# On the Energy-Distortion Tradeoff of Gaussian Broadcast Channels with Feedback

**Yonathan Murin [1], Yonatan Kaspi [2], Ron Dabora [3],* and Deniz Gündüz [4]**

[1] Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA; moriny@stanford.edu
[2] Goldman Sachs, New York, NY 10282, USA; yonikaspi@gmail.com
[3] Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Be'er-Sheva 8410501, Israel
[4] Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK; d.gunduz@imperial.ac.uk
* Correspondence: ron@ee.bgu.ac.il; Tel.: +972-8-646-1511

**Abstract:** This work studies the relationship between the energy allocated for transmitting a pair of correlated Gaussian sources over a two-user Gaussian broadcast channel with noiseless channel output feedback (GBCF) and the resulting distortion at the receivers. Our goal is to characterize the minimum transmission energy required for broadcasting a pair of source samples, such that each source can be reconstructed at its respective receiver to within a target distortion, when the source-channel bandwidth ratio is not restricted. This minimum transmission energy is defined as the energy-distortion tradeoff (EDT). We derive a lower bound and three upper bounds on the optimal EDT. For the upper bounds, we analyze the EDT of three transmission schemes: two schemes are based on separate source-channel coding and apply encoding over multiple samples of source pairs, and the third scheme is a joint source-channel coding scheme that applies uncoded linear transmission on a single source-sample pair and is obtained by extending the Ozarow–Leung (OL) scheme. Numerical simulations show that the EDT of the OL-based scheme is close to that of the better of the two separation-based schemes, which makes the OL scheme attractive for energy-efficient, low-latency and low-complexity source transmission over GBCFs.

**Keywords:** Gaussian broadcast channel with feedback; correlated sources; joint source-channel coding; energy efficiency; energy-distortion tradeoff

## 1. Introduction

This work studies the energy-distortion tradeoff (EDT) for the transmission of a pair of correlated Gaussian sources over a two-user Gaussian broadcast channel (GBC) with noiseless, causal feedback, referred to as the GBCF. The EDT was originally proposed in [1] to characterize the minimum energy-per-source sample required to achieve a target distortion at the receiver, without constraining the source-channel bandwidth ratio. In many practical scenarios, e.g., satellite broadcasting [2], sensor networks measuring physical processes [3,4] and wireless body-area sensor networks [5–7], correlated observations need to be transmitted over noisy channels. Moreover, in various emerging applications, particularly in the context of the Internet of Things, the sampling rates are low; and hence, the channel bandwidth for transmission is much larger than the rate of the sources. Consequently, the main fundamental limitation for the communication system is the available energy per source sample. For example, in wireless body-area sensor networks, wireless computing devices located on, or inside, the human body measure physiological parameters, which typically exhibit correlations as they originate from the same source. These devices commonly have limited energy supply due to

their size and are also subject to transmission power constraints due to safety restrictions, while bandwidth can be relatively large as communications takes place over short distances [8–10]. In this application, transmission of correlated parameters measured by a single sensor to different devices can be modeled as a BC with correlated sources. As an example for such a setting, consider a sensor measuring heart rate, as well as cardiac output (volume of blood outputted from the heart per unit time), which are correlated parameters (see, e.g., [10] (Section 2.6)), where the heart rate measurements are communicated to a smart watch (e.g., for the purpose of fitness tracking), while the cardiac output is communicated to a smart phone (e.g., for health monitoring and reporting purposes).

It is well known that for lossy transmission of a Gaussian source over a Gaussian memoryless point-to-point channel, either with or without feedback, when the source-channel bandwidth ratio is fixed and the average power is finite, then separate source and channel coding (SSCC) achieves the minimum possible average mean square error (MSE) distortion [11] (Theorem 3). In [1] (Cor. 1), it is further shown that SSCC is optimal also in the sense of EDT: for any target MSE distortion level, the minimal transmission energy is achieved by optimal lossy compression [12] (Chapter 13) followed by the most energy efficient channel code [13]. While [1] (Cor. 1) considered unbounded number of source samples, more recent works [14] (Theorem 9) and [15] showed that similar observations hold also for the point-to-point channel with a finite number of source samples. Except for a few special scenarios, e.g., [16–18] and the references therein, the optimality of SSCC does not generalize to multiuser networks, and a joint design of the source and channel codes can improve the performance.

An example for a setting in which SSCC is sub-optimal is the transmission of a pair of correlated Gaussian sources over a GBC where the bandwidths of the source and the channel match (i.e., on average, a single source sample pair is transmitted over a single use of the channel). The complete characterization of the achievable distortion pairs for this problem was given in [19], which also showed that a joint source-channel coding (JSCC) transmission scheme is optimal while separation-based schemes cannot achieve the optimal performance. JSCC for the transmission of correlated sources over GBCs with a source-channel bandwidth mismatch was recently studied in [20], where novel hybrid digital/analog coding schemes were proposed and shown to be superior to other schemes known in the literature. It should be noted that the transmission of correlated sources over GBCs is an important communications scenario, which applies to a vast number of practical applications, including broadcasting video [21,22], images [23] and physical measurements [24].

The impact of feedback on lossy JSCC over multiuser channels was considered in relatively few works. Several achievability schemes and a set of necessary conditions for losslessly transmitting a pair of discrete and memoryless correlated sources over a multiple-access channel (MAC) with feedback were presented in [25]. Lossy transmission of correlated Gaussian sources over a two-user Gaussian MAC with feedback was studied in [26], in which sufficient conditions, as well as necessary conditions for the achievability of an MSE distortion pair were derived for the case in which the source and channel bandwidths match. The work [26] also showed that for the symmetric setting, if the channel signal-to-noise ratio (SNR) is low enough, then uncoded transmission is optimal. While [26] considered source-channel coding with a unit bandwidth ratio, [1] studied the EDT for the transmission of correlated Gaussian sources over a two-user Gaussian MAC with and without feedback, when the bandwidth ratio is not restricted. Lastly, [27] improved the lower bound derived in [1] for the two-user Gaussian MAC without feedback and extended the results to more than two users.

While EDT analysis has gained some attention in recent years, the EDT of broadcast channels was considered only for GBCs without feedback. In particular, the work [15] studied the transmission of Gaussian sources over a GBC and characterized the energy-distortion exponents, namely, the exponential rate of decay of the square-error distortion as the available energy-to-noise ratio increases without bound. For GBCFs, the existing literature mainly focused on channel coding aspects, considering independent and uniformly distributed messages. A key work in this context is the work of Ozarow and Leung (OL) [28], which obtained inner and outer bounds on the capacity region of the two-user GBCF, by extending the point-to-point transmission strategy of Schalkwijk–Kailath

(SK) [29]. The work [30] extended the OL scheme for two-user GBCFs by using estimators with memory (at the receivers) instead of the memoryless estimators used in the original OL scheme of [28]. In contrast to the point-to-point case [29], for GBCFs, both the scheme of [28] and the scheme of [30] are generally suboptimal. While the analysis and construction of the OL scheme [28] are carried out in an estimation theoretic framework, the works [31,32] approached the problem of channel coding for the GBCF within a control theoretic framework. Specifically, [32] proposed a transmission scheme based on linear quadratic Gaussian (LQG) control theory, that achieves rate pairs outside the achievable rate region of the OL code developed in [28]. Recently, it was shown in [33,34] that, for the two-user GBCF when the noise components at the receivers are mutually independent with equal variances, the LQG scheme of [32] achieves the maximal sum-rate among all possible linear-feedback schemes. Finally, it was shown in [35] that the capacity of GBCFs with independent noises at the receivers and only a common message cannot be achieved using a linear feedback scheme. Instead, the work [35] presented a capacity-achieving non-linear feedback scheme.

JSCC for the transmission of correlated Gaussian sources over GBCFs when the number of transmitted symbols is finite (referred to as the finite horizon regime) was previously considered in [36], which studied the minimal number of channel uses required to achieve a target MSE distortion pair. Three linear encoding schemes based on uncoded transmission were considered: the first scheme was a JSCC scheme based on the coding scheme of [28], to which we shall refer as the OL scheme; the second scheme was a JSCC scheme based on the scheme of [32], to which we shall refer as the LQG scheme; and the third scheme was a JSCC scheme whose parameters are obtained using dynamic programming (DP) (in the present work we discuss only the former OL and LQG schemes since the scheme based on DP becomes analytically and computationally infeasible as the number of channel uses goes to infinity). We note that linear and uncoded transmission, as implemented in the OL and in the LQG schemes, has important advantages, including low computational complexity, short coding delays and small storage requirements, which make this type of coding very desirable. We further note that although the LQG channel coding scheme of [32] for the two-user GBCF (with two messages) achieves the largest rate region out of all known channel coding schemes, in [36], it was shown that when the time horizon is finite, JSCC based on the OL scheme can achieve MSE distortion pairs lower than the JSCC based on the LQG scheme. In the present work, we analyze lossy source coding over GBCFs using SSCC and JSCC schemes based on a different performance metric: the EDT.

We note here that, as discussed above, noiseless feedback has been studied extensively in wireless Gaussian networks. An immediate benefit of this analysis is that the performance obtained for noiseless feedback serves as an upper bound on the performance for channels with noisy feedback. The analysis of noiseless feedback scenarios also leads to guidelines and motivation, which then can be applied to channels with noisy feedback. Indeed, the works [37,38], which studied channel coding for point-to-point Gaussian channels with noisy feedback and for GBCs with noisy feedback, respectively, considered transmission schemes, which are based on the SK [29] and on the OL schemes [28], respectively, originally developed for noiseless feedback scenarios. In [37,38], the noise in the feedback links was handled by applying modulo-lattice precoding in both the direct and feedback links. It is shown in [37,38] that, while having noise in the feedback links results in a performance degradation compared to the case of noiseless feedback [37] (Section V.D), many of the benefits of noiseless feedback can be carried over to the more practical setup of noisy feedback, thereby further motivating the current work. It follows that the analysis of noiseless feedback models provides practically relevant insights while facilitating simpler analysis.

**Main contributions:** In this work, the EDT for GBCFs is studied for the first time. We derive lower and upper bounds on the minimum energy per source pair required to achieve a target MSE distortion at each receiver, for the problem of transmitting a pair of Gaussian sources over a two-user GBCF, without constraining the number of channel uses per source sample. The new lower bound is based on cut-set arguments, while the upper bounds are obtained using three transmission schemes: two SSCC schemes and an uncoded JSCC scheme. The first SSCC scheme jointly compresses the two

source sequences into a single bit stream, and transmits this stream to both receivers as a common message. The second SSCC scheme separately encodes each source sequence into two distinct bit streams, and broadcasts them via the LQG channel code of [32]. It is shown that in terms of the minimum energy-per-bit, the LQG code provides no gain compared to orthogonal transmission, from which we conclude that the first SSCC scheme, that jointly compresses the sequences into a single stream, is more energy efficient. As both SSCC schemes apply coding over multiple samples of the source pairs, they require high computational complexity, long delays and large storage space. We then consider the uncoded JSCC OL scheme presented in [36]. For this scheme, we first consider the case of fixed SNR and derive an upper bound on the number of channel uses required to achieve a target distortion pair. When the SNR approaches zero, the required number of channel uses grows, and the derived bound becomes tight. At the limiting scenario of SNR → 0, this provides a simple upper bound on the EDT. While our primary focus in this work is on the analysis of the three schemes mentioned above, such an analysis is a first step towards identifying schemes that would achieve improved EDT performance in GBCFs.

Numerical results indicate that the SSCC scheme based on joint compression achieves better EDT compared to the JSCC OL scheme; yet, the gap is quite small. Moreover, in delay-sensitive applications, there is a constraint on the maximal allowed latency in transmitting each source sample to the destination. In such scenarios, coding over large blocks of independent and identically distributed (i.i.d.) pairs of source samples is not possible, and instantaneous transmission of each observed pair of source samples via the JSCC-OL scheme may be preferable in order to satisfy the latency requirement, while maintaining high energy efficiency.

The rest of this paper is organized as follows: The problem formulation is detailed in Section 2. The lower bound on the minimum energy per source sample is derived in Section 3. Upper bounds on the minimum energy per source sample are derived in Sections 4 and 5. Numerical results are detailed in Section 6, and concluding remarks are provided in Section 7.

## 2. Problem Definition

### 2.1. Notation

We use capital letters to denote random variables, e.g., $X$, and boldface letters to denote column random vectors, e.g., $\mathbf{X}$; the $k^{\text{th}}$ element of a vector $\mathbf{X}$ is denoted by $X_k, k \geq 1$, and we use $X_k^j$, with $j \geq k$, to denote $(X_k, X_{k+1}, ..., X_j)$. We use sans-serif fonts to denote matrices, e.g., Q. We use $h(\cdot)$ to denote differential entropy, $I(\cdot;\cdot)$ to denote mutual information, and $X \leftrightarrow Y \leftrightarrow Z$ to denote a Markov chain, as defined in [12] (Chapters 9 and 2). We use $\mathbb{E}\{\cdot\}$, $(\cdot)^T$, $\log(\cdot)$, $\mathcal{R}$ and $\mathcal{N}$ to denote expectation, transpose, natural base logarithm, the set of real numbers and the set of non-negative integers, respectively. We let $\mathcal{O}(g_1(P))$ denote the set of functions $g_2(P)$ such that $\limsup_{P \to 0} |g_2(P)/g_1(P)| < \infty$. Finally, we define $\text{sgn}(x)$ as the sign of $x \in \mathcal{R}$, with $\text{sgn}(0) \triangleq 1$.

### 2.2. Problem Setup

The two-user GBCF is depicted in Figure 1, with all of the signals being real. In this work, we consider the symmetric setting in which the sources have the same variances and the noises have the same variances. The encoder observes $m$ i.i.d. realizations of a correlated and jointly Gaussian pair of sources $(S_{1,j}, S_{2,j}) \sim \mathcal{N}(0, Q_s), j = 1, \ldots, m$, where $Q_s \triangleq \sigma_s^2 \cdot \begin{bmatrix} 1 & \rho_s \\ \rho_s & 1 \end{bmatrix}, |\rho_s| < 1$. The task of the encoder (transmitter) is to generate a transmitted signal that will facilitate decoding of the sequence of the $i^{\text{th}}$ source, $S_{i,1}^m, i = 1, 2$, at the $i$th decoder (receiver), denoted by $\text{Rx}_i$, whose channel output at time $k$ is given by:

$$Y_{i,k} = X_k + Z_{i,k}, \quad i = 1, 2, \tag{1}$$

for $k = 1, \ldots, n$. The noise sequences $\{Z_{1,k}, Z_{2,k}\}_{k=1}^n$ are i.i.d. over $k = 1, 2, \ldots, n$, with $(Z_{1,k}, Z_{2,k}) \sim \mathcal{N}(0, \mathsf{Q}_z)$, where $\mathsf{Q}_z \triangleq \sigma_z^2 \cdot \begin{bmatrix} 1 & \rho_z \\ \rho_z & 1 \end{bmatrix}$, $|\rho_z| < 1$.
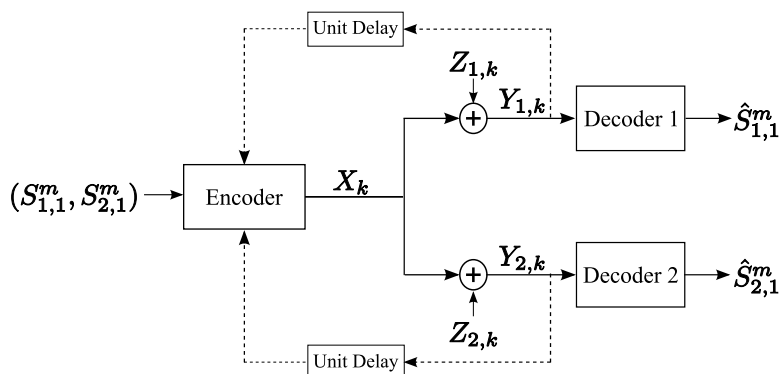


**Figure 1.** Gaussian broadcast channel with correlated sources and feedback links. $\hat{S}_{1,1}^m$ and $\hat{S}_{2,1}^m$ are the reconstructions of $S_{1,1}^m$ and $S_{2,1}^m$, respectively.

Let $\mathbf{Y}_k \triangleq (Y_{1,k}, Y_{2,k})$. The encoder maps the observed pair of source sequences and the noiseless causal channel outputs obtained through the feedback links into a channel input via: $X_k = f_k(S_{1,1}^m, S_{2,1}^m, \mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_{k-1}), f_k : \mathcal{R}^{2(m+k-1)} \mapsto \mathcal{R}$. Rx$_i$, $i = 1, 2$, uses its channel output sequence $Y_{i,1}^n$ to estimate $S_{i,1}^m$ via $\hat{S}_{i,1}^m = g_i(Y_{i,1}^n), g_i : \mathcal{R}^n \mapsto \mathcal{R}^m$.

We study the symmetric GBCF with parameters $(\sigma_s^2, \rho_s, \sigma_z^2, \rho_z)$, and define a $(D, E, m, n)$ code to be a collection of $n$ encoding functions $\{f_k\}_{k=1}^n$ and two decoding functions $g_1, g_2$, such that the MSE distortion satisfies:

$$\sum_{j=1}^m \mathbb{E}\left\{ (S_{i,j} - \hat{S}_{i,j})^2 \right\} \leq mD, \quad 0 < D \leq \sigma_s^2, \quad i = 1, 2, \tag{2}$$

and the energy of the transmitted signals satisfies:

$$\sum_{k=1}^n \mathbb{E}\left\{ X_k^2 \right\} \leq mE. \tag{3}$$

Our objective is to characterize the minimal $E$, for a given target MSE $D$ at each user, such that for all $\epsilon > 0$, there exist $m, n$ and a $(D + \epsilon, E + \epsilon, m, n)$ code. We call this minimal value the EDT and denote it by $E(D)$.

**Remark 1** (Energy constraint vs. power constraint). *The constraint (3) reflects the energy per source sample rather than per channel use. Note that by defining $P \triangleq \frac{m}{n} E$, the constraint (3) can be equivalently stated as $\frac{1}{n} \sum_{k=1}^n \mathbb{E}\left\{ X_k^2 \right\} \leq P$ which is the well known average power constraint. Yet, since there is no constraint on the ratio between $m$ and $n$, given a finite energy $E$, when the number of channel uses per source sample $\frac{n}{m}$ goes to infinity, the classical average power constraint goes to zero. We also note that $E(D)$ can be obtained by evaluating the power-distortion tradeoff, namely, the minimal power required to achieve a given distortion at each receiver (see, e.g., [39] (Section II) for the definition of achievable distortion and power for a GBC with a given set of scenario parameters), in the limit $\frac{n}{m} \to \infty$. This approach was indeed used in [15] to derive energy-distortion exponents for GBCs without feedback. However, to the best of our knowledge, there are no tight bounds on the power-distortion tradeoff for GBCFs. Moreover, for the GBCF, we show next that directly characterizing $E(D)$ leads to significantly simpler results.*

### 3. Lower Bound on $E(D)$

Our first result is a lower bound on $E(D)$. First, we define $R_{S_1}(D)$ as the rate-distortion function for the source variable $S_1$, and $R_{S_1,S_2}(D)$ as the rate distortion function for jointly compressing the pair of sources $(S_1, S_2)$. Using [40] (Section III.B), we can write these functions explicitly as:

$$R_{S_1}(D) \triangleq \frac{1}{2} \log_2 \left( \frac{\sigma_s^2}{D} \right) \tag{4a}$$

$$R_{S_1,S_2}(D) \triangleq \begin{cases} \frac{1}{2} \log_2 \left( \frac{\sigma_s^2(1+|\rho_s|)}{2D - \sigma_s^2(1-|\rho_s|)} \right), & D > \sigma_s^2(1 - |\rho_s|) \\ \frac{1}{2} \log_2 \left( \frac{\sigma_s^4(1-\rho_s^2)}{D^2} \right), & D \le \sigma_s^2(1 - |\rho_s|) \end{cases}. \tag{4b}$$

Note that [40] (Section III.B) uses the function $R_{S_1,S_2}(D_1, D_2)$ as it allows for a different distortion constraint for each source. For the present setup, in which the same distortion constraint is applied to both sources, $R_{S_1,S_2}(D)$ can be obtained by setting $D_1 = D_2 = D$ in [40] (Equation (10)), and thus, we use the simplified notation $R_{S_1,S_2}(D)$. Next, define:

$$E_{\text{lb}}(D) = \sigma_z^2 \cdot \log_e 2 \cdot \max \left\{ 2R_{S_1}(D), (1 + \rho_z)R_{S_1,S_2}(D) \right\}. \tag{5}$$

The lower bound on the EDT is stated in the following theorem:

**Theorem 1.** *The EDT $E(D)$ satisfies $E(D) \ge E_{\text{lb}}(D)$.*

**Remark 2** (Different approaches for deriving a lower bound). *The work [27] presented a novel technique for lower bounding the EDT in a Gaussian MAC. Applying this technique to the symmetric GBCF results in the lower bound reported in Theorem 1. The work [39] presented a lower bound on the distortion achievable in sending correlated Gaussian sources over a GBC (without feedback). This bound uses the entropy power inequality while relying on the fact that GBCs are degraded. As GBCFs are not degraded, it is not clear if the technique used in [39] can be used for deriving lower bounds on the EDT for GBCFs.*

**Proof of Theorem 1.** As we consider a symmetric setting, in the following, we focus on the distortion at Rx$_1$, and derive two different lower bounds. The first lower bound is obtained by identifying the minimal energy required in order to achieve an MSE distortion of $D$ at Rx$_1$, while ignoring Rx$_2$. The second lower bound is obtained by considering the transmission of both sources over a point-to-point channel with two outputs $Y_1$ and $Y_2$. We begin with the following lemma:

**Lemma 1.** *If for any $\epsilon > 0$, a $(D + \epsilon, E + \epsilon, m, n)$ code exists, then the rate-distortion functions in (4) are upper bounded by:*

$$R_{S_1}(D) \le \frac{1}{m} \sum_{k=1}^{n} I(X_k; Y_{1,k}) \tag{6a}$$

$$R_{S_1,S_2}(D) \le \frac{1}{m} \sum_{k=1}^{n} I(X_k; Y_{1,k}, Y_{2,k}). \tag{6b}$$

**Proof.** The proof is provided in Appendix A. □

Now, for achievable $(D, E, m, n)$ fix $\epsilon > 0$ and consider a $(D + \epsilon, E + \epsilon, m, n)$ code. For the right-hand side of (6a), we write:

$$\frac{1}{m}\sum_{k=1}^{n} I(X_k; Y_{1,k}) \overset{(a)}{\leq} \frac{1}{m}\sum_{k=1}^{n}\frac{1}{2}\log_2\left(1 + \frac{\text{var}\{X_k\}}{\sigma_z^2}\right)$$

$$\overset{(b)}{\leq} \frac{1}{m}\sum_{k=1}^{n}\frac{1}{2}\frac{\text{var}\{X_k\}}{\sigma_z^2 \cdot \log_e 2}$$

$$\overset{(c)}{\leq} \frac{E+\epsilon}{2\sigma_z^2 \cdot \log_e 2}, \tag{7}$$

where (a) follows by considering the point-to-point channel from $X_k$ to $Y_{1,k}$ and noting that the capacity of this additive white Gaussian noise channel, subject to an input variance variance constraint $P_k$, is $\frac{1}{2}\log\left(1 + \frac{P_k}{\sigma_z^2}\right)$. Thus, given $X_k$ with variance $\text{var}(X_k)$, then setting $P_k = \text{var}(X_k)$, it follows that $I(X_k; Y_{1,k}) \leq \frac{1}{2}\log_2\left(1 + \frac{\text{var}\{X_k\}}{\sigma_z^2}\right)$; (b) follows from changing the logarithm base and from the inequality $\log_e(1+x) \leq x, \forall x \geq 0$; and (c) follows by noting that (3) implies $\sum_{k=1}^{n}\text{var}\{X_k\} \leq m(E+\epsilon)$. Combining with (6a), we obtain $R_{S_1}(D+\epsilon) \leq \frac{E+\epsilon}{2\sigma_z^2 \cdot \log_e 2}$, which implies that $2\sigma_z^2 \cdot \log_e 2 \cdot R_{S_1}(D) \leq E+\epsilon$. Since this holds for every $\epsilon > 0$, we arrive at the first term on the right-hand-side (RHS) of (5).

Next, the RHS of (6b) can be upper bounded by considering a Gaussian single-input-multiple-output channel with two receive antennas. Then, the mutual information $I(X_k; Y_{1,k}, Y_{2,k})$ is upper bounded by the capacity of the channel subject to the variance of $X_k$:

$$\frac{1}{m}\sum_{k=1}^{n} I(X_k; Y_{1,k}, Y_{2,k}) \leq \frac{1}{m}\sum_{k=1}^{n}\frac{1}{2}\log_2\left(\frac{|Q_{Y_k}|}{|Q_{Z_k}|}\right), \tag{8}$$

where (8) follows from [12] (Theorem 9.6.5), combined with [12] (Theorem 9.4.1) for jointly Gaussian random variables, and by defining $Z_k = (Z_{1,k}, Z_{2,k})$ and the covariance matrices $Q_{Y_k} \triangleq \mathbb{E}\left\{Y_k Y_k^T\right\}$ and $Q_{Z_k} \triangleq \mathbb{E}\left\{Z_k Z_k^T\right\}$. To explicitly write $Q_{Y_k}$, we note that $\mathbb{E}\{Y_{i,k}^2\} = \mathbb{E}\left\{(X_k + Z_{i,k})^2\right\} = \mathbb{E}\left\{X_k^2\right\} + \sigma_z^2$ for $i = 1, 2$, and similarly, $\mathbb{E}\left\{Y_{1,k}Y_{2,k}\right\} = \mathbb{E}\left\{X_k^2\right\} + \rho_z\sigma_z^2$. We also have $\mathbb{E}\{Z_{i,k}^2\} = \sigma_z^2$ and $\mathbb{E}\left\{Z_{1,k}Z_{2,k}\right\} = \rho_z\sigma_z^2$. Thus, we obtain $|Q_{Y_k}| = 2\mathbb{E}\{X_k^2\}\sigma_z^2(1-\rho_z) + \sigma_z^4(1-\rho_z^2)$ and $|Q_{Z_k}| = \sigma_z^4(1-\rho_z^2)$. Plugging these expressions into (8) results in:

$$\frac{1}{m}\sum_{k=1}^{n}\frac{1}{2}\log_2\left(\frac{|Q_{Y_k}|}{|Q_{Z_k}|}\right) \leq \frac{1}{m}\sum_{k=1}^{n}\frac{\mathbb{E}\left\{X_k^2\right\}}{\sigma_z^2(1+\rho_z)\log_e 2}$$

$$\leq \frac{E+\epsilon}{\sigma_z^2(1+\rho_z)\log_e 2}, \tag{9}$$

where the inequalities follow the same arguments as those leading to (7). Combining with (6b), we obtain $R_{S_1,S_2}(D) \leq \frac{E+\epsilon}{\sigma_z^2(1+\rho_z)\log_e 2}$, which implies that $2\sigma_z^2(1+\rho_z)\log_e 2 \cdot R_{S_1,S_2}(D) \leq E+\epsilon$. Since this holds for every $\epsilon > 0$, we obtain the second term on the RHS of (5). This concludes the proof. $\square$

In the next sections, we study three achievability schemes which lead to upper bounds on $E(D)$. While these schemes have simple constructions, analyzing their achievable EDT is novel and challenging.

## 4. Upper Bounds on *E(D)* via SSCC

SSCC in multiuser scenarios carries the advantages of modularity and ease of integration with the layered architecture, which is the fundamental design architecture in many practical communications systems. In this section, we analyze the EDT of two SSCC schemes. The first scheme takes advantage of the correlation between the sources and ignores the correlation between the noise components, while the second scheme ignores the correlation between the sources and aims at utilizing the correlation between the noise components.

### 4.1. The SSCC-$\rho_s$ Scheme

This scheme utilizes the correlation between the sources by first jointly encoding both source sequences into a single bit stream via the source coding scheme proposed in [41] (Theorem 6); see also [40] (Theorem III.1). For a given distortion $D$, the minimum required compression bit rate is given by the rate-distortion function stated in (4b). The bit stream generated through compression is then encoded via a channel code designed for sending a common message over the GBC (without feedback), and the corresponding codeword is transmitted to both receivers. Note that the optimal code for transmitting a common message over GBCFs with $\rho_z \neq 0$ is not known, but, when $\rho_z = 0$, the optimal code for sending a common message over the GBCF is known to be the optimal point-to-point channel code which ignores the feedback [35] (Equation (13)). Thus, SSCC-$\rho_s$ uses the correlation between the sources, but ignores the correlation between the noises at the receivers. The following theorem characterizes the minimum energy per source sample achieved by this scheme.

**Theorem 2.** *The SSCC-$\rho_s$ scheme achieves the following EDT:*

$$E_{\text{sep}}^{(\rho_s)}(D) = \begin{cases} \sigma_z^2 \log_e \left( \frac{\sigma_s^2(1+|\rho_s|)}{2D - \sigma_s^2(1-|\rho_s|)} \right), & D > \sigma_s^2(1-|\rho_s|) \\ \sigma_z^2 \log_e \left( \frac{\sigma_s^4(1-\rho_s^2)}{D^2} \right), & D \le \sigma_s^2(1-|\rho_s|) \end{cases}. \tag{10}$$

**Proof.** The optimal rate for jointly encoding the source sequences into a single-bit stream is $R_{S_1,S_2}(D)$, given in (4b) [40] (Section III.B). Note that from this stream both source sequences can be recovered to within a distortion $D$. The encoded bit stream is then transmitted to both receivers via a capacity-achieving point-to-point channel code [12] (Theorem 10.1.1) (note that this code does not exploit the causal feedback [12] (Theorem 8.12.1)). Let $E_{b\,\min}^{\text{common}}$ denote the minimum energy-per-bit required for reliable transmission over the Gaussian point-to-point channel [13]. From [13] (p. 1025), we have $E_{b\,\min}^{\text{common}} = 2\sigma_z^2 \log_e 2$. As the considered scheme is based on source-channel separation, the achievable EDT is given by $E(D) = E_{b\,\min}^{\text{common}} \times R_{S_1,S_2}(D)$, where $R_{S_1,S_2}(D)$ is stated in (4b). This results in the EDT in (10). $\square$

**Remark 3** (EDT of GBC without feedback). *A basic question that may arise is about the EDT for transmitting a pair of correlated Gaussian sources over the GBC without feedback. The work [15] studied asymmetric GBCs, namely, when the noises have different variances, and used bounds derived in [39] to characterize the energy-distortion exponents. It is not clear whether the techniques used to derive the bounds in [39] can be used for the symmetric setting discussed in the current work. For the symmetric setting, the transmission of correlated Gaussian source over the GBC has been studied in [42]. Applying the results of [42] (Footnote 2) to the current case leads to the EDT of the SSCC-$\rho_s$ scheme, which indeed does not exploit feedback.*

### 4.2. The SSCC-$\rho_z$ Scheme

This scheme aims at utilizing the correlation between the noises at the receivers, that is available at the encoder through the feedback links, for generating the channel symbols, while avoiding using the correlation between the sources for compression. As in this section we focus on separation-based schemes, the correlation between the noises at the receivers can be utilized only via the channel code. Our results show that in terms of EDT (or minimum required energy per pair of encoded bits), even the best known channel code cannot utilize the correlation between the noises at the receivers.

In the SSCC-$\rho_z$ scheme each of the source sequences is first compressed using the optimal rate-distortion source code for scalar Gaussian sources [12] (Theorem 13.3.2). Then, the resulting compressed bit streams are sent over the GBCF using the best known channel code for transmission over the GBCF, which is the LQG channel coding scheme of [32], that generally utilizes the correlation between the noises at the receivers, as is evident from [32] (IV.B) and in particular from [32]

(Equations (23) and (24)). The following theorem characterizes the minimum energy per source sample required by this scheme.

**Theorem 3.** *The SSCC-$\rho_z$ scheme achieves the EDT:*

$$E_{\text{sep}}^{(\rho_z)}(D) = 2\sigma_z^2 \log_e \left( \frac{\sigma_s^2}{D} \right). \tag{11}$$

**Proof.** The encoder separately compresses each source sequence at rate $R_{S_1}(D)$, where $R_{S_1}(D)$ is given in (4a). Thus, from each encoded stream the corresponding source sequence can be recovered to within a distortion $D$. Next, the two compressed bit streams are broadcast to their corresponding receivers using the LQG scheme of [32]. Let $E_{b\,\text{min}}^{\text{LQG}}$ denote the minimum energy per pair of encoded bits required by the LQG scheme. $E_{b\,\text{min}}^{\text{LQG}}$ is given in the following lemma:

**Lemma 2.** *For the symmetric setting, the minimum energy per pair of encoded bits required by the LQG scheme is given by:*

$$E_{b\,\text{min}}^{\text{LQG}} = 2\sigma_z^2 \log_e 2. \tag{12}$$

**Proof.** The proof is provided in Appendix B.  □

Since two bit streams are transmitted, the achievable EDT is given by $E_{\text{sep}}^{(\rho_z)}(D) = E_{b\,\text{min}}^{\text{LQG}} \times 2R_{S_1}(D)$, yielding the EDT in (11).  □

**Remark 4** (SSCC-$\rho_z$ vs. time-sharing). *Note that $E_{b\,\text{min}}^{\text{LQG}}$ in (12) is independent of $\rho_z$, and therefore, even though in general the LQG scheme is capable of utilizing the correlation between the noises at the receivers, in terms of minimum energy per pair of encoded bits it cannot (recall that the LQG scheme is the best known channel coding scheme for the GBCF). Therefore, $E_{\text{sep}}^{(\rho_z)}(D)$ is also independent of $\rho_z$, and the SSCC-$\rho_z$ scheme does not take advantage of the correlation between the noises at the receivers to improve the minimum energy per source sample needed in the symmetric setting. Indeed, an EDT of $E_{\text{sep}}^{(\rho_z)}(D)$ can also be achieved by transmitting the two bit streams via time sharing over the GBCF without using the feedback. In this context, we recall that [43] (Prop. 1) also stated that in Gaussian broadcast channels without feedback, time sharing is asymptotically optimal as the power tends to zero.*

**Remark 5** (Relationship between $E_{\text{sep}}^{(\rho_s)}(D)$, $E_{\text{sep}}^{(\rho_z)}(D)$ and $E_{\text{lb}}(D)$). *We observe that $E_{\text{sep}}^{(\rho_s)}(D) \leq E_{\text{sep}}^{(\rho_z)}(D)$. For $D \leq \sigma_s^2(1 - |\rho_s|)$ this relationship directly follows from the expressions of $E_{\text{sep}}^{(\rho_s)}(D)$ and $E_{\text{sep}}^{(\rho_z)}(D)$. For $D > \sigma_s^2(1 - |\rho_s|)$ the above relationship holds if the polynomial $q(D) = D^2(1 + |\rho_s|) - 2\sigma_s^2 D + \sigma_s^4(1 - |\rho_s|)$ is positive. This is satisfied as the discriminant of $q(D)$ is negative. We thus conclude that it is preferable to use the correlation between the sources than the correlation between the noise components. We further note that as $D \to 0$, the gap between $E_{\text{sep}}^{(\rho_s)}(D)$ and $E_{\text{sep}}^{(\rho_z)}(D)$ is bounded. On the other hand, as $D \to 0$, the gap between $E_{\text{sep}}^{(\rho_s)}(D)$ and $E_{\text{lb}}(D)$ is not bounded (note that when $\rho_z = 0$, the RHS of (5) is given by $2\sigma_z^2 \cdot \log_e 2 \cdot R_{S_1}(D)$).*

**Remark 6** (Relevance to more than two users). *The lower bound presented in Theorem 1 can be extended to the case of $K > 2$ sources using the results of [41] (Theorem 1) and [44]. The upper bound of Theorem 2 can also be extended in a relatively simple manner to $K > 2$ sources, again, using [41] (Theorem 1). The upper bound in Theorem 3 can be extended to $K > 2$ sources by using the LQG scheme for $K > 2$ [32] (Theorem 1), or by using time-sharing.*

## 5. Upper Bound on *E(D)* via the OL Scheme

Next, we derive a third upper bound on $E(D)$ by applying uncoded JSCC transmission based on the OL scheme [36] (Section 3). This scheme sequentially transmits the source pairs $(S_{1,j}, S_{2,j})$,

$j = 1, 2, \ldots, m$, without source coding. Thus, the delay introduced by the OL scheme is significantly lower than the delay introduced by the schemes discussed in Section 4. We note that the OL scheme is designed for a fixed $P = E/n$, and from condition (3) we obtain that $P = E/n \geq \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left\{X_k^2\right\}$. An upper bound on $E(D)$ can now be obtained by first calculating the minimal number of channel uses required by the OL scheme to achieve the target distortion $D$, which we denote by $K_{\mathrm{OL}}(P, D)$, and then determining the required energy via $\sum_{k=1}^{K_{\mathrm{OL}}(P,D)} \mathbb{E}\left\{X_k^2\right\}$.

*5.1. JSCC Based on the OL Scheme*

In the OL scheme, each receiver recursively estimates its intended source samples. At each time index, the transmitter uses the feedback to compute the estimation errors at the receivers at the previous time index, and transmits a linear combination of these errors. The scheme is terminated after $K_{\mathrm{OL}}(P, D)$ channel uses, when the target MSE $D$ is achieved at each receiver.

*Setup and Initialization*: Let $\hat{S}_{i,k}$ be the estimate of $S_i$ at $\mathrm{Rx}_i$ after receiving the $k^{\mathrm{th}}$ channel output $Y_{i,k}$, $\epsilon_{i,k} \triangleq \hat{S}_{i,k} - S_i$ be the estimation error after $k$ transmissions, and define $\hat{\epsilon}_{i,k-1} \triangleq \hat{S}_{i,k-1} - \hat{S}_{i,k}$. It follows that $\epsilon_{i,k} = \epsilon_{i,k-1} - \hat{\epsilon}_{i,k-1}$. Next, define $\alpha_{i,k} \triangleq \mathbb{E}\{\epsilon_{i,k}^2\}$ to be the MSE at $\mathrm{Rx}_i$ after $k$ transmissions, $\rho_k \triangleq \frac{\mathbb{E}\{\epsilon_{1,k}\epsilon_{2,k}\}}{\sqrt{\alpha_{1,k}\alpha_{2,k}}}$ to be the correlation between the estimation errors after $k$ transmissions, and $\Psi_k \triangleq \sqrt{\frac{P}{2(1+|\rho_k|)}}$. For initialization, set $\hat{S}_{i,0} = 0$ and $\epsilon_{i,0} = -S_i$, $i = 1, 2$; thus, $\rho_0 = \rho_s$. Note that for this setup and initializations, we have $\alpha_{1,k} = \alpha_{2,k} \triangleq \alpha_k$.

*Encoding*: At the $k^{\mathrm{th}}$ channel use the transmitter sends $X_k = \frac{\Psi_{k-1}}{\sqrt{\alpha_k}}\left(\epsilon_{1,k-1} + \epsilon_{2,k-1} \cdot \mathrm{sgn}(\rho_{k-1})\right)$, and the corresponding channel outputs are given by (1).

*Decoding*: Each receiver computes $\hat{\epsilon}_{i,k-1}, i = 1, 2$, based only on $Y_{i,k}$ via $\hat{\epsilon}_{i,k-1} = \frac{\mathbb{E}\{\epsilon_{i,k-1}Y_{i,k}\}}{\mathbb{E}\{Y_{i,k}^2\}}Y_{i,k}$, which can be explicitly computed as in [28] (p. 669). Then, similarly to [45] (Equation (7)), the estimate of the source $S_i$ is given by $\hat{S}_{i,k} = \sum_{m=1}^{k} \hat{\epsilon}_{i,m-1}$. Let $\Theta \triangleq P + \sigma_z^2(2 - \rho_z)$ and $\nu_z \triangleq \sigma_z^4(1 - \rho_z)^2$. The instantaneous MSE $\alpha_k$ is given by the recursive expression [28] (Equation (5)):

$$\alpha_k = \alpha_{k-1}\frac{\sigma_z^2 + \Psi_{k-1}^2(1 - \rho_{k-1}^2)}{P + \sigma_z^2}, \quad i = 1, 2, \tag{13}$$

where the recursive expression for $\rho_k$ is given by [28] (Equation (7)):

$$\rho_k = \frac{(\rho_z\sigma_z^2\Theta + \nu_z)\rho_{k-1} - \Psi_{k-1}^2\Theta(1 - \rho_{k-1}^2)\mathrm{sgn}(\rho_{k-1})}{(P + \sigma_z^2)(\sigma_z^2 + \Psi_{k-1}^2(1 - \rho_{k-1}^2))}. \tag{14}$$

**Remark 7** (Initialization of the OL scheme). *Note that in the above OL scheme we do not apply the initialization procedure described in [28] (p. 669), as it optimizes the achievable rate rather than the distortion. Instead, we set $\epsilon_{i,0} = -S_i$ and $\rho_0 = \rho_s$, thus, taking advantage of the correlation between the sources. Moreover, in Appendix C, it is explicitly shown that for the OL scheme, in the low SNR regime, the impact of the correlation between the sources on the distortion at the receivers lasts over a large number of channel transmissions. It thus follows that the proposed initialization clearly exploits the correlation between the sources.. We further note that [36] (Section III.B) considered several initialization methods for the OL scheme and showed that setting $\epsilon_{i,0} = -S_i$ and $\rho_0 = \rho_s$ outperforms the other studied initialization approaches.*

Let $E_{\mathrm{OL,min}}(D)$ denote the minimal energy per source pair required to achieve MSE $D$ at each receiver using the OL scheme. Since in the OL scheme $\mathbb{E}\left\{X_k^2\right\} = P, \forall k$, we have $E_{\mathrm{OL,min}}(D) = \min_P \{P \cdot K_{\mathrm{OL}}(P, D)\}$. From (13) one observes that the MSE value at time instant $k$ depends on $\rho_{k-1}$ and the MSE at time $k - 1$. Due to the non-linear recursive expression for $\rho_k$ in (14), it is very complicated to obtain an explicit analytical characterization for $K_{\mathrm{OL}}(P, D)$. For any fixed $P$,

we can upper bound $E_{\text{OL,min}}(D)$, and therefore $E(D)$, via upper bounding $P \cdot K_{\text{OL}}(P, D)$. In [36] (Theorem 1) we showed that $K_{\text{OL}}(P, D) \leq \frac{2(P + \sigma_z^2)}{P} \log\left(\frac{\sigma_s^2}{D}\right)$, which leads to the upper bound: $E_{\text{OL,min}}(D) \leq \min_P 2(P + \sigma_z^2) \log\left(\frac{\sigma_s^2}{D}\right) \rightarrow_{P \to 0} E_{\text{sep}}^{(\rho_z)}(D)$. However, when $P \to 0$, the upper bound $K_{\text{OL}}(P, D) \leq \frac{2(P + \sigma_z^2)}{P} \log\left(\frac{\sigma_s^2}{D}\right)$ is not tight This can be seen by considering a numerical example: Let $\sigma_s^2 = 1$, $\rho_s = 0.9$, $\sigma_z^2 = 1$, $\rho_z = 0.7$, $D = 1$, and consider two possible values for $P$: $P_1 = 10^{-4}$ and $P_2 = 10^{-6}$. Via numerical simulations one can find that $K_{\text{OL}}(P_1, D) = 38,311$, while the upper bound is 46,058. For $P_2$ we have $K_{\text{OL}}(P_2, D) = 3,830,913$, while the upper bound is 4,605,176. Thus, the gap between $K_{\text{OL}}(P, D)$ and the above bound increases as $P$ decreases. For this reason, in the next subsection we derive a tighter upper bound on $K_{\text{OL}}(P, D)$ whose ratio to $K_{\text{OL}}(P, D)$ approaches 1 as $P \to 0$. This bound is then used to derive a tighter upper bound on $E_{\text{OL,min}}(D)$.

*5.2. A New Upper Bound on $K_{\text{OL}}(P, D)$*

Following ideas from [1] (Theorem 7), we assume a fixed $\sigma_z^2$ and approximate the recursive relationships for $\rho_k$ and $\alpha_k$ given in (13) and (14) for small values of $\frac{P}{\sigma_z^2}$. We note that while [1] (Theorem 7) obtained only asymptotic expressions for $\rho_k$ and $\alpha_k$ for $\frac{P}{\sigma_z^2} \to 0$, in the following we derive tight bounds for these quantities and obtain an upper bound on $K_{\text{OL}}(P, D)$ which is valid for small values of $\frac{P}{\sigma_z^2} > 0$. Then, letting $\frac{P}{\sigma_z^2} \to 0$, the derived upper bound on $K_{\text{OL}}(P, D)$ yields an upper bound on $E_{\text{OL,min}}(D)$, and therefore on $E(D)$.

First, define: $\psi_1 \triangleq 2|\rho_z| + 5(1 - \rho_z)$, $\psi_2 \triangleq \frac{\min\{2 - \rho_z, 2(1 - \rho_z)\}}{2\sigma_z^2}$ 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 the positive quantities $B_1(P)$ and $B_2(P)$:

$$B_1(P) \triangleq \frac{(8 + \psi_1) P^3 + 24\sigma_z^2 P^2 + 12\sigma_z^4 \psi_1 P + 4\sigma_z^6 \left(4\sigma_z^2 \psi_1 + 8\right)}{8\sigma_z^{10}} P^2, \tag{15a}$$

$$B_2(P) \triangleq \frac{P + 2\sigma_z^2}{2\sigma_z^6} P^2, \tag{15b}$$

and finally, we define the quantities:

$$\bar{\rho}(P) \triangleq \frac{P(3 - \rho_z)^2}{8\sigma_z^2} + B_1(P), \tag{16a}$$

$$F_1(P) \triangleq \frac{\rho_s}{P\psi_2 - B_1(P)} \cdot \psi_3 \cdot \left(\frac{(3 - \rho_z)^2 P}{8\sigma_z^2} + B_1(P)\right)^2, \tag{16b}$$

$$F_2(P) \triangleq \frac{\rho_s}{P\psi_2 - B_1(P)} \frac{B_1(P)}{\psi_2 2\sigma_z^2}, \tag{16c}$$

$$F_3(P) \triangleq \frac{\rho_s}{P\psi_2 - B_1(P)} \cdot \left(\frac{\left(\frac{(3-\rho_z)^2 P}{8\sigma_z^2} + B_1(P)\right)^2}{(1 - \rho_z)^2} + \frac{B_1(P)}{1 - \rho_z} + B_2(P)\right), \tag{16d}$$

$$F_4(P) \triangleq \frac{P}{2\sigma_z^2} \left(-1 + \bar{\rho}(P) + \frac{2\sigma_z^2}{P} B_2(P)\right), \tag{16e}$$

$$\rho_*^{\text{lb}}(P, D) \triangleq 2 - \rho_z + \frac{\sigma_s^2}{D} (\rho_z + |\rho_s| - 2) e^{F_3(P)}, \tag{16f}$$

$$D_{\text{th}}^{\text{ub}} \triangleq \frac{\sigma_s^2 (2 - \rho_z - |\rho_s|) e^{F_3(P)}}{2 - \rho_z}, \tag{16g}$$

$$D_{\text{th}}^{\text{lb}} \triangleq \frac{\sigma_s^2 (2 - \rho_z - |\rho_s|) e^{-F_3(P)}}{2 - \rho_z}. \tag{16h}$$

For small values of $\frac{P}{\sigma_z^2}$, the following theorem provides a tight upper bound on $K_{\text{OL}}(P, D)$:

**Theorem 4.** *Let $P$ satisfy the conditions $\bar{\rho}(P) + \frac{2\sigma_z^2}{P}B_2(P) < 1$ and $B_1(P) < P\psi_2$. The OL scheme achieves MSE $D$ at each receiver within $K_{\mathrm{OL}}(P,D) \leq K_{\mathrm{OL}}^{\mathrm{ub}}(P,D)$ channel uses, where, $K_{\mathrm{OL}}^{\mathrm{ub}}(P,D)$ is given by:*

$$K_{\mathrm{OL}}^{\mathrm{ub}}(P,D) = \begin{cases} \frac{2\sigma_z^2}{P(3-\rho_z)} \log\left( \frac{(2-\rho_z-\rho_*^{\mathrm{lb}}(P,D))(1+|\rho_s|)}{(2-\rho_z-|\rho_s|)(1+\rho_*^{\mathrm{lb}}(P,D))} \right) + \frac{2\sigma_z^2}{P}\left( F_1(P) + F_2(P) \right), & D > D_{\mathrm{th}}^{\mathrm{ub}}, \quad (17a) \\[3mm] \left( \log\left( \frac{D(2-\rho_z-\bar{\rho}(P))}{\sigma_s^2(2-\rho_z-|\rho_s|)} \right) - F_3(P) \right) \frac{1}{F_4(P)} \\[2mm] \quad + \frac{2\sigma_z^2}{P(3-\rho_z)} \log\left( \frac{(2-\rho_z)(1+|\rho_s|)}{2-\rho_z-|\rho_s|} \right) + \frac{2\sigma_z^2}{P}\left( F_1(P) + F_2(P) \right), & D < D_{\mathrm{th}}^{\mathrm{lb}}. \quad (17b) \end{cases}$$

**Proof outline.** Let $\rho_s \geq 0$ (otherwise replace $S_1$ with $-S_1$). From [28] (p. 669) it follows that $\rho_k$ monotonically decreases with $k$ until it crosses zero. Let $K_{\mathrm{th}} \triangleq \min\{k \in \mathcal{N} : \rho_{k+1} < 0\}$ be the largest time index $k$ for which $\rho_k \geq 0$. In the proof of Theorem 4 we show that, for sufficiently small $\frac{P}{\sigma_z^2}$, $|\rho_k| \leq \bar{\rho}(P), \forall k \geq K_{\mathrm{th}}$. Hence, $\rho_k$ decreases until time $K_{\mathrm{th}}$ and then it has a bounded magnitude (larger than zero). This implies that the behavior of $\alpha_k$ is different in the regions $k \leq K_{\mathrm{th}}$ and $k > K_{\mathrm{th}}$. Let $\tilde{D}_{\mathrm{th}}$ be the MSE after $K_{\mathrm{th}}$ channel uses. We first derive upper and lower bounds on $\tilde{D}_{\mathrm{th}}$, denoted by $D_{\mathrm{th}}^{\mathrm{ub}}$ and $D_{\mathrm{th}}^{\mathrm{lb}}$, respectively. Consequently, we arrive at the two cases in Theorem 4: (17a) corresponds to the case of $K_{\mathrm{OL}}(P,D) < K_{\mathrm{th}}$, while (17b) corresponds to the case $K_{\mathrm{OL}}(P,D) > K_{\mathrm{th}}$. The detailed proof is provided in Appendix C.　□

**Remark 8** (Bandwidth used by the OL scheme). *Note that as $P \to 0$, $K_{\mathrm{OL}}^{\mathrm{ub}}$ increases to infinity. Since, as $P \to 0$, $\frac{K_{\mathrm{OL}}}{K_{\mathrm{OL}}^{\mathrm{ub}}} \to 1$, it follows that as $P \to 0$, $K_{\mathrm{OL}} \to \infty$. Assuming the source samples are generated at a fixed rate, this implies that the bandwidth used by the OL scheme increases to infinity as $P \to 0$.*

**Remark 9** (Theorem 4 holds for non-asymptotic values of $P$). *Note that the conditions on $P$ in Theorem 4 can be written as $P < P_{\mathrm{th}}$ with $P_{\mathrm{th}}$ depending explicitly on $\sigma_z^2$ and $\rho_z$. Plugging $B_1(P)$ in (15) into the condition $B_1(P) < P\psi_2$, we obtain the condition: $(8 + \psi_1)P^4 + 24\sigma_z^2 P^3 + 12\sigma_z^4\psi_1 P^2 + 4\sigma_z^6\left(4\sigma_z^2\psi_1 + 8\right)P < 8\psi_2\sigma_z^{10}$. We note that, in this formulation the coefficients of $P^m, m = 1,2,3,4$, are all positive. Therefore, the left-hand-side (LHS) is monotonically increasing with $P$, and since $8\psi_2\sigma_z^{10}$ is constant, the condition $B_1(P) < P\psi_2$ is satisfied if $P < P_{\mathrm{th},2}$, for some threshold $P_{\mathrm{th},2}$. Following similar arguments, the same conclusion holds for $\bar{\rho}(P) + \frac{2\sigma_z^2}{P}B_2(P) < 1$ with some threshold $P_{\mathrm{th},1}$ instead of $P_{\mathrm{th},2}$. Thus, by setting $P_{\mathrm{th}} = \min\{P_{\mathrm{th},1}, P_{\mathrm{th},2}\}$ we obtain that the conditions in Theorem 4 restrict the range of power constraint values $P$ for which the theorem holds for some $P < P_{\mathrm{th}}$.*

*5.3. An Upper Bound on $E_{\mathrm{OL,min}}(D)$*

Next, we let $P \to 0$, and use $K_{\mathrm{OL}}^{\mathrm{ub}}(P,D)$ derived in Theorem 4 to obtain an upper bound on $E_{\mathrm{OL,min}}(D)$, and therefore on $E(D)$. This upper bound is stated in the following theorem.

**Theorem 5.** *Let $D_{\mathrm{th}} \triangleq \frac{\sigma_s^2(2-\rho_z-|\rho_s|)}{2-\rho_z}$. Then, $E_{\mathrm{OL,min}}(D) \leq E_{\mathrm{OL}}(D)$, where*

$$E_{\mathrm{OL}}(D) = \begin{cases} \frac{2\sigma_z^2}{3-\rho_z} \log\left( \frac{\sigma_s^2(1+|\rho_s|)}{D+(2-\rho_z)(D-\sigma_s^2)+\sigma_s^2\cdot|\rho_s|} \right), & D \geq D_{\mathrm{th}}, \\[3mm] 2\sigma_z^2\left( \log\left( \frac{(2-\rho_z-|\rho_s|)\sigma_s^2}{(2-\rho_z)D} \right) + \frac{1}{3-\rho_z} \log\left( \frac{(2-\rho_z)(1+|\rho_s|)}{2-\rho_z-|\rho_s|} \right) \right), & D < D_{\mathrm{th}}. \end{cases} \quad (18)$$

**Proof.** We evaluate $P \cdot K_{\mathrm{OL}}^{\mathrm{ub}}(P,D)$ for $P \to 0$. Note that $B_i(P) \in \mathcal{O}(P^2), i = 1,2$, which implies that $F_j(P) \in \mathcal{O}(P), j = 1,2,3,4$. To see why this holds, consider, for example, $F_1(P)$:

$$F_1(P) = \underbrace{\frac{\rho_s \cdot \psi_3}{P\psi_2 - B_1(P)}}_{F_1^{(a)}(P)} \underbrace{\left( \frac{(3-\rho_z)^2 P}{8\sigma_z^2} + B_1(P) \right)^2}_{F_1^{(b)}(P)}.$$

Since $\rho_s, \psi_2$, and $\psi_3$ are constants, and since $B_1(P) \in \mathcal{O}(P^2)$, we have $F_1^{(a)}(P) \in \mathcal{O}(1/P)$. Now, since $\frac{(3-\rho_z)^2}{8\sigma_z^2}$ is constant we have $F_1^{(b)}(P) \in \mathcal{O}(P^2)$. Taking the product of these two asymptotics we conclude that $F_1(P) \in \mathcal{O}(P)$.

Now, for $D \geq D_{\text{th}}$ we bound the minimum $E(D)$ as follows: First, for $D \geq D_{\text{th}}^{\text{ub}}$ defined in (16g), we multiply both sides of (17a) by $P$. As $F_1(P), F_2(P) \in \mathcal{O}(P)$, then, as $P \to 0$, we obtain:

$$P \cdot K_{\text{OL}}^{\text{ub}}(P, D) = \frac{2\sigma_z^2}{3 - \rho_z} \log \left( \frac{(2 - \rho_z - \rho_*^{\text{lb}}(P, D))(1 + |\rho_s|)}{(2 - \rho_z - |\rho_s|)(1 + \rho_*^{\text{lb}}(P, D))} \right) + \mathcal{O}(P)$$

$$\xrightarrow[P \to 0]{(a)} \frac{2\sigma_z^2}{3 - \rho_z} \log \left( \frac{\sigma_s^2(1 + |\rho_s|)}{D + (2 - \rho_z)(D - \sigma_s^2) + \sigma_s^2 \cdot |\rho_s|} \right),$$

where (a) follows from (16f) by noting that $F_3(P) \in \mathcal{O}(P)$, and therefore, when $P \to 0$, $F_3(P) \to 0$. This implies that as $P \to 0$ we have $\rho_*^{\text{lb}}(P, D) \to 2 - \rho_z + \frac{\sigma_s^2}{D}(\rho_z + |\rho_s| - 2)$. Finally, note that for $P \to 0$ we have $D_{\text{th}}^{\text{ub}} \to D_{\text{th}}$.

Next, for $D < D_{\text{th}}$ we bound the minimum $E(D)$ by first noting that since $\bar{\rho}(P) \in \mathcal{O}(P)$ and $\frac{2\sigma_z^2}{P} B_2(P) \in \mathcal{O}(P)$, then $F_4(P) \in \mathcal{O}(P)$. Now, for $D < D_{\text{th}}^{\text{lb}}$ defined in (16h), multiplying both sides of (17b) by $P$, we obtain:

$$P \cdot K_{\text{OL}}^{\text{ub}}(P, D) = 2\sigma_z^2 \left( \log \left( \frac{D(2 - \rho_z - \bar{\rho}(P))}{\sigma_s^2(2 - \rho_z - |\rho_s|)} \right) + \mathcal{O}(P) \right) \cdot \frac{1}{-1 + \mathcal{O}(P)}$$

$$+ \frac{2\sigma_z^2}{3 - \rho_z} \log \left( \frac{(2 - \rho_z)(1 + |\rho_s|)}{2 - \rho_z - |\rho_s|} \right) + \mathcal{O}(P)$$

$$\xrightarrow[P \to 0]{(a)} 2\sigma_z^2 \left( \log \left( \frac{(2 - \rho_z - |\rho_s|)\sigma_s^2}{(2 - \rho_z)D} \right) + \frac{1}{3 - \rho_z} \log \left( \frac{(2 - \rho_z)(1 + |\rho_s|)}{2 - \rho_z - |\rho_s|} \right) \right),$$

where (a) follows from the fact that $\bar{\rho}(P) \in \mathcal{O}(P)$, see (16a). This concludes the proof. $\square$

**Remark 10** (Performance for extreme correlation values). *Similarly to Remark 5, as $D \to 0$, the gap between $E_{\text{OL}}(D)$ and $E_{\text{lb}}(D)$ is not bounded, which is in contrast to the situation for the OL-based JSCC for the Gaussian MAC with feedback, cf. [1] (Remark 6). When $\rho_s = 0$ we obtain that $E_{\text{OL}}(D) = E_{\text{sep}}^{(\rho_s)}(D) = E_{\text{sep}}^{(\rho_z)}(D)$, for all $0 \leq D \leq \sigma_s^2$, which follows as the sources are independent. When $|\rho_s| \to 1$ and $\rho_z \to 1$ then $E_{\text{OL}}(D) \to E_{\text{lb}}(D) \approx \sigma_z^2 \log \left( \frac{\sigma_s^2}{D} \right)$, in this case we also have $E_{\text{sep}}^{(\rho_s)}(D) \to E_{\text{lb}}(D)$ and $E_{\text{sep}}^{(\rho_z)}(D) \approx 2E_{\text{OL}}(D)$.*

**Remark 11** (Comparison of the OL scheme and the separation-based schemes). *From (10) and (18), it follows that if $D < \sigma_s^2(1 - |\rho_s|)$ then $E_{\text{OL}}(D) - E_{\text{sep}}^{(\rho_s)}(D)$ is given by:*

$$E_{\text{OL}}(D) - E_{\text{sep}}^{(\rho_s)}(D)$$
$$= 2\sigma_z^2 \left( \log \left( \frac{(2 - \rho_z - |\rho_s|)\sigma_s^2}{2 - \rho_z} \right) + \frac{1}{3 - \rho_z} \log \left( \frac{(2 - \rho_z)(1 + |\rho_s|)}{2 - \rho_z - |\rho_s|} \right) - \frac{1}{2} \log \left( \sigma_s^4(1 - \rho_s^2) \right) \right). \quad (19)$$

*Note that $E_{\text{OL}}(D) - E_{\text{sep}}^{(\rho_s)}(D)$ is independent of $D$ in this range. Similarly, from (11) and (18) it follows that if $D < D_{\text{th}}$ then $E_{\text{sep}}^{(\rho_z)}(D) - E_{\text{OL}}(D)$ is independent of $D$ and is given by:*

$$E_{\text{sep}}^{(\rho_z)}(D) - E_{\text{OL}}(D) = 2\sigma_z^2 \left( \log \left( \frac{2 - \rho_z}{2 - \rho_z - |\rho_s|} \right) + \frac{1}{3 - \rho_z} \log \left( \frac{2 - \rho_z - |\rho_s|}{(2 - \rho_z)(1 + |\rho_s|)} \right) \right). \quad (20)$$

*Note that in both cases the gap decreases as $|\rho_s|$ decreases, which follows as the scenario approaches the transmission of independent sources. The gap also increases as $\rho_z$ decreases.*

**Remark 12** (Uncoded JSCC via the LQG scheme). *In this work, we do not include an analysis of the EDT of JSCC using the LQG scheme, $E_{\mathrm{LQG}}(D)$, because JSCC-LQG does not lend itself to a concise analytical treatment, and, moreover, our numerical study demonstrated that, in terms of EDT, JSCC-LQG is generally inferior to JSCC-OL. To elaborate on these aspects, we first recall that the LQG scheme of [32] was already applied to the transmission of correlated Gaussian sources over GBCFs in [36] (Section IV). It follows from the derivations in [36] that $E_{\mathrm{LQG}}(D)$ is expressed as the solution of an optimization problem which does not have an explicit analytic solution. It is also shown in [36] that, for a finite duration of transmission and low transmission power, when the covariance matrix of the sources is different from the covariance matrix of the steady-state of the LQG scheme, then the JSCC-OL scheme outperforms the JSCC-LQG scheme, which stands in contrast to the results of [33] for the channel coding problem. This surprising conclusion carries over to the EDT as well. Indeed, using the results of [36] we carried out an extensive numerical study of JSCC-LQG, the outcome of which was that the JSCC-LQG scheme of [36] (Section IV) achieves roughly the same minimum energy as the SSCC-$\rho_z$ scheme. Since in Section 6 we show that the JSCC-OL scheme outperforms the SSCC-$\rho_z$ scheme in terms of the EDT, we decided to exclude the JSCC-LQG scheme from the numerical comparisons reported in Section 6.*

## 6. Numerical Results

In the following, we numerically compare $E_{\mathrm{lb}}(D), E_{\mathrm{sep}}^{(\rho_s)}(D), E_{\mathrm{sep}}^{(\rho_z)}(D)$ and $E_{\mathrm{OL}}(D)$. We set $\sigma_s^2 = \sigma_z^2 = 1$ and consider several values of $\rho_z$ and $\rho_s$. Figure 2a depicts $E_{\mathrm{lb}}(D), E_{\mathrm{sep}}^{(\rho_s)}(D), E_{\mathrm{sep}}^{(\rho_z)}(D)$ and $E_{\mathrm{OL}}(D)$ for $\rho_z = 0.5$, and for two values of $\rho_s$: $\rho_s = 0.2$ and $\rho_s = 0.9$. As $E_{\mathrm{sep}}^{(\rho_z)}(D)$ is not a function of $\rho_s$, it is plotted only once. It can be observed that when $\rho_s = 0.2$, then $E_{\mathrm{sep}}^{(\rho_s)}(D), E_{\mathrm{sep}}^{(\rho_z)}(D)$ and $E_{\mathrm{OL}}(D)$ are almost the same. This follows because when the correlation between the sources is low, the gain from utilizing this correlation is also low. Furthermore, when $\rho_s = 0.2$ the gap between the lower bound and the upper bounds is evident. On the other hand, when $\rho_s = 0.9$, both SSCC-$\rho_s$ and OL significantly improve upon SSCC-$\rho_z$. This follows as SSCC-$\rho_z$ does not take advantage of the correlation among the sources. It can further be observed that when the distortion is low, there is a small gap between OL and SSCC-$\rho_s$, while when the distortion is high, OL and SSCC-$\rho_s$ require roughly the same amount of energy per source-pair sample. This is also supported by Figure 2c. We conclude that as the SSCC-$\rho_s$ scheme encodes over long sequences of source samples, it better exploits the correlation among the sources compared to the OL scheme.

Figure 2b depicts $E_{\mathrm{lb}}(D), E_{\mathrm{sep}}^{(\rho_s)}(D), E_{\mathrm{sep}}^{(\rho_z)}(D)$ and $E_{\mathrm{OL}}(D)$ vs. $D$, for $\rho_s = 0.8$, and for $\rho_z \in \{-0.9, 0.9\}$. As $E_{\mathrm{sep}}^{(\rho_s)}(D)$ and $E_{\mathrm{sep}}^{(\rho_z)}(D)$ are not functions of $\rho_z$, we plot them only once. It can be observed that when $\rho_z = 0.9$, $E_{\mathrm{lb}}(D), E_{\mathrm{sep}}^{(\rho_s)}(D)$ and $E_{\mathrm{OL}}(D)$ are very close to each other, as was analytically concluded in Remark 10. On the other hand, for $\rho_z = -0.9$ the gap between the bounds is large.

Note that while analytically comparing $E_{\mathrm{sep}}^{(\rho_s)}(D), E_{\mathrm{sep}}^{(\rho_z)}(D)$ and $E_{\mathrm{OL}}(D)$ for any $D$ is difficult, our numerical simulations suggest the relationship $E_{\mathrm{sep}}^{(\rho_s)}(D) \leq E_{\mathrm{OL}}(D) \leq E_{\mathrm{sep}}^{(\rho_z)}(D)$, for all values of $D, \rho_s, \rho_z$. For example, Figure 2c depicts the difference $E_{\mathrm{OL}}(D) - E_{\mathrm{sep}}^{(\rho_s)}(D)$ for $\rho_z = 0.5$, and for all values of $D$ and $|\rho_s|$. It can be observed that for low values of $|\rho_s|$, or for high values of $D$, $E_{\mathrm{sep}}^{(\rho_s)}(D) \approx E_{\mathrm{OL}}(D)$. On the other hand, when the correlation among the sources is high and the distortion is low, then the SSCC-$\rho_s$ scheme improves upon the OL scheme. When $D < \sigma_s^2(1 - |\rho_s|)$ we can use (19) to analytically compute the gap between the energy requirements of the two schemes. For instance, at $\rho_s = 0.99$ and $D < 0.02$ the gap is approximately 3.173. Figure 2d depicts the difference $E_{\mathrm{sep}}^{(\rho_z)}(D) - E_{\mathrm{OL}}(D)$ for $\rho_z = 0.5$. It can be observed that larger $|\rho_s|$ results in a larger gap. Again we can use (20) to analytically compute the gap between the energy requirements of the two schemes for a certain range of distortion values: At $\rho_s = 0.99$ and $D < 0.34$, the gap is approximately 0.744. Finally, as stated in Remark 12, the LQG scheme achieves approximately the same minimum energy as the SSCC-$\rho_z$ scheme, hence, OL is expected to outperform LQG. This is in accordance with [36] (Section VI), which shows that for low values of $P$, OL outperforms LQG, but, is in contrast to the channel coding problem in which the LQG scheme of [32] is known to achieve higher rates compared to the OL scheme of [28].
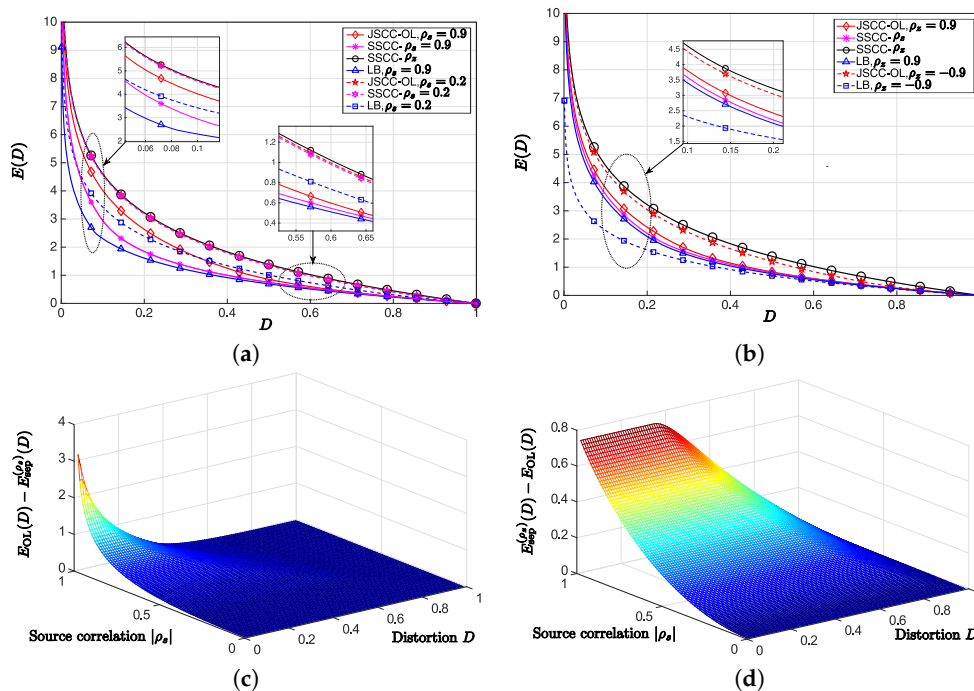
**Figure 2.** Numerical results. (**a**) Upper and lower bounds on $E(D)$ for $\sigma_s^2 = \sigma_z^2 = 1$, and $\rho_z = 0.5$. Solid lines correspond to $\rho_s = 0.9$, while dashed lines correspond to $\rho_s = 0.2$. (**b**) Upper and lower bounds on $E(D)$ for $\sigma_s^2 = \sigma_z^2 = 1$, $\rho_s = 0.8$. Solid lines correspond to $\rho_z = 0.9$, while dashed lines correspond to $\rho_z = -0.9$. (**c**) Normalized excess energy requirement of the OL scheme over the SSCC-$\rho_s$ scheme, $\rho_z = 0.5$. (**d**) Normalized excess energy requirement of the SSCC-$\rho_z$ scheme over the OL scheme, $\rho_z = 0.5$.

## 7. Conclusions and Future Work

In this work, we studied the EDT for sending correlated Gaussian sources over GBCFs, without constraining the source-channel bandwidth ratio. In particular, we first derived a lower bound on the minimum energy per source pair sample using information theoretic tools and then presented upper bounds on the minimum energy per source pair sample by analyzing three transmission schemes. The first scheme, SSCC-$\rho_s$, jointly encodes the source sequences into a single bit stream, while the second scheme, SSCC-$\rho_z$, separately encodes each of the sequences, thus, it does not exploit the correlation among the sources. We further showed that the LQG channel coding scheme of [32] achieves the same minimum energy-per-bit as orthogonal transmission, and therefore, in terms of the minimum energy-per-bit, it does not take advantage of the correlation between the noises at the receivers. We also concluded that SSCC-$\rho_s$ outperforms SSCC-$\rho_z$.

The third scheme analyzed is the OL scheme for which we first derived an upper bound on the number of channel uses required to achieve a target distortion pair, which, in the limit $P \to 0$, leads to an upper bound on the minimum energy per source pair sample. Numerical results indicate that SSCC-$\rho_s$ outperforms the OL scheme, as well. On the other hand, the gap between the energy requirements of the two schemes is rather small. We note that the SSCC-$\rho_s$ scheme implements coding over blocks of samples of source pairs, which introduces high computational complexity, large delays and requires a large amount of storage space. On the other hand, the OL scheme applies linear and uncoded transmission to each source pair sample separately, which requires low computational complexity, short delays and limited storage space. Our results demonstrate that the OL scheme provides an attractive alternative for energy efficient transmission over GBCFs.

Finally, we note that for the Gaussian MAC with feedback, OL-based JSCC is very close to the lower bound, cf. [1] (Figure 4), while, as indicated in Section 6, for the GBCF, the gap between

the OL-JSCC and the lower bound is larger. This difference is also apparent in the channel coding problem for GBCFs, namely between the achievable rate region of the OL scheme and the tightest outer bound (note that while the OL strategy achieves the capacity of the Gaussian MAC with feedback [32] (Section V.A), for the GBCF the OL strategy is sub-optimal [28]). Therefore, it is interesting to see if the duality results between the Gaussian MAC with feedback and the GBCF, presented in [33,34] for the channel coding problem, can be extended to JSCC and if the approach of [33,34] facilitates a tractable EDT analysis. We consider this as a direction for future work.

**Author Contributions:** Yonathan Murin developed this work in discussion with Yonatan Kaspi, Ron Dabora, and Deniz Gündüz. Yonathan Murin wrote the paper with comments from Yonatan Kaspi, Ron Dabora, and Deniz Gündüz. All authors have read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| EDT | Energy distortion tradeoff |
| GBC | Gaussian broadcast channel |
| GBC | Gaussian broadcast channel with feedback |
| JSCC | Joint source-channel coding |
| LHS | Left-hand side |
| LQG | Linear quadratic Gaussian |
| MAC | Multiple access channel |
| MSE | Mean square error |
| OL | Ozarow–Leung |
| RHS | Right-hand side |
| SK | Schalkwijk–Kailath |
| SNR | Signal-to-noise ratio |
| SSCC | Separate source-channel coding |

## Appendix A. Proof of Lemma 1

We begin with the proof of (6a). From [12] (Theorem 13.2.1) we have:

$$R_{S_1}(D) = \inf_{P_{\hat{S}_1|S_1} : \mathbb{E}\{(\hat{S}_1 - S_1)^2\} \leq D} I(\hat{S}_1; S_1). \tag{A1}$$

Now, for any $\epsilon > 0$ we write:

$$m \cdot R_{S_1}(D + \epsilon) \overset{(a)}{\leq} \inf_{P_{\hat{S}_{1,1}^m|S_{1,1}^m} : \sum_{j=1}^m \mathbb{E}\{(\hat{S}_{1,j} - S_{1,j})^2\} \leq m(D+\epsilon)} \sum_{j=1}^m I(\hat{S}_{1,j}; S_{1,j}|S_{1,1}^{j-1}) \overset{(b)}{\leq} I(\hat{S}_{1,1}^m; S_{1,1}^m), \tag{A2}$$

where (a) follows from the convexity of the mutual information $I(\hat{S}_1; S_1)$ in the conditional distribution $P_{\hat{S}_1|S_1}$, and from the assumption that the sources are memoryless; (b) is due to the non-negativity of mutual information combined with the chain rule for mutual information. Next, we upper bound $I(\hat{S}_{1,1}^m; S_{1,1}^m)$ as follows:

$$I(\hat{S}_{1,1}^m; S_{1,1}^m) \overset{(a)}{\leq} I(Y_{1,1}^n; S_{1,1}^m) \overset{(b)}{\leq} \sum_{k=1}^n I(X_k; Y_{1,k}), \tag{A3}$$

where (a) follows from the data processing inequality [12] (Section 2.8), by noting that $S_1^m \leftrightarrow Y_1^n \leftrightarrow \hat{S}_1^m$; (b) follows from the fact that conditioning reduces entropy, and from the fact that since the channel is memoryless, then $Y_{1,k}$ depends on $(S_1^m, X_k, Y_{1,1}^{k-1})$ only through the channel input $X_k$, see (1). By combining (A1)–(A3) we obtain (6a).

Next, we prove (6b). From [40] (Theorem III.1) we have:

$$R_{S_1,S_2}(D) = \inf_{\substack{P_{\hat{S}_1,\hat{S}_2|S_1,S_2}: \\ \mathbb{E}\{(\hat{S}_i - S_i)^2\} \leq D, i=1,2}} I(\hat{S}_1, \hat{S}_2; S_1, S_2). \tag{A4}$$

Again, for any $\epsilon > 0$, we write:

$$m \cdot R_{S_1,S_2}(D + \epsilon) \overset{(a)}{\leq} \inf_{\substack{P_{\hat{S}_{1,1}^m, \hat{S}_{2,1}^m|S_{1,1}^m, S_{2,1}^m}: \\ \sum_{j=1}^m \mathbb{E}\{(\hat{s}_{j,i} - S_{j,i})^2\} \leq m(D+\epsilon), i=1,2}} \sum_{j=1}^m I(\hat{S}_{1,j}, \hat{S}_{2,j}; S_{1,j}, S_{2,j}) \overset{(b)}{\leq} I(\hat{S}_{1,1}^m, \hat{S}_{2,1}^m; S_{1,1}^m, S_{2,1}^m), \tag{A5}$$

where (a) is due to the convexity of the mutual information $I(\hat{S}_1, \hat{S}_2; S_1, S_2)$ in the conditional distribution $P_{\hat{S}_1,\hat{S}_2|S_1,S_2}$, and (b) follows from the memorylessness of the sources, the chain rule for mutual information, and from the fact that it is non-negative. Next, we upper bound $I(\hat{S}_{1,1}^m, \hat{S}_{2,1}^m; S_{1,1}^m, S_{2,1}^m)$ as follows:

$$I(\hat{S}_1^m, \hat{S}_2^m; S_1^m, S_2^m) \overset{(a)}{\leq} I(Y_1^n, Y_2^n; S_1^m, S_2^m) \overset{(b)}{\leq} \sum_{k=1}^n I(X_k; Y_{1,k}, Y_{2,k}), \tag{A6}$$

where (a) follows from the data processing inequality [12] (Section 2.8), by noting that we have $(S_1^m, S_2^m) \leftrightarrow (Y_1^n, Y_2^n) \leftrightarrow (\hat{S}_1^m, \hat{S}_2^m)$; (b) follows from the fact that conditioning reduces entropy, and from the fact that the channel is memoryless, thus, $Y_{1,k}$ and $Y_{2,k}$ depend on $(S_1^m, S_2^m, X_k, Y_{1,1}^{k-1}, Y_{2,1}^{k-1})$ only through the channel input $X_k$, see (1). By combining (A4)–(A6) we obtain (6b). This concludes the proof of the lemma.

### Appendix B. Proof of Lemma 2: Minimum Energy-Per-Bit for the LQG Scheme

We first note that by following the approach taken in the achievability part of [48] (Theorem 1) it can be shown that for the symmetric GBCF with symmetric rates, the minimum energy-per-bit is given by:

$$E_{b\,\min}^{\text{LQG}} = \lim_{P \to 0} \frac{P}{R_{\text{LQG}}^{\text{sum}}(P)}, \tag{A7}$$

where $R_{\text{LQG}}^{\text{sum}}(P)$ is the sum rate achievable by the LQG scheme. Let $x_0$ be the unique positive real root of the third order polynomial $p(x) = (1 + \rho_z)x^3 + (1 - \rho_z)x^2 - \left(1 + \rho_z + \frac{2P}{\sigma_z^2}\right)x - (1 - \rho_z)$. From [32] (Equation (26)), for the symmetric GBCF, the achievable per-user rate of the LQG scheme is $R_{\text{LQG}}(P) = \frac{1}{2}\log_2(x_0)$ bits. We now follow the approach taken in [36] (Appendix A.3) and bound $x_0$ using Budan's theorem [49].

Explicitly writing the derivatives of $p(x)$ and evaluating the sequence $p^{(i)}(1)$, $i = 0, 1, 2, 3$, we have $V(1) = 1$. Next, we let $\chi = \frac{2P}{\alpha\sigma_z^2}$ where $\alpha > 0$ is a real constant. Setting $x = 1 + \chi$ we obtain $p(1+\chi) = (1+\rho_z)\chi^3 + (4 + 2\rho_z - \alpha)\chi^2 + (4 - \alpha)\chi$, $p^{(1)}(1+\chi) = 3(1+\rho_z)\chi^2 + (8 + 4\rho_z - \alpha)\chi + 4$, and $p^{(2)}(1+\chi), p^{(3)}(1+\chi) > 0$. Note that we are interested in the regime $P \to 0$ which implies that $\chi \to 0$. Now, for $\chi$ small enough we have $p^{(1)}(1+\chi) \approx 4 > 0$. Furthermore, when $\chi \to 0$ we have $p^{(0)}(1+\chi) = p^{(0)}\left(1 + \frac{2P}{\alpha\sigma_z^2}\right) \approx (4 - \alpha)\frac{2P}{\alpha\sigma_z^2}$. Clearly, for any $0 < \alpha < 4$, $\lim_{P\to0} p^{(0)}(1 + \frac{2P}{\alpha\sigma_z^2}) > 0$, and when $\alpha > 4$, $\lim_{P\to0} p^{(0)}(1 + \frac{2P}{\alpha\sigma_z^2}) < 0$. Thus, letting $0 < \delta < 4$, Budan's theorem implies that

when $P \to 0$, the number of roots of $p(x)$ in the interval $\left(1 + \frac{2P}{(4+\delta)\sigma_z^2}, 1 + \frac{2P}{(4-\delta)\sigma_z^2}\right)$ is 1. From Descartes' rule [50] (Section 1.6.3), we know that there is a unique positive root, thus, as this holds for any $0 < \delta < 4$, we conclude that $\lim_{P \to 0} x_0 = 1 + \frac{P}{2\sigma_z^2}$. Plugging the value of $x_0$ into (A7), and considering the sum-rate, we obtain:

$$E_{b\,\min}^{\mathrm{LQG}} = \lim_{P \to 0} \frac{P}{\log_2\left(1 + \frac{P}{2\sigma_z^2}\right)} = 2\sigma_z^2 \log_e 2. \tag{A8}$$

This concludes the proof.

### Appendix C. Proof of Theorem 4

First, note that if $\rho_s < 0$, we can replace $S_1$ with $-S_1$, which changes only the sign of $\rho_s$ in the joint distribution of the sources. Note that changing the sign of $\rho_{k-1}$ in (14) only changes the sign of $\rho_k$ while $|\rho_k|$ remains unchanged. Hence, $\alpha_k$ in (13) is not affected by changing the sign of $\rho_s$. Therefore, in the following we assume that $0 \leq \rho_s < 1$. To simplify the notation we also omit the dependence of $K_{\mathrm{OL}}(P, D)$ on $P$ and $D$, and write $K_{\mathrm{OL}}$. For characterizing the termination time of the OL scheme we first characterize the temporal evolution of $\rho_k$. From [28] (p. 669), $\rho_k$ decreases (with $k$) until it crosses zero. Let $K_{\mathrm{th}} \triangleq \min\{k : \rho_{k+1} < 0\}$, regardless of whether the target MSE was achieved or not. We begin our analysis with the case $K_{\mathrm{OL}} \leq K_{\mathrm{th}}$.

*Appendix C.1. The Case of $K_{\mathrm{OL}} \leq K_{\mathrm{th}}$*

From (14) we write the (first order) Maclaurin series expansion [50] (Chapter 7.3.3.3) of $\rho_{k+1} - \rho_k$ in the parameter $P$:

$$\rho_{k+1} - \rho_k = \frac{-P\left((1 - \rho_k^2)\mathrm{sgn}(\rho_k) + (1 - \rho_z)(\mathrm{sgn}(\rho_k) + \rho_k)\right)}{2\sigma_z^2} + \mathrm{Res}_1(P, k), \tag{A9}$$

where $\mathrm{Res}_1(P, k)$ is the remainder of the first order Maclaurin series expansion. The following lemma upper bounds $|\mathrm{Res}_1(P, k)|$:

**Lemma 3.** *For any $k$, we have $|\mathrm{Res}_1(P, k)| \leq B_1(P)$, where $B_1(P)$ is defined in (15).*

**Proof.** Let $\varphi(P, k) \triangleq \rho_{k+1} - \rho_k$. From Taylor's Theorem [50] (Subsection 6.1.4.5) it follows that $\mathrm{Res}_1(P, k) = \frac{\partial^2 \varphi(x,k)}{2\partial x^2} \cdot P^2$, for some $0 \leq x \leq P$. In the following we upper bound $\frac{\partial^2 \varphi(x,k)}{\partial x^2}$, for $0 \leq x \leq P$: Let $b_2 \triangleq (1 - \rho_k^2)(\mathrm{sgn}(\rho_k) + \rho_k)$, $b_1 \triangleq \rho_z \sigma_z^2(1 - \rho_k^2)(\mathrm{sgn}(\rho_k) + \rho_k) + \sigma_z^2(1 - \rho_z)(2(\mathrm{sgn}(\rho_k) + \rho_k) + \rho_k(1 - \rho_k^2))$, $a_2 \triangleq (1 - \rho_k^2)$, $a_1 \triangleq \sigma_z^2\left(2(1 + |\rho_k|) + 1 - \rho_k^2\right)$, and $a_0 \triangleq 2\sigma_z^4(1 + |\rho_k|)$ (note that in order to simplify the expressions we ignore the dependence of $b_2, b_2, a_2, a_1$, and $a_0$ on $k$). Using (14), the expression $\rho_{k+1} - \rho_k$ can now be explicitly written as $\varphi(P, k) = \frac{-b_2 P^2 - b_1 P}{a_2 P^2 + a_1 P + a_0}$, from which we obtain:

$$\frac{\partial^2 \varphi(x,k)}{\partial x^2} = 2\left(\frac{(a_1 a_2 b_2 - a_2^2 b_1)x^3 + 3a_0 a_2 b_2 x^2 + 3a_0 a_2 b_1 x + a_0 a_1 b_1 - a_0^2 b_2}{(a_2 x^2 + a_1 x + a_0)^3}\right).$$

Since $a_1, a_2 > 0$, we lower bound the denominator of $\frac{\partial^2 \varphi(x,k)}{\partial x^2}$ in the range $0 \leq x \leq P$ by $(a_2 x^2 + a_1 x + a_0)^3 \geq a_0^3 = 8\sigma_z^{12}$. Next, we upper bound each of the terms in the numerator of $\frac{\partial^2 \varphi(x,k)}{\partial x^2}$. For the coefficient of $x^3$ we write $|a_1 a_2 b_2 - a_2^2 b_1| \leq 4\sigma_z^2 \cdot 2 + |\rho_z|\sigma_z^2 \cdot 2 + \sigma_z^2(1 - \rho_z) \cdot 5 = \sigma_z^2(8 + \psi_1)$, where the inequality follows from the fact that $3 + 2|\rho_k| - \rho_k^2 \leq 4$. For the coefficient of $x^2$ we write $|3a_0 a_2 b_2| \leq 24\sigma_z^4$. For the coefficient of $x$ we write $|3a_0 a_2 b_1| \leq 12\sigma_z^6(2|\rho_z| + 5(1 - \rho_z)) = 12\sigma_z^6 \psi_1$. Finally, for the constant term we write $|a_0 a_1 b_1 - a_0^2 b_2| \leq 4\sigma_z^8(4\sigma_z^2 \psi_1 + 8)$. Collecting the above bounds on the terms of the numerator, and the bound on the denominator, we obtain $|\mathrm{Res}_1(P, k)| \leq B_1(P)$, concluding the proof of the lemma. □

Note that for $k \leq K_{\text{th}}$ we have $\rho_k > 0$. Hence, (A9) together with Lemma 3 imply that, for $k \leq K_{\text{th}}$ we have:

$$\frac{|\rho_{k+1} - \rho_k|}{P} \leq \frac{(1 + \rho_k)(2 - \rho_z - \rho_k)}{2\sigma_z^2} + \frac{B_1(P)}{P}.$$

Next, note that the function $f(x) \triangleq (1 + x)(2 - \rho_z - x), 0 \leq x < 1$ satisfies:

$$\min\{2 - \rho_z, 2(1 - \rho_z)\} \leq f(x) \leq \frac{(3 - \rho_z)^2}{4}, 0 \leq x < 1. \tag{A10}$$

The lower bound on $f(x)$ follows from the fact that $f(x)$ is concave, and the upper bound is obtained via: $\max_{x \in \mathcal{R}} f(x)$. When $B_1(P) < \psi_2 P$ then we have $\frac{\min\{2 - \rho_z, 2(1 - \rho_z)\}}{2\sigma_z^2} > \frac{B_1(P)}{P}$, hence $\frac{\min\{2 - \rho_z, 2(1 - \rho_z)\}}{2\sigma_z^2} - \frac{B_1(P)}{P} > 0$. Thus, we can combine the lower and upper bounds on $\text{Res}_1(P, k)$, and the bound on $\frac{(1 + \rho_k)(2 - \rho_z - \rho_k)}{2\sigma_z^2}$ to obtain the following lower and upper bounds on $\frac{|\rho_{k+1} - \rho_k|}{P}$:

$$\frac{\min\{2 - \rho_z, 2(1 - \rho_z)\}}{2\sigma_z^2} - \frac{B_1(P)}{P} \leq \frac{|\rho_{k+1} - \rho_k|}{P} \leq \frac{(3 - \rho_z)^2}{8\sigma_z^2} + \frac{B_1(P)}{P}. \tag{A11}$$

Now, recalling that $\rho_0 = \rho_s$, the fact that the bound in (A11) does not depend on $k$ results in the following upper bound on $K_{\text{th}}$:

$$K_{\text{th}} \leq \frac{\rho_s}{\frac{P}{2\sigma_z^2} \min\{2 - \rho_z, 2(1 - \rho_z)\} - B_1(P)} = \frac{\rho_s}{P\psi_2 - B_1(P)}. \tag{A12}$$

Next, using the fact that $\rho_k \geq 0$ for $k < K_{\text{th}}$, we rewrite (A9) as follows:

$$\frac{\rho_{k+1} - \rho_k}{(1 + \rho_k)(2 - \rho_z - \rho_k)} = -\frac{P}{2\sigma_z^2} + \frac{\text{Res}_1(P, k)}{(1 + \rho_k)(2 - \rho_z - \rho_k)},$$

which implies that for $K_{\text{OL}} \leq K_{\text{th}}$ we have:

$$\sum_{k=0}^{K_{\text{OL}}-1} \frac{\rho_{k+1} - \rho_k}{(1 + \rho_k)(2 - \rho_z - \rho_k)} = -\frac{K_{\text{OL}} P}{2\sigma_z^2} + \sum_{k=0}^{K_{\text{OL}}-1} \frac{\text{Res}_1(P, k)}{(1 + \rho_k)(2 - \rho_z - \rho_k)}. \tag{A13}$$

Observe that $\frac{\text{Res}_1(P,k)}{(1 + \rho_k)(2 - \rho_z - \rho_k)} \in \mathcal{O}(P^2)$, which follows from the fact that $0 < (1 + \rho_k)(2 - \rho_z - \rho_k)$ is lower and upper bounded independent of $P$ and $\rho_k$ (see (A10)), and from the fact that $|\text{Res}_1(P, k)| \in \mathcal{O}(P^2)$. Next, we focus on the LHS of (A13) and write:

$$\sum_{k=0}^{K_{\text{OL}}-1} \frac{\rho_{k+1} - \rho_k}{(1 + \rho_k)(2 - \rho_z - \rho_k)} = \sum_{k=0}^{K_{\text{OL}}-1} \frac{1}{(1 + \rho_k)(2 - \rho_z - \rho_k)} \int_{\rho_k}^{\rho_{k+1}} d\rho. \tag{A14}$$

Since $|\rho_z| < 1$, it follows that $\frac{1}{f(x)} = \frac{1}{(1 + x)(2 - x - \rho_z)}$ is continuous, differentiable and bounded over $0 \leq x < 1$, which implies that there exists a constant $c_0$ such that:

$$\max_{x \in [\rho_{k+1}, \rho_k]} \left| \frac{1}{f(x)} - \frac{1}{f(\rho_k)} \right| \leq c_0 |\rho_{k+1} - \rho_k|. \tag{A15}$$

The constant $c_0$ is upper bounded in the following Lemma 4. Note that (A15) constitutes an upper bound on the maximal magnitude of the difference between $\frac{1}{f(\rho_{k+1})}$ and $\frac{1}{f(\rho_k)}$.

**Lemma 4.** *The constant $c_0$, in (A15), satisfies: $c_0 \leq \max\left\{ \frac{|\rho_z - 1|}{(2 - \rho_z)^2}, \frac{1 + \rho_z}{4(1 - \rho_z)^2} \right\} \triangleq \psi_3$.*

**Proof.** Since $0 \leq \rho_k, \rho_{k+1} < 1$, the mean-value theorem [50] (Section 6.1.4) implies: $c_0 \leq \max_{x \in [0,1]} \left| \left( \frac{1}{f(x)} \right)' \right|$.

Writing $\left( \frac{1}{f(x)} \right)'$ explicitly we have: $\left( \frac{1}{f(x)} \right)' = \frac{2x - 1 + \rho_z}{((1+x)(2-x-\rho_z))^2} \triangleq g_0(x)$. To maximize $g_0(x)$ over $x \in [0,1]$,

we compute $g_0'(x) = -\frac{2(3 - 3\rho_z + \rho_z^2 - 3(1-\rho_z)x + 3x^2)}{((1+x)(-2+x+\rho_z))^3}$. Setting $g_0'(x) = 0$ requires $g_1(x) = x^2 - (1-\rho_z)x + 1 -$

$\rho_z + \frac{\rho_z^2}{3} = 0$. Since for all $|\rho_z| < 1$ the roots of $g_1(x)$ are complex (the determinant of $g_1(x)$ is equal

to $-\frac{\rho_z^2}{3} + 2\rho_z - 3 < 0, |\rho_z| < 1$.), then $g_0'(x)$ is not equal to 0 in the interval $x \in [0,1]$, and hence its

maximal value is achieved at one of the boundaries of the interval [0,1]. This concludes the proof of

the lemma. $\square$

Next, we write the LHS of (A14) as follows:

$$
\sum_{k=0}^{K_{\mathrm{OL}}-1} \frac{\rho_{k+1} - \rho_k}{(1+\rho_k)(2-\rho_z-\rho_k)} \overset{(a)}{=} \sum_{k=0}^{K_{\mathrm{OL}}-1} \int_{\rho_k}^{\rho_{k+1}} \frac{d\rho}{(1+\rho_k)(2-\rho_z-\rho_k)}
$$

$$
\overset{(b)}{\leq} \sum_{k=0}^{K_{\mathrm{OL}}-1} \int_{\rho_k}^{\rho_{k+1}} \frac{d\rho}{(1+\rho)(2-\rho_z-\rho)} + \sum_{k=0}^{K_{\mathrm{OL}}-1} \int_{\rho_k}^{\rho_{k+1}} \psi_3 \cdot |\rho_{k+1} - \rho_k| d\rho
$$

$$
\leq \int_{\rho_s}^{\rho_{K_{\mathrm{OL}}}} \frac{d\rho}{(1+\rho)(2-\rho_z-\rho)} + \sum_{k=0}^{K_{\mathrm{OL}}-1} \psi_3 \cdot |\rho_{k+1} - \rho_k|^2
$$

$$
\overset{(c)}{\leq} \frac{1}{\rho_z - 3} \log \left( \frac{(2 - \rho_z - \rho_{K_{\mathrm{OL}}})(1 + \rho_s)}{(2 - \rho_z - \rho_s)(1 + \rho_{K_{\mathrm{OL}}})} \right) + F_1(P), \tag{A16}
$$

where (a) follows from (A14); (b) follows from (A15) which implies that $\forall x \in [\rho_{k+1}, \rho_k] : \frac{1}{f(\rho_k)} \leq$

$\frac{1}{f(x)} + c_0 |\rho_{k+1} - \rho_k|$, and from Lemma 4; (c) follows from explicitly calculating the integral, and by

multiplying (A12) by the RHS of (A11) to bound the summation, and then using the upper

bounds (A11) and (A12) which leads to an upper bound on the second summation by $F_1(P)$, which is

defined in (16b). By following arguments similar to those leading to (A16) the summation at the LHS

of (A14) can be lower bounded via:

$$
\sum_{k=0}^{K_{\mathrm{OL}}-1} \frac{\rho_{k+1} - \rho_k}{(1+\rho_k)(2-\rho_z-\rho_k)} \geq \frac{1}{\rho_z - 3} \log \left( \frac{(2 - \rho_z - \rho_{K_{\mathrm{OL}}})(1 + \rho_s)}{(2 - \rho_z - \rho_s)(1 + \rho_{K_{\mathrm{OL}}})} \right) - F_1(P). \tag{A17}
$$

Next, consider again the RHS of (A13). Using the bound (A10) and Lemma 3, we can write:

$$
-\frac{K_{\mathrm{OL}} P}{2\sigma_z^2} + \sum_{k=0}^{K_{\mathrm{OL}}-1} \frac{\mathrm{Res}_1(P)}{(1+\rho_k)(2-\rho_z-\rho_k)} \leq -\frac{K_{\mathrm{OL}} P}{2\sigma_z^2} + \sum_{k=0}^{K_{\mathrm{OL}}-1} \frac{B_1(P)}{\min\{2-\rho_z, 2(1-\rho_z)\}}
$$

$$
\overset{(a)}{\leq} -\frac{K_{\mathrm{OL}} P}{2\sigma_z^2} + F_2(P), \tag{A18}
$$

where (a) follows from (A12), the LHS of (A10) and Lemma 4, and from the definitions of $\psi_2$ and $F_2(P)$

in Section 5.2. Plugging the lower bound (A17) and the upper bound (A18) into (A13) we arrive at an

upper bound on $K_{\mathrm{OL}}$ when $K_{\mathrm{OL}} < K_{\mathrm{th}}$:

$$
K_{\mathrm{OL}} \leq \frac{2\sigma_z^2}{P} \frac{1}{3 - \rho_z} \log \left( \frac{(2 - \rho_z - \rho_{K_{\mathrm{OL}}})(1 + \rho_s)}{(2 - \rho_z - \rho_s)(1 + \rho_{K_{\mathrm{OL}}})} \right) + \frac{2\sigma_z^2}{P} \left( F_1(P) + F_2(P) \right). \tag{A19}
$$

We emphasize that the above expressions hold only for $K_{\text{OL}} \leq K_{\text{th}}$, and we note that these expressions depend on $\rho_{K_{\text{OL}}}$. As $\rho_{K_{\text{OL}}}$ is unknown, in the following we bound its value. For this purpose, we set $\alpha_{K_{\text{OL}}} = D$ in (13) and write:

$$
\log\left(\frac{D}{\sigma_s^2}\right) = \sum_{k=0}^{K_{\text{OL}}-1} \log\left(\frac{2\sigma_z^2(1+|\rho_k|) + P(1-\rho_k^2)}{2(P+\sigma_z^2)(1+|\rho_k|)}\right)
$$

$$
\stackrel{(a)}{=} \sum_{k=0}^{K_{\text{OL}}-1} -\frac{P}{2\sigma_z^2}(1+|\rho_k|) + \sum_{k=0}^{K_{\text{OL}}-1} \text{Res}_2(P,k), \tag{A20}
$$

where (a) follows from the first order Maclaurin series expansion of $\log\left(\frac{2\sigma_z^2(1+|\rho_k|)+P(1-\rho_k^2)}{2(P+\sigma_z^2)(1+|\rho_k|)}\right)$ in the parameter $P$, and $\text{Res}_2(P,k)$ is the remainder term. Note that this holds for any $K_{\text{OL}}$, irrespective whether it is smaller or larger than $K_{\text{th}}$. The following lemma upper bounds $|\text{Res}_2(P,k)|$:

**Lemma 5.** *For any $k$ we have $|\text{Res}_2(P,k)| \leq B_2(P)$, where $B_2(P)$ is defined in (15).*

**Proof outline.** We follow the technique used in the proof of Lemma 3. We let $\varphi(P,k) \triangleq \log\left(\frac{2\sigma_z^2(1+|\rho_k|)+P(1-\rho_k^2)}{2(P+\sigma_z^2)(1+|\rho_k|)}\right)$, and use Taylor's theorem to write $\text{Res}_2(P,k) = \frac{\partial^2\varphi(x,k)}{2\partial x^2} \cdot P^2$ for some $0 \leq x \leq P$. Then, we upper bound $\frac{\partial^2\varphi(x,k)}{\partial x^2}$ in the range $0 \leq x \leq P$. $\square$

Next, we focus on the first summation on the RHS of (A20): From (A9), and for $k \leq K_{\text{th}}$, we have $\frac{\rho_{k+1}-\rho_k}{2-\rho_z-\rho_k} = -\frac{P}{2\sigma_z^2}(1+\rho_k) + \frac{\text{Res}_1(P,k)}{2-\rho_z-\rho_k}$. Hence, we write the first summation on the RHS of (A20), for $K_{\text{OL}} \leq K_{\text{th}}$ as:

$$
\sum_{k=0}^{K_{\text{OL}}-1} -\frac{P}{2\sigma_z^2}(1+|\rho_k|) = \sum_{k=0}^{K_{\text{OL}}-1} \frac{\rho_{k+1}-\rho_k}{2-\rho_z-\rho_k} - \sum_{k=0}^{K_{\text{OL}}-1} \frac{\text{Res}_1(P,k)}{2-\rho_z-\rho_k}. \tag{A21}
$$

Similarly to (A16) we write:

$$
\sum_{k=0}^{K_{\text{OL}}-1} \frac{\rho_{k+1}-\rho_k}{2-\rho_z-\rho_k} \leq \int_{\rho_s}^{\rho_{K_{\text{OL}}}} \frac{1}{2-\rho_z-\rho} d\rho + F_{3,1}(P)
$$

$$
= \log\left(\frac{2-\rho_z-\rho_s}{2-\rho_z-\rho_{K_{\text{OL}}}}\right) + F_{3,1}(P), \tag{A22}
$$

where

$$
F_{3,1}(P) \stackrel{(a)}{=} \underbrace{\frac{\rho_s}{P\psi_2 - B_1(P)}}_{(*)} \times \underbrace{\max_{x\in[0,1]} \left|\left(\frac{1}{2-\rho_z-x}\right)'\right| \cdot \left(\frac{(3-\rho_z)^2 P}{8\sigma_z^2} + B_1(P)\right)^2}_{(**)}
$$

$$
\stackrel{(b)}{=} \frac{\rho_s}{P\psi_2 - B_1(P)} \cdot \frac{1}{(1-\rho_z)^2} \cdot \left(\frac{(3-\rho_z)^2 P}{8\sigma_z^2} + B_1(P)\right)^2.
$$

Here, in step (a) $(*)$ is obtained as $K_{\text{OL}} \leq K_{\text{th}}$, where $K_{\text{th}}$ is upper bounded as in (A12), and $(**)$ follows from bounding $|\frac{1}{2-\rho_z-\rho} - \frac{1}{2-\rho_z-\rho_k}| \leq d_0|\rho_{k+1}-\rho_k|$, where $d_0$ is found using a similar approach to the one in the proof of Lemma 4. Then, applying arguments similar to those leading to (A16), we plug the upper bound on $|\rho_{k+1}-\rho_k|$ stated in the RHS of (A11), and combine with the bound on $d_0$ to obtain $(**)$. Step (b) follows from the fact that $\frac{\partial^2}{\partial x^2}\frac{1}{2-\rho_z-x} > 0, x \in [0,1]$ which implies that $\frac{\partial}{\partial x}\frac{1}{2-\rho_z-x}$ is increasing with $x \in [0,1]$, and therefore, its maximal value is achieved at $x = 1$.

For the second term on the RHS of (A21), noting that for $|\rho_k| < 1, 0 < \frac{1}{2-\rho_z-\rho_k} < \frac{1}{1-\rho_z}$, we write:

$$\sum_{k=0}^{K_{\mathrm{OL}}-1} \frac{\mathrm{Res}_1(P,k)}{2-\rho_z-\rho_k} \leq \frac{\rho_s}{P\psi_2 - B_1(P)} \frac{B_1(P)}{1-\rho_z} \triangleq F_{3,2}(P). \tag{A23}$$

Now, we consider the second term on the RHS of (A20). From (A12) and Lemma 5 we obtain:

$$\sum_{k=0}^{K_{\mathrm{OL}}-1} \mathrm{Res}_2(P,k) \leq K_{\mathrm{OL}} \cdot B_2(P) \leq \frac{\rho_s B_2(P)}{P\psi_2 - B_1(P)} \triangleq F_{3,3}(P). \tag{A24}$$

Therefore, from (A20)–(A24) using the definition of $F_3(P)$ in (16d), we obtain:

$$\log\left(\frac{D}{\sigma_s^2}\right) \leq \log\left(\frac{2-\rho_z-\rho_s}{2-\rho_z-\rho_{K_{\mathrm{OL}}}}\right) + F_3(P). \tag{A25a}$$

By following similar arguments for lower bounding $\log\left(\frac{D}{\sigma_s^2}\right)$, we also obtain:

$$\log\left(\frac{D}{\sigma_s^2}\right) \geq \log\left(\frac{2-\rho_z-\rho_s}{2-\rho_z-\rho_{K_{\mathrm{OL}}}}\right) - F_3(P). \tag{A25b}$$

From (A25a), we can extract the following lower bound on $\rho_{K_{\mathrm{OL}}}$: $\rho_{K_{\mathrm{OL}}} \geq 2 - \rho_z + \frac{\sigma_s^2}{D}(\rho_z + \rho_s - 2)e^{F_3(P)} \triangleq \rho_*^{\mathrm{lb}}(P,D)$. Similarly, from (A25b), we can extract the following upper bound on $\rho_{K_{\mathrm{OL}}}$: $\rho_{K_{\mathrm{OL}}} \leq 2 - \rho_z + \frac{\sigma_s^2}{D}(\rho_z + \rho_s - 2)e^{-F_3(P)} \triangleq \rho_*^{\mathrm{ub}}(P,D)$. Up to this point we assumed that $K_{\mathrm{OL}} \leq K_{\mathrm{th}}$ and therefore $\rho_{K_{\mathrm{OL}}} \geq 0$. Hence, we restricted our attention only to values of $D$ for which $\rho_*^{\mathrm{lb}}(D) \geq 0$, which is satisfied for $D \geq \frac{\sigma_s^2(2-\rho_z-\rho_s)e^{F_3(P)}}{2-\rho_z} = D_{\mathrm{th}}^{\mathrm{ub}}$. We conclude that if $D \geq D_{\mathrm{th}}^{\mathrm{ub}}$, we can obtain an upper bound on $K_{\mathrm{OL}}$ plugging $\rho_*^{\mathrm{lb}}(P,D)$ into (A19):

$$K_{\mathrm{OL}} \leq \frac{2\sigma_z^2}{P} \frac{1}{3-\rho_z} \log\left(\frac{(2-\rho_z-\rho_*^{\mathrm{lb}}(P,D))(1+\rho_s)}{(2-\rho_z-\rho_s)(1+\rho_*^{\mathrm{lb}}(P,D))}\right) + \frac{2\sigma_z^2}{P}\left(F_1(P) + F_2(P)\right). \tag{A26}$$

This corresponds to the bounds (17a). In the next subsection, we consider the case of $K_{\mathrm{OL}} > K_{\mathrm{th}}$.

*Appendix C.2. The Case of $K_{\mathrm{OL}} > K_{\mathrm{th}}$*

For upper bounding $K_{\mathrm{OL}}$ when $K_{\mathrm{OL}} > K_{\mathrm{th}}$, we first derive an upper bound on $|\rho_k|$ for $k \geq K_{\mathrm{th}}$. From (A9) we have for any $k$:

$$\begin{aligned}|\rho_{k+1} - \rho_k| &\leq \left|-\frac{P}{2\sigma_z^2}\left((1-\rho_k^2)\mathrm{sgn}(\rho_k) + (1-\rho_z)(\mathrm{sgn}(\rho_k)+\rho_k)\right)\right| + |\mathrm{Res}_1(P,k)|\\ &\stackrel{(a)}{\leq} \frac{P}{2\sigma_z^2}\left|\left((1-|\rho_k|^2) + (1-\rho_z)(1+|\rho_k|)\right)\right| + B_1(P)\\ &\stackrel{(b)}{=} \frac{P}{2\sigma_z^2}(1+|\rho_k|)(2-\rho_z-|\rho_k|) + B_1(P),\end{aligned}$$

where (a) follows from Lemma 3, and (b) follows since $|\rho_k|$ is non-negative. Thus, we can use the upper bound in (A10) to further bound:

$$\frac{P}{2\sigma_z^2}(1+|\rho_k|)(2-\rho_z-|\rho_k|) + B_1(P) \leq \frac{P(3-\rho_z)^2}{8\sigma_z^2} + B_1(P) \triangleq \bar{\rho}(P). \tag{A27}$$

Note that this bound holds for every $k$, regardless of the value of $K_{\mathrm{OL}}$. Further note that the condition $\bar{\rho}(P) + \frac{2\sigma_z^2}{P}B_2(P) < 1$ implies that $\bar{\rho}(P) < 1$. The following lemma uses (A27) to bound $|\rho_k|, k \geq K_{\mathrm{th}}$.

**Lemma 6.** *For $k \geq K_{\text{th}}$ it holds that $|\rho_k| \leq \bar{\rho}(P)$.*

**Proof.** We first recall that $\rho_{K_{\text{th}}} > 0$ while $\rho_{K_{\text{th}}+1} < 0$. Therefore, the bound $|\rho_{k+1} - \rho_k| \leq \bar{\rho}(P)$ combined with $|\rho_{K_{\text{th}}+1} - \rho_{K_{\text{th}}}| = |\rho_{K_{\text{th}}}| + |\rho_{K_{\text{th}}+1}|$ implies that $|\rho_{K_{\text{th}}}| \leq \bar{\rho}(P)$ as well as $|\rho_{K_{\text{th}}+1}| \leq \bar{\rho}(P)$. From [28] (p. 669) we have that if $\rho_k > 0$ then $\rho_{k+1} < \rho_k$, and if $\rho_k < 0$ then $\rho_{k+1} > \rho_k$. Note that these statements hold for every $k$. We now prove by induction the statement: Suppose $|\rho_{K_{\text{th}}+\Delta}| < \bar{\rho}(P)$, for $\Delta > 0$, then $|\rho_{K_{\text{th}}+\Delta+1}| < \bar{\rho}(P)$. Note that the induction assumption is satisfied for $\Delta = 1$. If $\rho_{K_{\text{th}}+\Delta} < 0$, then $\rho_{K_{\text{th}}+\Delta} < \rho_{K_{\text{th}}+\Delta+1}$, which implies that $|\rho_{K_{\text{th}}+\Delta+1}| \leq \bar{\rho}(P)$ since $|\rho_{k+1} - \rho_k| \leq \bar{\rho}(P)$. If $\rho_{K_{\text{th}}+\Delta} > 0$, then $\rho_{K_{\text{th}}+\Delta} > \rho_{K_{\text{th}}+\Delta+1}$, which again, implies that $|\rho_{K_{\text{th}}+\Delta+1}| \leq \bar{\rho}(P)$ since $|\rho_{k+1} - \rho_k| \leq \bar{\rho}(P)$. Thus, by induction we conclude that $|\rho_{K_{\text{th}}}| \leq \bar{\rho}(P), \forall k \geq K_{\text{th}}$. $\square$

Next, we characterize a lower bound on the distortion achieved after $K_{\text{th}}$ time steps. Recall that for $K_{\text{OL}} \leq K_{\text{th}}$ we have $\rho_{K_{\text{OL}}} \leq \rho_*^{\text{ub}}(P, D)$, where $\rho_*^{\text{ub}}(P, D)$ is defined in Appendix C.1. By setting $\rho_*^{\text{ub}}(P, D) = 0$, we obtain $D = \frac{\sigma_s^2(2 - \rho_z - \rho_s)e^{-F_3(P)}}{2 - \rho_z} \triangleq D_{\text{th}}^{\text{lb}}$. Thus, $D_{\text{th}}^{\text{lb}}$ constitutes a lower bound on $D_{\text{th}}$.

Now, we are ready to analyze the case of $K_{\text{OL}} > K_{\text{th}}$. We first note that (A20) holds for any value of $K_{\text{OL}}$. Hence, we write:

$$\log\left(\frac{D}{\sigma_s^2}\right) = \sum_{k=0}^{K_{\text{th}}-1}\left(-\frac{P}{2\sigma_z^2}(1 + |\rho_k|) + \text{Res}_2(P, k)\right) + \sum_{k=K_{\text{th}}}^{K_{\text{OL}}-1}\left(-\frac{P}{2\sigma_z^2}(1 + |\rho_k|) + \text{Res}_2(P, k)\right). \quad \text{(A28)}$$

For the second term on the RHS of (A28), we write:

$$\sum_{k=K_{\text{th}}}^{K_{\text{OL}}-1}\left(-\frac{P}{2\sigma_z^2}(1 + |\rho_k|) + \text{Res}_2(P, k)\right) \overset{(a)}{\leq} (K_{\text{OL}} - K_{\text{th}})\frac{P}{2\sigma_z^2}\left(-1 + \left(\bar{\rho}(P) + \frac{2\sigma_z^2}{P}B_2(P)\right)\right)$$

$$= (K_{\text{OL}} - K_{\text{th}})F_4(P).$$

where (a) follows from Lemma 5, as the lemma holds for any $k$, and from the fact that $|\rho_k| \leq \bar{\rho}(P), \forall k \geq K_{\text{th}}$. Since the sum in (A20) is negative, we require $F_4(P) < 0$, which results in $\bar{\rho}(P) + \frac{2\sigma_z^2}{P}B_2(P) < 1$. Now, we write (A28) as:

$$\sum_{k=K_{\text{th}}}^{K_{\text{OL}}-1}\left(-\frac{P}{2\sigma_z^2}(1 + |\rho_k|) + \text{Res}_2(P, k)\right) = \log\left(\frac{D}{\sigma_s^2}\right) - \sum_{k=0}^{K_{\text{th}}-1}\left(-\frac{P}{2\sigma_z^2}(1 + |\rho_k|) + \text{Res}_2(P, k)\right), \quad \text{(A29)}$$

and note that since (A20)–(A25) hold for $K_{\text{OL}} \leq K_{\text{th}}$, then replacing $K_{\text{OL}}$ with $K_{\text{th}}$ in (A20)–(A25) and $\rho_{K_{\text{OL}}}$ with $\rho_{K_{\text{th}}}$ we can bound:

$$\sum_{k=0}^{K_{\text{th}}-1}\left(-\frac{P}{2\sigma_z^2}(1 + |\rho_k|) + \text{Res}_2(P, k)\right) \leq \log\left(\frac{2 - \rho_z - \rho_s}{2 - \rho_z - \bar{\rho}(P)}\right) + F_3(P),$$

where we used the fact that $0 < \rho_{K_{\text{th}}} \leq \bar{\rho}(P)$. Thus, to obtain an upper bound on $K_{\text{OL}}$ we write:

$$(K_{\text{OL}} - K_{\text{th}})F_4(P) \geq \log\left(\frac{D}{\sigma_s^2}\right) - \log\left(\frac{2 - \rho_z - \rho_s}{2 - \rho_z - \bar{\rho}(P)}\right) - F_3(P). \quad \text{(A30)}$$

Finally, plugging $\rho_{K_{\text{th}}}$ instead of $\rho_{K_{\text{OL}}}$ in (A19), we obtain an upper bound on $K_{\text{th}}$. Since the function $\frac{(2 - \rho_z - x)(1 + \rho_s)}{(2 - \rho_z - \rho_s)(1 + x)}$ in (A19) monotonically decreases with $x$, using the lower bound $K_{\text{th}} \geq 0$, we obtain an explicit upper bound on $K_{\text{th}}$. Combining this upper bound on $K_{\text{th}}$ with (A30) we obtain the following upper bound on $K_{\text{OL}}$:

$$K_{\mathrm{OL}} \leq \left( \log \left( \frac{D(2 - \rho_z - \bar{\rho}(P))}{\sigma_s^2(2 - \rho_z - \rho_s)} \right) - F_3(P) \right) \frac{1}{F_4(P)}$$
$$+ \frac{2\sigma_z^2}{P} \frac{1}{3 - \rho_z} \log \left( \frac{(2 - \rho_z)(1 + \rho_s)}{2 - \rho_z - \rho_s} \right) + \frac{2\sigma_z^2}{P} \left( F_1(P) + F_2(P) \right),$$

where since $F_4(P) < 0$, dividing by $F_4(P)$ changes the direction of the inequality. This concludes the proof.

## References

1. Jain, A.; Gündüz, D.; Kulkarni, S.R.; Poor, H.V.; Verdú, S. Energy-distortion tradeoffs in Gaussian joint source-channel coding problems. *IEEE Trans. Inf. Theory* **2012**, *58*, 3153–3168.
2. Alagöz, F.; Gür, G. Energy efficiency and satellite networking: A holistic overview. *Proc. IEEE* **2011**, *99*, 1954–1979.
3. Rault, T.; Bouabdallah, A.; Challal, Y. Energy efficiency in wireless sensor networks: A top-down survey. *Comput. Netw.* **2014**, *67*, 104–122.
4. He, T.; Krishnamurthy, S.; Stankovic, J.A.; Abdelzaher, T.; Luo, L.; Stoleru, R.; Yan, T.; Gu, L. Energy-efficient surveillance system using wireless sensor networks. In Proceedings of the International Conference on Mobile Aystems, Applications, and Services (MobiSys), Boston, MA, USA, 6–9 June 2004.
5. Blessy, J.; Anpalagan, A. Body area sensor networks: Requirements, operations, and challenges. *IEEE Potentials* **2014**, *33*, 21–25.
6. Hanson, M.A.; Powell, H.C., Jr.; Barth, A.T.; Ringgenberg, K.; Calhoun, B.H.; Aylor, J.H.; Lach, J. Body area sensor networks: Challenges and opportunities. *Computer* **2009**, *42*, 58–65.
7. Ntouni, G.D.; Lioumpas, A.S.; Nikita, K.S. Reliable and energy-efficient communications for wireless biomedical implant systems. *IEEE J. Biomed. Health Inform.* **2014**, *18*, 1848–1856.
8. Guha, A.N. Joint Network-Channel-Correlation Decoding in Wireless Body Area Networks. Ph.D. Thesis, University of Nebraska, Lincoln, NE, USA, May 2014.
9. Deligiannis, N.; Zimos, E.; Ofrim, D.M.; Andreopoulos, Y.; Munteanu, A. Distributed joint source-channel coding with raptor codes for correlated data gathering in wireless sensor networks. In Proceedings of the International Conference on Body Area Networks (BodyNets), London, UK, 29 September–1 October 2014.
10. Brandão, P. *Abstracting Information on Body Area Networks*; Technical Report; University of Cambridge: Cambridge, UK, 2012.
11. Shannon, C.E. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.* **1959**, *4*, 142–163.
12. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 1st ed.; John Wiley & Sons: New York, NY, USA, 1991.
13. Verdú, S. On channel capacity per unit cost. *IEEE Trans. Inf. Theory* **1990**, *36*, 1019–1030.
14. Kostina, V.; Polyanskiy, Y.; Verdú, S. Joint source-channel coding with feedback. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015.
15. Köken, E.; Tuncel, E.; Gündüz, D. Energy-distortion exponents in lossy transmission of Gaussian sources over Gaussian channels. *IEEE Trans. Inf. Theory* **2017**, *63*, 1227–1236.
16. Murin, Y.; Dabora, R.; Gündüz, D. Source-channel coding theorems for the multiple-access relay channel. *IEEE Trans. Inf. Theory* **2013**, *59*, 5446–5465.
17. Gündüz, D.; Erkip, E.; Goldsmith, A.; Poor, H.V. Source and channel coding for correlated sources over multiuser channels. *IEEE Trans. Inf. Theory* **2009**, *55*, 3927–3944.
18. Tian, C.; Chen, J.; Diggavi, S.N.; Shamai, S. Optimality and approximate optimality of source-channel separation in networks. *IEEE Trans. Inf. Theory* **2014**, *60*, 904–918.
19. Tian, C.; Diggavi, S.N.; Shamai, S. The achievable distortion region of sending a bivariate Gaussian source on the Gaussian broadcast channel. *IEEE Trans. Inf. Theory* **2011**, *57*, 6419–6427.
20. Koken E.; Tuncel, E. Joint source-channel coding for broadcasting correlated sources. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016.
21. Ji, W.; Li, Z.; Chen, Y. Joint source-channel coding and optimization for layered video broadcasting to heterogeneous devices. *IEEE Trans. Multimed.* **2012**, *14*, 443–455.

22. Wu, J.; Shang, Y.; Huang, J.; Zhang, X.; Cheng, B.; Chen, J. Joint source-channel coding and optimization for mobile video streaming in heterogeneous wireless networks. *J. Wirel. Commun. Netw.* **2013**, *2013*, doi:10.1186/1687-1499-2013-283.

23. Bovik, A. *Handbook of Image and Video Processing*, 2nd ed.; Elsevier Academic Press: San Diego, CA, USA, 2005.

24. Mohassel, R.R.; Fung, A.S.; Mohammadi, F.; Raahemifar, K. A survey on advanced metering infrastructure and its application in smart grids. In Proceedings of the IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE), Toronto, ON, Canada, 4–7 May 2014.

25. Ong L.; Motani, M. Coding strategies for multiple-access channels with feedback and correlated sources. *IEEE Trans. Inf. Theory* **2007**, *53*, 3476–3497.

26. Lapidoth, A.; Tinguely, S. Sending a bivariate Gaussian source over a Gaussian MAC with feedback. *IEEE Trans. Inf. Theory* **2010**, *56*, 1852–1864.

27. Jiang, N.; Yang, Y.; Høst-Madsen, A.; Xiong, Z. On the minimum energy of sending correlated sources over the Gaussian MAC. *IEEE Trans. Inf. Theory* **2014**, *60*, 6254–6275.

28. Ozarow, L.H.; Leung-Yan-Cheong, S. An achievable region and outer bound for the Gaussian broadcast channel with feedback. *IEEE Trans. Inf. Theory* **1984**, *30*, 667–671.

29. Schalkwijk, J.P.M.; Kailath, T. A coding scheme for additive white noise channels with feedback—Part I: No bandwidth constraint. *IEEE Trans. Inf. Theory* **1966**, *12*, 172–182.

30. Murin, Y.; Kaspi, Y.; Dabora, R. On the Ozarow-Leung scheme for the Gaussian broadcast channel with feedback. *IEEE Signal Proc. Lett.* **2015**, *22*, 948–952.

31. Elia, N. When Bode meets Shannon: Control oriented feedback communication schemes. *IEEE Trans. Automat. Control* **2004**, *49*, 1477–1488.

32. Ardestanizadeh, E.; Minero, P.; Franceschetti, M. LQG control approach to Gaussian broadcast channels with feedback. *IEEE Trans. Inf. Theory* **2012**, *58*, 5267–5278.

33. Amor, S.B.; Steinberg, Y.; Wigger, M. MIMO MAC-BC duality with linear-feedback coding schemes. *IEEE Trans. Inf. Theory* **2015**, *61*, 5976–5998.

34. Amor, S.B.; Wigger, M. Linear-feedback MAC-BC duality for correlated BC-noises, and iterative coding. In Proceedings of the 53rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, 29 September–2 October 2015.

35. Wu, Y.; Minero, P.; Wigger, M. Insufficiency of linear-feedback schemes in Gaussian broadcast channels with common message. *IEEE Trans. Inf. Theory* **2014**, *60*, 4553–4566.

36. Murin, Y.; Kaspi, Y.; Dabora, R.; Gündüz, D. Finite-length linear schemes for joint source-channel coding over Gaussian broadcast channels with feedback. *IEEE Trans. Inf. Theory* **2017**, *63*, 2737–2772.

37. Ben-Yishai, A.; Shayevitz, O. Interactive schemes for the AWGN channel with noisy feedback. *IEEE Trans. Inf. Theory* **2017**, *63*, 2409–2427.

38. Ben-Yishai, A.; Shayevitz, O. The AWGN BC with MAC feedback: a reduction to noiseless feedback via interaction. In Proceedings of the 2015 IEEE Information Theory Workshop, Jerusalem, Israel, 26 April–1 May 2015.

39. Behroozi, H.; Alajaji, F.; Linder, T. On the performance of hybrid digital/analog coding for broadcasting correlated Gaussian sources. *IEEE Trans. Commun.* **2011**, *59*, 3335–3343.

40. Lapidoth, A.; Tinguely, S. Sending a bivariate Gaussian source over a Gaussian MAC. *IEEE Trans. Inf. Theory* **2010**, *56*, 2714–2752.

41. Xiao, J.J.; Luo, Z.Q. Compression of correlated Gaussian sources under individual distortion criteria. In Proceedings of the 43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, 28–30 September 2015.

42. Bross, S.I.; Lapidoth, A.; Tinguely, S. Broadcasting correlated Gaussians. *IEEE Trans. Inf. Theory* **2010**, *56*, 3057–3068.

43. Lapidoth, A.; Telatar, İ.E.; Urbanke, R. On wide-band broadcast channels. *IEEE Trans. Inf. Theory* **2003**, *49*, 3250–3258.

44. Gastpar, M. Cut-set arguments for source-channel networks. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Chicago, IL, USA, 27 June–2 July 2004.

45. Gallager, R.G.; Nakiboğlu, B. Variations on a theme by Schalkwijk and Kailath. *IEEE Trans. Inf. Theory* **2010**, *56*, 6–17.

46. Murin, Y.; Kaspi, Y.; Dabora, R.; Gündüz, D. On the transmission of a bivariate Gaussian source over the Gaussian broadcast channel with feedback. In Proceedings of the 2015 IEEE Information Theory Workshop, Jerusalem, Israel, 26 April–1 May 2015.

47. Murin, Y.; Kaspi, Y.; Dabora, R.; Gündüz, D. Energy-distortion tradeoff for the Gaussian broadcast channel with feedback. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016.

48. El Gamal, A.; Mohseni, M.; Zahedi, S. Bounds on capacity and minimum energy-per-bit for AWGN relay channels. *IEEE Trans. Inf. Theory* **2006**, *52*, 1545–1561.

49. Wilf, H.S. Budan's theorem for a class of entire functions. *Proc. Am. Math. Soc.* **1962**, *13*, 122–125.

50. Bronshtein, I.N.; Semendyayev, K.A.; Musiol, G.; Muehlig, H. *Handbook of Mathematics*, 5th ed.; Springer: Berlin/Heidelberg, Germany, 2007.