Re^{2k} \rightarrow \Re$. The transmitted signal is subject to a per-symbol average power constraint defined as:

$$\mathbb{E}\left\{X_k^2\right\} \leq P, \quad \forall k = 1, 2, \ldots, K. \tag{2}$$

For a specific set of parameters $(P, \sigma_z^2, \rho_z, \sigma_s^2, \rho_s)$, we define a $(D, K)$ code, $0 < D \leq \sigma_s^2$, to be a collection of $K$ encoding functions, each satisfying (2), and two decoding functions such that:

$$\mathbb{E}\{(S_i - \hat{S}_{i,K})^2\} \leq D, \quad i = 1, 2. \tag{3}$$

*Our objective is to identify, for a given D, the minimal number of channel uses $K$ for which a $(D, K)$ code exists.*

In the sequel, we let $K_{\text{SCHEME}}$ denote the minimal number of channel uses required to achieve a target MSE $D$ using the scheme "SCHEME". Next, we consider transmission based on the LQG and OL schemes, specialized to the symmetric setting.

## III. Transmission Based on the LQG Scheme

### A. A Brief Overview of the LQG Scheme

We use the LQG scheme of [6], after adaptation to the transmission of Gaussian sources as described in [10]. Consider a two-dimensional unstable dynamical system which is stabilized by a controller observing the entire system state vector at time $k$, $\mathbf{U}_k = [U_{1,k}, U_{2,k}]^T$. The controller outputs a scalar signal $X_k$, which is corrupted by additive Gaussian noises. The dynamics of the system is given by:

$$\mathbf{U}_1 = \mathbf{S}, \;\; \mathbf{U}_k = \mathsf{A}\mathbf{U}_{k-1} + \mathbf{Y}_{k-1}, \;\; k = 2, 3, \ldots, K, \tag{4}$$

where $\mathbf{Y}_k$ is given in (1). For the symmetric setting we let $\mathsf{A} = \text{diag}(a, -a), a \in \Re, |a| > 1$.

***Encoding***: In the corresponding communications problem, the encoder consists of the system given in (4) and of the controller. At each time index, the encoder recursively computes $\mathbf{U}_k$ and transmits $X_k$. In this work we use the linear controller presented in [6, Lemma 4], which is given by $X_k = -\mathbf{C}^T \mathbf{U}_k$, where $\mathbf{C} = [c, -c]^T$ is such that all the eigenvalues of the matrix $\mathsf{M} \triangleq \mathsf{A} - \mathbf{B}\mathbf{C}^T$ have magnitudes smaller than 1. The coefficients $a$ and $c$ are determined as in [6, Lemma 4].

***Decoding***: We use the decoder of [10, Thm. 1], which first estimates $U_{i,k}$ via $\hat{U}_{i,1} = 0, \hat{U}_{i,k} = (-1)^{i-1} \cdot a \cdot \hat{U}_{i,k-1} + Y_{i,k-1}$, $k = 2, 3, \ldots, K$, and then applies minimum MSE (MMSE) estimation of $S_i$ from $\hat{U}_{i,k+1}$ via [10, Eq. (8)]. The LQG scheme is terminated after $K$ channel uses, where $K$ is chosen such that the target MSE $D$ is achieved at each receiver.

### B. Finite Horizon Analysis of the LQG Scheme

Let $P_k$ denote the instantaneous average transmission power. Note that in the LQG scheme $P_k$ changes over time and converges to $P$ as $k \rightarrow \infty$. In fact, in the LQG scheme $P_k$ may be larger than $P$ and (2) may not be satisfied. Therefore, given $P, \sigma_z^2$ and $\rho_z$, some sources must be scaled before transmission in order for the constraint (2) to be satisfied. Let $[\lambda, -\lambda]^T$ be the eigenvalues of $\mathsf{M}$, and $\mathsf{V} \triangleq \begin{bmatrix} v_1 & v_2 \\ v_2 & v_1 \end{bmatrix}$ be a matrix whose columns are the corresponding eigenvectors of $\mathsf{M}$. Furthermore, define $\mu_1 = 2c^2\sigma_s^2(1-\rho_s)$, $\mu_2 = 2c^2\sigma_s^2(1+\rho_s)a^4$, and $\mu_3 = \frac{2c^2\sigma_z^2((1-\rho_z)\lambda^2 + (1+\rho_z)a^4)}{1-\lambda^4}$. The following proposition characterizes the sources which satisfy (2):

*Proposition* 1. The LQG scheme satisfies the per-symbol average power constraint (2) *iff* $\mu_1 \leq P$ and $\mu_2 \leq \mu_3$.

*Proof outline:* We show that $P_k = P + (\mu_1 - P)\lambda^{2(k-1)}$, for odd $k$'s, and that $P_k = P + (\mu_2 - \mu_3)\lambda^{2(k-1)}$, for even $k$'s. Since $\lambda < 1$, it follows that (2) is satisfied iff $\mu_1 \leq P$ and $\mu_2 \leq \mu_3$. The detailed proof can be found in [13]. ■

Next, we explicitly characterize $K_{\text{LQG}}$. First, we define the function $\Phi(\varsigma, \rho) \triangleq \frac{\varsigma^2\left(\left(v_1^2 + v_2^2 - 2\rho v_1 v_2\right)^2 + 4(1-\rho^2)v_1^2 v_2^2\right)}{\det^2(\mathsf{V})}$. We also define the terms: $\Psi_0 \triangleq \frac{\sigma_z^2 + \lambda^2 \Phi(\sigma_z, \rho_z)}{1-\lambda^4}$, $\Psi_1 \triangleq \frac{\Phi(\sigma_z, \rho_z) + \sigma_z^2 \lambda^2}{1-\lambda^4}$, $\Gamma_s = \frac{\sigma_s^2(v_1^2 + v_2^2 - 2\rho_s v_1 v_2)}{v_1^2 - v_2^2}$, $\Upsilon_0 = \Psi_0 \cdot (D - \sigma_s^2) - D\sigma_s^2$, $\Upsilon_1 = \Psi_0 \cdot (\sigma_s^2 - D) + 2D\sigma_s^2$, $\Upsilon_2 = (\Phi(\sigma_s, \rho_s) - \Psi_1)(\sigma_s^2 - D) - \Gamma_s^2$, and $\Upsilon_3 = \Psi_0 \cdot (\sigma_s^2 - D) + 2D\Gamma_s$. Finally, let $n$ be a positive integer and define the functions $f^{(e)}(n) \triangleq 2\left\lceil \frac{n}{2} \right\rceil$, and $f^{(o)}(n) \triangleq 2\left\lceil \frac{n-1}{2} \right\rceil + 1$. The following theorem characterizes $K_{\text{LQG}}$:

$$B_\rho(P) \triangleq \frac{(8+\psi_1)\,P^3+24\sigma_z^2P^2+12\sigma_z^4\psi_1P+4\sigma_z^6\left(4\sigma_z^2\psi_1+8\right)}{8\sigma_z^{10}}P^2, \quad B_\alpha(P) \triangleq \frac{P+2\sigma_z^2}{2\sigma_z^6}P^2$$

*Theorem* 1. Let $(x_1^{(e)}, x_2^{(e)})$ and $(x_1^{(o)}, x_2^{(o)})$ denote the roots of the polynomials $P^{(e)}(x) \triangleq \Upsilon_0 x^2+\Upsilon_1 x-D\sigma_s^2$, and $P^{(o)}(x) \triangleq \Upsilon_2 x^2+\Upsilon_3 x-D\sigma_s^2$, respectively. Furthermore, define:

$$x_0^{(e)} \triangleq \begin{cases} \min\{x_1^{(e)}, x_2^{(e)}\}, & \frac{-\Upsilon_1^2}{4D\sigma_s^2} \le \Upsilon_0 \\ a^{-4}, & \text{otherwise.} \end{cases}$$

$$x_0^{(o)} \triangleq \begin{cases} \min\{x_1^{(o)}, x_2^{(o)}\}, & \frac{-\Upsilon_3^2}{4D\sigma_s^2} \le \Upsilon_2 \le 0, \\ a^{-2}, & \Upsilon_2 < \frac{-\Upsilon_3^2}{4D\sigma_s^2}, \\ \max\{x_1^{(o)}, x_2^{(o)}\}, & \text{otherwise.} \end{cases}$$

Then, $K_{\text{LQG}}$ is given by:

$$K_{\text{LQG}}=\min\left\{f^{(e)}\left(\left\lceil -\frac{\log x_0^{(e)}}{2\log|a|}\right\rceil\right), f^{(o)}\left(\left\lceil -\frac{\log x_0^{(o)}}{2\log|a|}\right\rceil\right)\right\}. \quad (5)$$

*Proof outline:* Recall that the scheme is terminated when $D\ge\mathbb{E}\{(S_i-\hat{S}_{i,k})^2\}$. We show that for odd $k$'s this condition can be formulated as $P^{(o)}(x)\le 0$, where $x=a^{-2k}$. Similar observation holds for even $k$'s with $P^{(e)}(x)\le 0$. Then, if $\Upsilon_2\ge\frac{-\Upsilon_3^2}{4D\sigma_s^2}$, we choose $x_0^{(o)}$ among the possible two real roots of $P^{(o)}(x)$, based on the concavity of $P^{(o)}(x)$. Otherwise we set $x_0^{(o)}=a^{-2}$ which results in $K_{\text{LQG}}=1$. Similarly we find $x_0^{(e)}$, while noting that $\Upsilon_0\le 0$. Finally, $x_0^{(o)}$ and $x_0^{(e)}$ are translated to $K_{\text{LQG}}$ as in (5). The detailed proof can be found in [13]. ■

## IV. Transmission Based on the OL Scheme

### A. A Brief Overview of the OL Scheme

We use the OL scheme of [2], adapted to the transmission of Gaussian sources as described in [10]. In the OL scheme, each receiver recursively estimates its intended source. The transmitter, using the NCF, tracks the estimation errors at the receivers and sends a linear combination of these errors. The scheme is terminated after $K$ channel uses, where $K$ is chosen such that the target MSE $D$ is achieved at each receiver.

***Setup*:** Let $\hat{S}_{i,k}$ be the estimate of $S_i$ at Rx$_i$ after receiving the $k$'th channel output $Y_{i,k}$. Let $\epsilon_{i,k}\triangleq\hat{S}_{i,k}-S_i$ be the estimation error after $k$ transmissions, and define $\hat{\epsilon}_{i,k-1}\triangleq\hat{S}_{i,k-1}-\hat{S}_{i,k}$. Hence, we have $\epsilon_{i,k}=\epsilon_{i,k-1}-\hat{\epsilon}_{i,k-1}$. Further define $\alpha_{i,k}\triangleq\mathbb{E}\{\epsilon_{i,k}^2\}$ to be the MSE at Rx$_i$ after $k$ transmissions. Note that as we consider the symmetric setting, then $\alpha_{1,k}=\alpha_{2,k}\triangleq\alpha_k$. Finally, define $\rho_k\triangleq\frac{\mathbb{E}\{\epsilon_{1,k}\epsilon_{2,k}\}}{\alpha_k}$ which is the correlation between the estimation errors.

***Encoding*:** Set $\hat{S}_{i,0}=0$ and $\epsilon_{i,0}=S_i$, thus, $\rho_0=\rho_s$. Furthermore, let $\Psi_k\triangleq\sqrt{\frac{P}{2(1+|\rho_k|)}}$. At the $k$'th channel use, $1\le k\le K$, the transmitter sends $X_k=\frac{\Psi_{k-1}}{\sqrt{\alpha_k}}(\epsilon_{1,k-1}+\epsilon_{2,k-1}\text{sgn}(\rho_{k-1}))$, and the corresponding channel outputs are given by (1).

***Decoding*:** Each receiver computes $\hat{\epsilon}_{i,k-1}, i=1,2$, based only on $Y_{i,k}$: $\hat{\epsilon}_{i,k-1}=\frac{\mathbb{E}\{\epsilon_{i,k-1}Y_{i,k}\}}{\mathbb{E}\{Y_{i,k}^2\}}Y_{i,k}$, see [2, pg. 669] for the explicit expressions. Let $\pi_0\triangleq P+\sigma_z^2$, $\Sigma\triangleq P+\sigma_z^2(2-\rho_z)$ and $\nu=\sigma_z^4(1-\rho_z)^2$. Then, the instantaneous MSE $\alpha_k$ is given by the recursive expression [2, Eq. (5)]:

$$\alpha_k=\alpha_{k-1}\frac{\sigma_z^2+\Psi_{k-1}^2(1-\rho_{k-1}^2)}{\pi_0}, \quad (6)$$

where the recursive expression for $\rho_k$ is given by [2, Eqn. (7)]:

$$\rho_k=\frac{(\rho_z\sigma_z^2\Sigma+\nu)\rho_{k-1}-\Psi_{k-1}^2\Sigma(1-\rho_{k-1}^2)\text{sgn}(\rho_{k-1})}{\pi_0(\sigma_z^2+\Psi_{k-1}^2(1-\rho_{k-1}^2))}. \quad (7)$$

*Remark* 1. In the above OL scheme we do not apply the initialization procedure described in [2, pg. 669]. Instead, we set $\epsilon_{i,0}=S_i$ and $\rho_0=\rho_s$, thus, taking advantage of the correlation among the sources.

### B. Finite Horizon Analysis of the OL Scheme

The instantaneous MSE in (6), $\alpha_k$, depends on $\rho_k$, thus, an explicit characterization of $K_{\text{OL}}$ requires explicitly characterizing $\rho_k$. However, this is highly complicated as $\rho_k$ is defined recursively in (7). In the following, we obtain upper and lower bounds on $K_{\text{OL}}$, such that the ratios between the values of the bounds and $K_{\text{OL}}$ approach 1 when the SNR approach 0, i.e., $\frac{P}{\sigma_z^2}\to 0$. In this regime these bounds improve upon those derived in [10, Thm. 4].

We follow the approach of [11, Thm. 7] and approximate the temporal behavior of $\rho_k$ and $\alpha_k$ based on (6) and (7), at low $\frac{P}{\sigma_z^2}$. We first define the following terms: $\psi_1\triangleq 2|\rho_z|+5(1-\rho_z)$, $\psi_2\triangleq\frac{\min\{2-\rho_z,2(1-\rho_z)\}}{2\sigma_z^2}$ and $\psi_3\triangleq\max\left\{\frac{1-\rho_z}{(2-\rho_z)^2},\frac{1+\rho_z}{4(1-\rho_z)^2}\right\}$. We further define, at the top of the page, $B_\rho(P)$ and $B_\alpha(P)$ which constitute upper bounds on the approximation errors of $\rho_{k+1}-\rho_k$ and $\frac{\alpha_{k+1}}{\alpha_k}$, respectively. Lastly, define $\bar{K}_0(P)\triangleq\frac{\rho_s}{P\psi_2(\rho_z)-B_\rho(P)}$, $\bar{\rho}(P)\triangleq\frac{P(3-\rho_z)^2}{8\sigma_z^2}+B_\rho(P)$, and the terms $F_j(P), j=1,2,\ldots,9$:

$$F_1(P)\triangleq\bar{K}_0(P)\psi_3\cdot\left(\frac{(3-\rho_z)^2P}{8\sigma_z^2}+B_\rho(P)\right)^2,$$

$$F_2(P)\triangleq\bar{K}_0(P)\frac{B_\rho(P)}{2\psi_2\sigma_z^2},$$

$$F_3(P)\triangleq\bar{K}_0(P)\frac{1}{(1-\rho_z)^2}\left(\frac{(3-\rho_z)^2P}{8\sigma_z^2}+B_\rho(P)\right)^2,$$

$$F_4(P)\triangleq\frac{\bar{K}_0(P)B_\rho(P)}{1-\rho_z},\ F_5(P)\triangleq\bar{K}_0(P)B_\alpha(P),$$

$$F_6(P)\triangleq\sum_{i=3}^{5}F_i(P),$$

$$F_7(P)\triangleq\frac{P}{2\sigma_z^2}\left(-1+\left(\bar{\rho}(P)+\frac{2\sigma_z^2}{P}B_\alpha(P)\right)\right),$$

$$F_8(P)\triangleq\frac{P}{2\sigma_z^2}\left(-1-\left(\bar{\rho}(P)+\frac{2\sigma_z^2}{P}B_\alpha(P)\right)\right),$$

$$F_9(P)\triangleq\frac{2\sigma_z^2}{P}\left(F_1(P)+F_2(P)\right).$$

Let $\rho_*^{\text{lb}}(D)\triangleq 2-\rho_z+\frac{\sigma_s^2}{D}(\rho_z+\rho_s-2)\,e^{F_6(P)}$ and $\rho_*^{\text{ub}}(D)\triangleq 2-\rho_z+\frac{\sigma_s^2}{D}(\rho_z+\rho_s-2)\,e^{-F_6(P)}$. The following theorem provides upper and lower bounds on $K_{\text{OL}}$:

*Theorem* 2. Let $P$ satisfy the conditions: $\bar{\rho}(P)+\frac{2\sigma_z^2}{P}B_\alpha(P)<1$ and $B_\rho(P)<P\psi_2$. Further let $D_0^{\text{ub}}\triangleq\frac{\sigma_s^2(2-\rho_z-\rho_s)e^{F_6(P)}}{2-\rho_z}$ and

$D_0^{\text{lb}} \triangleq \frac{\sigma_s^2(2-\rho_z-\rho_s)e^{-F_6(P)}}{2-\rho_z}$. Then, it follows that $K_{\text{OL}}^{\text{lb}} \leq K_{\text{OL}} \leq K_{\text{OL}}^{\text{ub}}$, where, for $D \geq D_0^{\text{ub}}$:

$$K_{\text{OL}}^{\text{ub}} = \frac{2\sigma_z^2}{P(3-\rho_z)} \log\left(\frac{(2-\rho_z-\rho_*^{\text{lb}}(D))(1+\rho_s)}{(2-\rho_z-\rho_s)(1+\rho_*^{\text{lb}}(D))}\right) + F_9(P), \quad \text{(8a)}$$

$$K_{\text{OL}}^{\text{lb}} = \frac{2\sigma_z^2}{P(3-\rho_z)} \log\left(\frac{(2-\rho_z-\rho_*^{\text{ub}}(D))(1+\rho_s)}{(2-\rho_z-\rho_s)(1+\rho_*^{\text{ub}}(D))}\right) - F_9(P), \quad \text{(8b)}$$

and for $D < D_0^{\text{lb}}$:

$$K_{\text{OL}}^{\text{ub}} = \left(\log\left(\frac{D(2-\rho_z-\bar{\rho}(P,\rho_z))}{\sigma_s^2(2-\rho_z-\rho_s)}\right) - F_6(P)\right)\frac{1}{F_7(P)} + \frac{2\sigma_z^2}{P(3-\rho_z)} \log\left(\frac{(2-\rho_z-\rho_*^{\text{lb}}(D_0^{\text{ub}}))(1+\rho_s)}{(2-\rho_z-\rho_s)(1+\rho_*^{\text{lb}}(D_0^{\text{ub}}))}\right) + F_9(P), \quad \text{(8c)}$$

$$K_{\text{OL}}^{\text{lb}} = \left(\log\left(\frac{D(2-\rho_z+\bar{\rho}(P,\rho_z))}{\sigma_s^2(2-\rho_z-\rho_s)}\right) + F_6(P)\right)\frac{1}{F_8(P)} + \frac{2\sigma_z^2}{P(3-\rho_z)} \log\left(\frac{(2-\rho_z-\rho_*^{\text{ub}}(D_0^{\text{lb}}))(1+\rho_s)}{(2-\rho_z-\rho_s)(1+\rho_*^{\text{ub}}(D_0^{\text{lb}}))}\right) - F_9(P). \quad \text{(8d)}$$

*Proof:* The proof can be found in [13]. ■

*Remark* 2. Let $\rho_s \geq 0$ (otherwise replace $S_1$ with $-S_1$). From [2, pg. 669] it follows that if $\rho_k > 0$ then $\rho_{k+1} < \rho_k$. This implies that $\rho_k$ decreases (with $k$) until it crosses zero. Let $K_0 \triangleq \min\{k : \rho_{k+1} < 0\}$ be the largest time index $k$ for which $\rho_k \geq 0$. In the proof of Thm. 2 we show that, for sufficiently small $\frac{P}{\sigma_z^2}$, $|\rho_k| \leq \bar{\rho}(P), \forall k \geq K_0$. Hence, $\rho_k$ decreases until time $K_0$ and then has a bounded magnitude. This implies that the behavior of $\alpha_k$ is different in the regimes $k \leq K_0$ and $k > K_0$. Let $D_0$ be the MSE after $K_0$ channel uses; then $D_0^{\text{ub}}$ and $D_0^{\text{lb}}$ constitute upper and lower bounds on $D_0$, respectively. This leads to the two cases in Thm. 2: (8a)–(8b) correspond to the case of $K_{\text{OL}} < K_0$, while (8c)–(8d) correspond to the case $K_0 < K_{\text{OL}}$. Lastly we note that $\bar{K}_0(P)$ constitutes an upper bound on $K_0$.

Next, we demonstrate these results via a numerical example.

### C. A Numerical Example

Let $K_{\text{OL}}^{\text{ub},[10]}$ and $K_{\text{OL}}^{\text{lb},[10]}$ denote the upper and lower bounds on $K_{\text{OL}}$ presented in [10, Thm. 4]. We consider a GBCF with $\sigma_s^2 = 1, \rho_s = 0.9, \sigma_z^2 = 1, \rho_z = 0.7$, $D = 0.1$, and two possible values of $P : P_1 = 10^{-4}$ and $P_2 = 10^{-6}$. Both $P_1$ and $P_2$ satisfy the conditions of Thm. 2. Table I details $K_{\text{OL}}, K_{\text{OL}}^{\text{ub},[10]}, K_{\text{OL}}^{\text{lb},[10]}, K_{\text{OL}}^{\text{ub}}$ and $K_{\text{OL}}^{\text{lb}}$ for $P_1$ and $P_2$:

| $P$ | $K_{\text{OL}}$ | $K_{\text{OL}}^{\text{ub},[10]}$ | $K_{\text{OL}}^{\text{lb},[10]}$ | $K_{\text{OL}}^{\text{ub}}$ | $K_{\text{OL}}^{\text{lb}}$ |
|---|---|---|---|---|---|
| $10^{-4}$ | 38311 | 46058 | 23026 | 38659 | 37960 |
| $10^{-6}$ | 3830913 | 4605176 | 2302586 | 3831260 | 3830563 |

TABLE I: Values of $K_{\text{OL}}$ and the proposed bounds.

Note that $K_{\text{OL}}^{\text{ub}} - K_{\text{OL}}$ is approximately the same for both $P_1$ and $P_2$, while $K_{\text{OL}}$ increases by approximately $10^2$. This holds for $K_{\text{OL}} - K_{\text{OL}}^{\text{lb}}$ as well. Combined with the fact that $B_\rho(P), B_\alpha(P) \to 0$ when $P \to 0$, this implies that $K_{\text{OL}}^{\text{ub}}/K_{\text{OL}}$ approaches 1 when the SNR approaches 0. Table I also indicates that the ratio $K_{\text{OL}}^{\text{ub},[10]}/K_{\text{OL}} \approx 1.2$ for both $P_1$ and $P_2$, yet, these bounds hold for any SNR.

The LQG and OL schemes described above are linear and memoryless. Next, we use DP to design a linear and memoryless transmission scheme, which outperforms OL and LQG, under the per-symbol average power constraint (2).

## V. Transmission Via the DP Scheme

### A. Problem Formulation - Revisited

We examine a problem complimentary to the one formulated in Section II: The number of channel uses is fixed to $K$, and we denote the MSE after $K$ channel uses by $D_K$. Our objective is to find a linear and memoryless transmission scheme which achieves the minimal MSE $D_{K,\text{min}}$ at each receiver. Let $\epsilon_{i,k-1}$ denote the information sent to $\text{Rx}_i$ at time $k$. As we focus on linear and memoryless transmission schemes, $\epsilon_{i,k}$ is of the form:

$$\epsilon_{i,k} = \beta_{i,k}\left(\epsilon_{i,k-1} - b_{i,k}Y_{i,k}\right), \quad \beta_{i,k}, b_{i,k} \in \Re. \quad (9)$$

Following [3], we let $m_k \in \{1, -1\}$ be a modulation coefficient. As we consider the symmetric setting, then $|\beta_{1,k}| = |\beta_{2,k}|$, $|b_{1,k}| = |b_{2,k}|$ and we let $b_{1,k} = b_k$. This leads to the following structure of $X_{k+1}$:

$$X_{k+1} = d_k\big((\epsilon_{1,k-1} - b_k Y_{1,k}) + m_k(\epsilon_{2,k-1} - m_{k-1} b_k Y_{2,k})\big), \quad (10)$$

where $d_k$ is chosen to minimize $D_K$ under the constraint $P_k \leq P$. In [13] we show that choosing $P_k = P$ is optimal. Next, we let $\alpha_k \triangleq \mathbb{E}\{\epsilon_{i,k}^2\}, i = 1, 2$, be the MSE after $k$ channel uses, and $r_k \triangleq \mathbb{E}\{\epsilon_{1,k}\epsilon_{2,k}\}$. Thus, $d_k$ is given by $d_k = \sqrt{\frac{P}{2(\alpha_k + m_k r_k)}}$. Finally, we initialize the transmission scheme via $\epsilon_{i,0} = S_i$, $\alpha_0 = \sigma_s^2$, and $r_0 = \rho_s \sigma_s^2$.

Our objective is to minimize $D_K$, over all possible vectors of estimation coefficients $\mathbf{b} = [b_1, b_2, \ldots, b_K] \in \Re^K$, and all possible vectors of modulation coefficients $\mathbf{m} = [m_0, m_1, \ldots, m_{K-1}] \in \{1, -1\}^K$. As the joint minimization of $D_K$ over $\mathbf{b}$ and $\mathbf{m}$ is complicated, we define $\alpha_{K,\text{min}}(\mathbf{m})$ to be the minimal achievable MSE after $K$ channel uses *for a given modulation vector* $\mathbf{m}$. We first *calculate* $\alpha_{K,\text{min}}(\mathbf{m})$, thereby arriving at the optimization problem:

$$D_{K,\text{min}} = \min_{\mathbf{m} \in \{1,-1\}^K} \alpha_{K,\text{min}}(\mathbf{m}), \quad (11)$$

which can be solved by searching over all possible $2^K$ modulation vectors. We refer to the transmission scheme which uses the optimal $\mathbf{b}$ and $\mathbf{m}$ as the *DP scheme*. Next, we present the algorithm for finding the minimizing $\mathbf{b}$ and the minimal $\alpha_{K,\text{min}}(\mathbf{m})$ for a given $\mathbf{m}$.

### B. Calculating the Minimizing $\mathbf{b}$ and $\alpha_{K,\text{min}}(\mathbf{m})$

Let $\mathbf{m}$ be a given modulation vector. Then, (1), (9), and (10) imply that $\alpha_k$ and $r_k$ are given by:

$$\alpha_k = \alpha_{k-1} + b_k^2 \cdot (P + \sigma_z^2) - b_k\sqrt{2P\left(\alpha_{k-1} + m_{k-1} r_{k-1}\right)} \quad \text{(12a)}$$

$$r_k = r_{k-1} + b_k^2 m_{k-1} \cdot (P + \rho_z \sigma_z^2) - b_k m_{k-1}\sqrt{2P\left(\alpha_{k-1} + m_{k-1} r_{k-1}\right)}. \quad \text{(12b)}$$

Therefore, $(\alpha_{k-1}, r_{k-1})$ can be treated as a state variable, which, given $b_k$ and $\mathbf{m}$, evolves deterministically. Thus, finding $\alpha_{K,\text{min}}(\mathbf{m})$ can be cast as a DP with state $(\alpha_{k-1}, r_{k-1})$, actions $b_k$, and cost function $\alpha_K$. Note that with this formulation $b_k$ is a function of only $(\alpha_{k-1}, r_{k-1})$, and the last action $b_K$ is the MMSE estimation coefficient for estimating $\epsilon_{1,K-1}$ from $Y_{1,K}$. Finally, the DP solution [12, Ch. 4] implies that $\alpha_k$ can be written as $\alpha_k = \eta_{k-1}\alpha_{k-1} + \theta_{k-1} m_{k-1} r_{k-1}$, where the sequences $\eta_k$ and $\theta_k, k = 1, 2, \ldots, K-1$, are obtained using backwards recursion (in time). The minimizing $\mathbf{b}$ and the sequences $\eta_k$ and $\theta_k$ are given in the following theorem:

*Theorem* 3. For a fixed $\mathbf{m}$, the sequences $\eta_k$ and $\theta_k, k = 1, 2, \ldots, K-1$, are defined through the backwards recursion (in time):

$$\eta_{k-1} = \eta_k - \tau(\eta_k, \theta_k, m_k, m_{k-1}) \tag{13a}$$

$$\theta_{k-1} = \theta_k m_k m_{k-1} - \tau(\eta_k, \theta_k, m_k, m_{k-1}), \tag{13b}$$

where $\tau(\eta_k, \theta_k, m_k, m_{k-1}) = \frac{P(\eta_k + \theta_k m_k m_{k-1})^2}{2(\eta_k(P+\sigma_z^2) + \theta_k m_k m_{k-1}(P+\rho_z \sigma_z^2))}$, $\eta_{K-1} = \left(1 - \frac{P}{2(P+\sigma_z^2)}\right)$, and $\theta_{K-1} = -\frac{P}{2(P+\sigma_z^2)}$. Furthermore, for $k=1, 2, \ldots, K-1$ the coefficients $b_k$ are given by $b_k = \sqrt{\frac{P(\alpha_{k-1} + m_{k-1} r_{k-1})}{2}} \frac{\eta_k + \theta_k m_k m_{k-1}}{\eta_k (P+\sigma_z^2) + \theta_k m_k m_{k-1}(P+\rho_z \sigma_z^2)}$, and $b_K = \sqrt{\frac{P(\alpha_{K-1} + m_{K-1} r_{K-1})}{2(P+\sigma_z^2)^2}}$. The corresponding MSE at time $K$ is the minimal MSE given $\mathbf{m}$.

*Proof:* The proof can be found in [13]. ■

Thm. 3 can be used for calculating the optimal $\mathbf{b}$ for a given $\mathbf{m}$. The procedure is summarized in Alg. 1:

**Algorithm 1** Calculating the Minimizing $\mathbf{a}$ and $\alpha_{K,\min}(\mathbf{m})$

$\eta_{K-1} \leftarrow \left(1 - \frac{P}{2(P+\sigma_z^2)}\right)$, $\theta_{K-1} \leftarrow -\frac{P}{2(P+\sigma_z^2)}$
Calculate the backwards recursions (13)
$\alpha_0 \leftarrow \sigma_s^2$, $r_0 \leftarrow \rho_s$
**for** $k = 1, 2, \ldots, K$ **do**
    Calculate $b_k$ as in Thm. 3
    Calculate $\alpha_k, r_k$ via (12)
**end for**

## VI. A Discussion and a Numerical Example

*Remark* 3. As we aim at minimizing $\alpha_{K,\min}(\mathbf{m})$, $b_K$ is the MMSE estimation coefficient for estimating $\epsilon_{1,K-1}$ from $Y_{1,K}$, given $\mathbf{m}$. On the other hand, for $k < K$, using the MMSE estimation coefficient is not necessarily optimal as the $b_k$'s affect the future time indices. With this observation, it is clear why the OL scheme, which applies the MMSE estimator for all $k$'s, is not optimal in the MMSE sense, even among the memoryless linear transmission schemes.

*Remark* 4. In [13] we show that choosing $P_k = P$ in the DP scheme is optimal. Thus, the DP scheme is the optimal scheme (in the sense of minimal $D_K$) among the class of schemes which can be formulated via (9)–(10), subject to the constraint $P_k \leq P$. From [6, Eqs. (36)–(37)] it follows that the LQG scheme can also be formulated via (9)–(10). Hence, we conclude that DP achieves MSE which is at least as low as any LQG scheme which satisfies the per-symbol average power constraint (2), see Prop. 1.

*Remark* 5. Note that any choice of $\mathbf{m}$ will result in an upper bound on $D_{K,\min}$. While finding $D_{K,\min}$ involves searching over all $2^K$ possible $\mathbf{m}$ sequences, the actual search can be shortened at the expense of a possibly larger MSE. Motivated by the alternating sign of $\rho_k$ in the OL and LQG schemes, for $k \to \infty$, (see [6, Eqs. (23), (36)–(37)]), we can choose $\mathbf{m}$ to start alternating signs after some $L \ll K$ channel uses, and essentially search only over $2^L$ sequences. Numerical simulations show that this approach can perform well, as indicated in Fig. 2.

Lastly, we demonstrate our results via a numerical example. Consider the transmission of a bivariate Gaussian source with $\sigma_s^2 = 1$ and $\rho_s = 0.2$, over a symmetric GBCF with $\sigma_z^2 = 1.3$, $\rho_z = 0$, and $P = 1$. For this setting the conditions of Prop. 1 are satisfied. Fig. 2 depicts the MSEs (6), [10, Eq. (9)], and (11), for this scenario, where the left figure depicts the MSEs for $5 \leq k \leq 30$ and the right figure depicts the MSEs for $705 \leq k \leq 730$. It can be observed that DP outperforms both OL and LQG: For $5 \leq K \leq 30$ OL is very close to DP, while LQG obtains MSEs higher by factor of 10. For $705 \leq K \leq 730$ LQG slightly outperforms OL and both perform worse than DP by a factor of 10. This indicates that the *slope of the line corresponding to DP is the same as the slope of LQG*, and both are smaller (and negative) than the slope of OL. This supports the asymptotic results of [6, Sec. V.A].

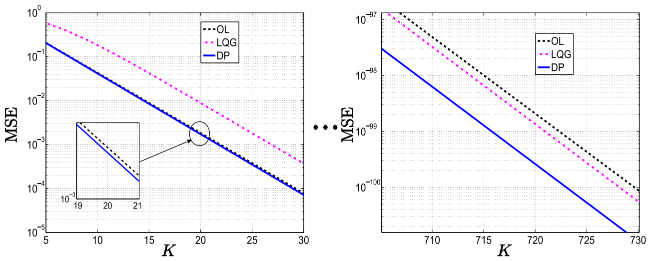


Fig. 2: MSE vs. number of channel uses for $\sigma_s^2 = 1, \rho_s = 0.2, \rho_z = 0, \sigma_z^2 = 1.3$ and $P = 1$. $m_k$ is forced to an alternating sequence starting from $L = 25$.

## References

[1] J. P. M. Schalkwijk and T. Kailath, "Coding scheme for additive noise channels with feedback-Part I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. 12, no. 2, pp. 172–182, Apr. 1966.

[2] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 667–671, Jul. 1984.

[3] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1423–1438, Jun. 2002.

[4] Y. Murin, Y. Kaspi, and R. Dabora, "On the Ozarow-Leung scheme for the Gaussian broadcast channel with feedback," *IEEE Sig. Proc. Letters*, vol. 22, no. 7, pp. 948–952, Jul. 2015.

[5] N. Elia, "When Bode meets Shannon: Control oriented feedback communication schemes," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1477–1488, Sep. 2004.

[6] E. Ardestanizadeh, P. Minero and M. Franceschetti, "LQG control approach to Gaussian broadcast channels with feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5267–5278, Aug. 2012.

[7] S. B. Amor, Y. Steinberg and M. Wigger, "Duality with linear-feedback schemes for the scalar Gaussian MAC and BC," in *Proc. Int. Zurich Seminar Commun.*, Zurich, Switzerland, Feb. 2014, pp. 25–28.

[8] M. Gastpar, A. Lapidoth, Y. Steinberg, and M. Wigger, "Coding schemes and asymptotic capacity for the Gaussian broadcast and interference channels with feedback," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 54–71, Jan. 2014.

[9] Y. Wu, P. Minero, and M. Wigger, "Insufficiency of linear-feedback schemes in Gaussian broadcast channels with common message," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4553–4566, Aug. 2014.

[10] Y. Murin, Y. Kaspi, R. Dabora and D. Gündüz, "Uncoded transmission of correlated Gaussian sources over broadcast channels with feedback," in *Proc. IEEE GlobalSIP Symp. on Network Theory*, Atlanta, GA, Dec. 2014, pp. 1063–1067.

[11] A. Jain, D. Gündüz, S. R. Kulkarni, V. Poor and S. Verdú, "Energy-distortion tradeoffs in Gaussian joint source-channel coding problems," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3153–3168, May 2012.

[12] D. P. Bertsekas, *Dynamic Programing and Optimal Control*, 3rd ed. Nashua, NH: Athena Sceintific, 2007.

[13] Y. Murin, Y. Kaspi, R. Dabora and D. Gündüz, "Joint source-channel coding for Gaussian broadcast channels with feedback," in preparation, available at http://www.ee.bgu.ac.il/∼moriny.