Source-Interference Recovery over Broadcast Channels: Asymptotic Bounds and Analog Codes

Ahmad Abou Saleh, Fady Alajaji, *Senior Member, IEEE*, and Wai-Yip Chan

***Abstract*—We consider the problem of joint recovery of a bivariate Gaussian source and of interference over the two-user Gaussian degraded broadcast channel in the presence of common interference. The interference, that is available non-causally at the encoder, is assumed to be Gaussian and correlated to the sources. The tradeoff between the distortion of the sources and the interference estimation error is studied; information-theoretic outer and inner bounds based on ideas from rate-distortion theory and hybrid coding are derived, respectively. More precisely, the outer bound is found by assuming additional knowledge at each user; the inner bound, however, is obtained by analyzing the distortion of a layered hybrid scheme based on proper power splitting, Costa and Wyner-Ziv coding. Low delay and complexity coding schemes based on analog mapping are next proposed. More specifically, parametric mappings based on linear and sawtooth curves are studied and optimized by minimizing an upper bound on the system's distortion; nonparametric mappings based on joint optimization between the encoder and the decoder using an iterative algorithm are designed. Numerical results show that for the special cases that are previously considered in [1] and [2] (with no fading), the derived outer bound is tighter and the proposed hybrid scheme has a lower complex structure with no loss in performance. In addition, the proposed low delay nonlinear schemes outperform the linear scheme and perform relatively close to the inner bound under certain system settings.**

***Index Terms*—Joint source-channel coding, distortion region, correlated interference, dirty paper coding, Shannon-Kotel'nikov mapping, low delay coding, degraded broadcast channels.**

I. Introduction

ONE main problem that arises in ad-hoc wireless networks is that different source-destination pairs interfere with each other. One such example, which consists of a network with five nodes, is illustrated in Fig. 1. In this network, node 1 collects three correlated measurements and aims to transmit them over two time instants. At the first time instant, node 1 transmits to node 5 with low transmission power (node 5 being close to the source node) one of the collected measurements. This transmission is also overheard by node 4 but not by nodes 2 and 3; note that nodes 2 and 3, being destination nodes, are also interested in the transmitted signal. At the following time instant, node 1 boosts its power and communicates the remaining two measurements to nodes 2 and 3. At the same time, node 4 acts as a relay and amplifies the overheard information (from the previous time instant) to nodes 2 and 3; this transmission interferes with the transmission from node 1 (the main source). Considering this scenario and assuming that node 1 uses an uncoded scheme (which is optimal for a Gaussian channel) at the first time instant, the main question is how to choose the coding structure (for the second time instant) at node 1 given that the transmission from node 4 is known to the source node 1. In this work, we develop a Gaussian model for this scenario and quantify its information-theoretic bounds. Practical low delay coding schemes based on analog techniques are also studied.

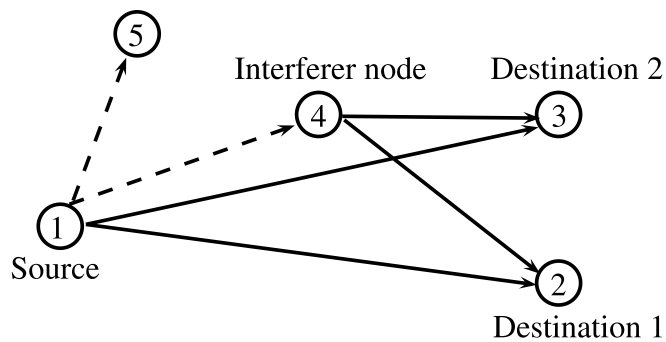

Fig. 1. Ad-hoc wireless networks with 5 nodes.

Transmission of information over noisy channels in the presence of interference has been widely studied. As mentioned above, one interesting problem is when the interference is known non-causally to the encoder. This problem is considered by Gel'fand and Pinsker for discrete memoryless channels [3]. In [4], Costa extends this result to the Gaussian case with additive interference. The authors in [5] investigate the transmission of a Gaussian source over a Gaussian channel with additive interference that is correlated to the source. In [6], the authors consider the joint transmission of source and interference over a Gaussian channel with additive interference known non-causally to the encoder; the optimal tradeoff between the source rate transmission and the mean square error (MSE) distortion from estimating the interference is characterized. In [7], the authors consider the same problem as in [6] but with imperfect knowledge of the interference at the transmitter side. In [2], we extend the joint source-interference transmission problem to the case of correlated source-interference and fading channels.

The traditional approach for analog source transmission over noisy channels is to use separate source and channel coders, also known as tandem coding [8]. This approach is optimal for point-to-point communications given unlimited delay and complexity in the coders. There are, however,

This work was supported in part by NSERC of Canada. The material in this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), Istanbul, Turkey, July 2013 [1].

A. Abou Saleh and W-Y. Chan are with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, K7L 3N6, Canada (e-mail: ahmad.abou.saleh@queensu.ca; chan@queensu.ca).

F. Alajaji is with the Department of Mathematics and Statistics, Queen's University, Kingston, ON, K7L 3N6, Canada (e-mail: fady@mast.queensu.ca).

two disadvantages associated with digital transmission: 1) the threshold effect and 2) the levelling-off effect. A family of hybrid digital-analog (HDA) schemes are introduced in [9]–[11] to overcome the threshold and the levelling-off effects. In [12]–[14], HDA schemes are also proposed for broadcast channels and Wyner-Ziv systems.

With the increasing popularity of wireless sensor networks, reliable transmission with low delay and complexity constraints is more relevant than ever. A sensor node, often conceived as having limited lifetime and processing power, communicates its sensed field information to a fusion centre over a noisy wireless channel. To meet these challenges, low delay analog coding, which tends to promote low power implementation, has been considered for several communication scenarios [15]–[21]. Memoryless low delay strategies are proposed for two-way relay channels in [22]. In [23], we propose to use analog coding techniques for the transmission of Gaussian sources over fading channels in the presence of interference. In [24], we tackle the problem of joint transmission of source and interference over a Gaussian channel in the presence of interference and propose to use low delay analog mappings. In [25], the authors propose low delay mappings for the problem of transmitting Gaussian sources over Gaussian channels in the presence of interference that is independent of the source.

For multi-terminal systems, tandem coding is no longer optimal; a joint source-channel coding (JSCC) scheme may be required to achieve optimality. One simple scenario where the tandem scheme is suboptimal concerns the broadcast of Gaussian sources over Gaussian channels [26]. For a single Gaussian source sent over a Gaussian broadcast channel with matched source-channel bandwidth, the optimal distortion region is known, and can be realized using a linear scheme [26]. For mismatched source-channel bandwidth, the best known coding schemes are based on JSCC with hybrid signalling [27]–[29]. One extension to this problem is the broadcasting of two correlated sources to two users, each of which is interested in recovering one of the two sources; in [30], it is proven that the linear scheme is optimal when the systems signal-to-noise ratio is below a certain threshold under matched bandwidth. In [31], a hybrid digital-analog scheme is proposed for the same matched bandwidth system and is shown to be optimal whenever the linear scheme of [30] is not, hence providing a complete characterization of the distortion region. Under mismatched bandwidth, various HDA schemes are proposed in [32], consisting of different combinations of several coding techniques using either superposition or dirty paper coding. Recently, in [33], a tandem scheme based on successive coding is studied and shown to outperform the HDA schemes of [32]. In [34], the authors extend the broadcast scenario to the case when the sources are vectors. In [1] we consider the transmission of a bivariate Gaussian source over a two-user channel in the presence of interference that is correlated to the sources; we derive inner and outer bounds on the source distortion region. In [35], [36], the authors investigate the problem of transmission of both messages and source/state sequences over broadcast and point-to-point channels, respectively.

In this work, we tackle the problem of joint recovery of a bivariate Gaussian source and of interference over the two-user broadcast channel in the presence of Gaussian interference that is known non-causally to the transmitter and is correlated to the source. More precisely, information theoretical limits are derived and low delay and complexity coding schemes are proposed based on analog coding. Different from previous work, our scenario considers broadcast channels, correlated interference and interference estimation at the receiver side. The recovery of source and interference finds application in sensor networks and relay channels. As illustrated in Fig. 1, the relay node amplifies the "interference" (another version of the main source) which interferes with the transmission of the source node; the sink node estimates all versions of the source (the main source and the interference). The rest of the paper is organized as follows. In Section II, we present the problem formulation and state the main contributions. In Section III, we derive an asymptotic outer bound on the system's distortion. In Section IV, we derive an asymptotic inner bound (achievable distortion region) by proposing a layered hybrid coding scheme. Section V presents practically implementable low-delay analog codes. Numerical results are included in Section VI. Finally, conclusions are drawn in Section VII.

II. Problem Formulation and Main Contributions

We consider the transmission of a pair of correlated Gaussian sources (V_1^K, V_2^K) over a two-user Gaussian broadcast channel in the presence of Gaussian interference S^K known non-causally to the transmitter (see Fig. 2), where $V_i^K = (V_i(1), ..., V_i(K)) \in \mathbb{R}^K$ is composed of independent and identically distributed (i.i.d.) samples, for $i = 1,\ 2$; S^K is similarly defined. The source pair vector (V_1^K, V_2^K) and the interference S^K are transformed into a K dimensional channel input $X^K \in \mathbb{R}^K$ via $\alpha(\cdot)$, a mapping from $(\mathbb{R}^K \times \mathbb{R}^K \times \mathbb{R}^K) \to \mathbb{R}^K$. User i $(i = 1, 2)$ receives the transmitted signal corrupted by additive white Gaussian noise W_i^K and interference S^K. The received vector at user i is given by

$$Y_i^K = X^K + S^K + W_i^K \quad (1)$$

where addition is component-wise, $X^K = \alpha(V_1^K, V_2^K, S^K)$, each sample in the interference vector S^K follows an i.i.d. Gaussian distribution with variance σ_S^2 $(S \sim \mathcal{N}(0, \sigma_S^2))$ and each sample in the additive noise W_i^K is drawn from an i.i.d. Gaussian distribution with variance $\sigma_{W_i}^2$ $(W_i \sim \mathcal{N}(0, \sigma_{W_i}^2))$ independently from both sources and interference.

Each user i aims to recover both the source V_i^K and the state interference S^K; the reconstructed source and interference at user i are denoted by $\hat{V}_i^K$ and $\hat{S}_i^K$, respectively. In this work, we assume that $(V_1(i), V_2(i), S(i)),\ i = 1, ..., K$, are correlated via the following covariance matrix

$$\Sigma_{V_1V_2S} = \begin{bmatrix} \sigma_{V_1}^2 & \rho_{V_1V_2}\sigma_{V_1}\sigma_{V_2} & \rho_{V_1S}\sigma_{V_1}\sigma_S \\ \rho_{V_1V_2}\sigma_{V_1}\sigma_{V_2} & \sigma_{V_2}^2 & \rho_{V_2S}\sigma_{V_2}\sigma_S \\ \rho_{V_1S}\sigma_{V_1}\sigma_S & \rho_{V_2S}\sigma_{V_2}\sigma_S & \sigma_S^2 \end{bmatrix} \quad (2)$$

where $\sigma_{V_1}^2$ and $\sigma_{V_2}^2$ are the variances of V_1 and V_2, respectively, $\rho_{V_1V_2}, \rho_{V_1S}$ and ρ_{V_2S} are the correlation coefficients between

V_1 and V_2, S and V_1 and S and V_2, respectively. The covariance matrix in (2) being assumed to be positive definite restricts the possible values of $\rho_{V_1V_2}, \rho_{V_1S}$ and ρ_{V_2S}.

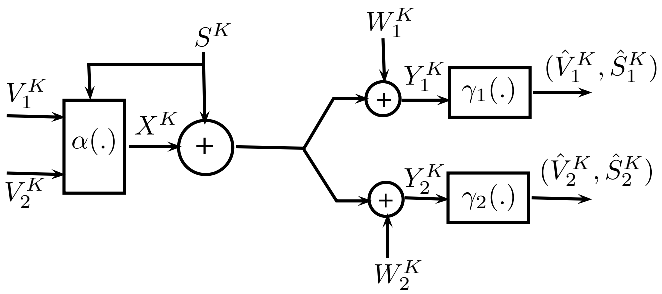

Fig. 2. System model structure.

The system operates under an average power constraint P given by

$$\frac{\mathbb{E}[||\alpha(V_1^K, V_2^K, S^K)||^2]}{K} \leq P \tag{3}$$

where $\mathbb{E}[(\cdot)]$ denotes the expectation operator. The reconstructed source and interference signals are given by $(\hat{V}_i^K, \hat{S}_i^K) = \gamma_i(Y_i^K) = (\gamma_i^{(v)}(Y_i^K), \gamma_i^{(s)}(Y_i^K))$, where the decoder functions $\gamma_i(.)$ are mappings from $\mathbb{R}^K \rightarrow (\mathbb{R}^K, \mathbb{R}^K)$. In this paper, we aim to find a source-channel encoder α and decoders γ_i $(i = 1, 2)$ that optimize the MSE distortion tradeoff from reconstructing the source and the state interference at both users. We first derive theoretical (outer and inner) asymptotic bounds on the optimal distortion region; these bounds give us a good understanding on the optimal system's behaviour (i.e., the optimal theoretical performance). We then construct efficient analog low-delay coding schemes that promote low power implementation and low complexity. At user i, the MSE distortion from reconstructing the source and the interference are denoted by

$$D_{v_i} = \frac{\mathbb{E}[||V_i^K - \hat{V}_i^K||^2]}{K}, \;\; D_{s_i} = \frac{\mathbb{E}[||S^K - \hat{S}_i^K||^2]}{K} \tag{4}$$

for $i = 1,\ 2$. We assume a degraded broadcast channel with $\sigma^2_{W_1} > \sigma^2_{W_2}$; hence user 1 is the weak user and user 2 is the strong one. For a given power constraint P, the distortion region is defined as the closure of all distortion quadruple $(\tilde{D}_{v_1}, \tilde{D}_{v_2}, \tilde{D}_{s_1}, \tilde{D}_{s_2})$ for which $(P, \tilde{D}_{v_1}, \tilde{D}_{v_2}, \tilde{D}_{s_1}, \tilde{D}_{s_2})$ is achievable, where a power-distortion set is achievable if for any $\delta > 0$, there exist sufficiently large integer K, encoding and decoding functions $(\alpha, \gamma_1, \gamma_2)$ satisfying (3), such that $D_{v_i} < \tilde{D}_{v_i} + \delta$ and $D_{s_i} < \tilde{D}_{s_i} + \delta$, for $i = 1,\ 2$. Our main contributions can be summarized as follows:

- We derive an outer bound on the system's distortion region for a Gaussian broadcast channel in the presence of interference. The outer bound on the distortion from reconstructing the source pairs is found by assuming additional knowledge about the source and the interference at the receiver side of each user; no additional knowledge is assumed when deriving a bound on the distortion from estimating the interference. The derived bound generalizes the one in [35] for the case of Gaussian correlated source-interference. Note that our problem resorts to several interesting scenarios studied in [1], [2], [6], [35]; numerical results show that the new derived outer bound is tighter than the ones obtained in [1] and [2] and is optimal for the Gaussian version of the problem considered in [6].
- Inner bounds are derived by proposing a hybrid coding scheme based on superposition coding [37], proper combination of power splitting, Wyner-Ziv [38] and Costa coding [4]. The proposed scheme reduces to the one in [31] which is optimal for the transmission of correlated sources over a Gaussian broadcast channel (with no interference) and to the one in [6] for the state amplification problem (for the case of uncorrelated interference). Moreover, the proposed encoder structure has a lower complexity than the one proposed in [1] with no loss in performance.
- After deriving information theoretical bounds, we propose low delay and complexity coding schemes based on analog coding. We first study a linear scheme and prove that the performance of any linear scheme is achieved using a single-letter linear code. To benefit from nonlinearity whenever possible, we then study a parametric analog mapping based on the sawtooth (modulo) function. We derive an upper bound on the system's distortion by assuming a suboptimal decoder at the receiver side; the optimized system parameters are found by minimizing the derived upper bound expression. Finally, whenever storage and offline design complexity are not an issue, a nonparametric mapping is designed; this is done by deriving the necessary conditions for optimality and proposing an iterative algorithm based on joint optimization between the transmitter and the receivers.

Relating Fig. 2 to Fig. 1, the interference S^K represents the transmission from the interferer node (which is an amplified version of the first sample of the collected measurements); the source pair (V_1^K, V_2^K), however, represents the remaining two samples of the collected measurements. Note that all measurements are correlated.

III. Distortion Outer Bound

In [39] and [32], outer bounds on the distortion region for sending correlated sources over the Gaussian degraded broadcast channel are obtained for the matched and mismatched source-channel bandwidth cases, respectively; this is done by assuming additional knowledge of the source V_1^K at the strong user. In [5], [40], several bounds are derived for Gaussian channels in the presence of interference that is known non-causally to the transmitter and correlated to the source. In [2], we derive an outer bound for the joint source-interference transmission problem over fading channels in the presence of interference. In [1], we derive an outer bound for the source transmission problem over Gaussian broadcast channels in the presence of interference. We herein generalize the problem considered in [1] and establish an outer bound on the distortion region of (V_1^K, V_2^K, S^K) for the interference broadcast channel. Since $S(i)$ and $V_1(i)$ are correlated for $i = 1, ..., K$, we have $S(i) = S_I(i) + S_D(i)$, with $S_D(i) = \frac{\rho_{V_1S}\sigma_S}{\sigma_{V_1}} V_1(i)$ and $S_I \sim \mathcal{N}(0, (1 - \rho^2_{V_1S})\sigma^2_S)$. Note that S_I and S_D are independent of each other. To derive a bound on D_{v_1} and D_{v_2}, we assume knowledge of a noisy version of V_1^K at the strong user and $\tilde{S}^K$, a partial knowledge about the interference, at both users, where $\tilde{S}^K = \beta_1 S_I^K + \beta_2 S_D^K$, β_1 and β_2 are real parameters with $\beta_1 \neq 0$; the noisy version of V_1^K is denoted by $\tilde{V}_1^K = \zeta V_1^K + \mathcal{V}^K$, where $\zeta \in [0,\ 1]$ and each sample of $\mathcal{V}^K$ is i.i.d. Gaussian with variance $\sigma_{V_1}^2(1-\zeta^2)$ that is independent of the source and the interference. The bounds on D_{s_1} and D_{s_2} are obtained by bounding the mutual information $I(S^K; \hat{S}_i^K)$ for $i = 1, 2$.

Definition 1 *Let MSE$(Y_2; \tilde{V}_1, \tilde{S})$ be the distortion incurred from estimating Y_2 based on $(\tilde{V}_1, \tilde{S})$ using a linear minimum MSE (LMMSE) estimator denoted by $\gamma_{lmse}(Y_2|\tilde{V}_1, \tilde{S})$ and MSE$(Y_1; \tilde{S})$ be the distortion incurred from estimating Y_1 using an LMMSE estimator based on $\tilde{S}$ denoted by $\gamma_{lmse}(Y_1|\tilde{S})$. These distortions, which are a function of β_1, β_2, ζ, $\mathbb{E}[XS_I]$ and $\mathbb{E}[XS_D]$, are given by MSE$(Y_2; \tilde{V}_1, \tilde{S}) = \mathbb{E}[(Y_2 - \gamma_{lmse}(Y_2|\tilde{V}_1, \tilde{S}))^2] = \mathbb{E}[Y_2^2] - \Gamma_b \Lambda_{\tilde{V}_1\tilde{S}}^{-1} \Gamma_b^T$ and MSE$(Y_1; \tilde{S}) = \mathbb{E}[(Y_1 - \gamma_{lmse}(Y_1|\tilde{S}))^2] = \left(\mathbb{E}[Y_1^2] - \frac{(\mathbb{E}[Y_1\tilde{S}])^2}{\mathbb{E}[\tilde{S}^2]}\right)$, where Γ_b is the correlation vector between Y_2 and $[\tilde{V}_1 \quad \tilde{S}]$, $\Lambda_{\tilde{V}_1\tilde{S}}$ is the covariance matrix of $[\tilde{V}_1 \quad \tilde{S}]$, $\mathbb{E}[Y_i^2] = P + \sigma_S^2 + 2(\mathbb{E}[XS_I + XS_D]) + \sigma_{W_i}^2$ for $i = 1, 2$, $\mathbb{E}[Y_1\tilde{S}] = \mathbb{E}[X(\beta_1 S_I + \beta_2 S_D)] + \mathbb{E}[\beta_1 S_I^2 + \beta_2 S_D^2]$ and $\mathbb{E}[\tilde{S}^2] = \mathbb{E}[\beta_1^2 S_I^2 + \beta_2^2 S_D^2]$. The set $\mathcal{A}$ denotes all i.i.d. Gaussian channel inputs X such that*

$$\begin{aligned} h(Y_2^K|S^K) &= \frac{K}{2}\log 2\pi e(\eta_1 P + \sigma_{W_2}^2), \\ h(Y_2^K|V_1^K, S^K) &= \frac{K}{2}\log 2\pi e(\eta_2 P + \sigma_{W_2}^2) \end{aligned} \tag{5}$$

for some $\eta_1 \in [0,\ 1]$ and $\eta_2 \in [0,\ \eta_1]$. These terms will be used next in Theorem 1.

Theorem 1 *The outer bound on the distortion region can be expressed as follows*

$$\begin{aligned} D_{v_2} &\geq \frac{Var(V_2|\tilde{V}_1, \tilde{S})\sigma_{W_2}^2}{\sup_{\mathbb{E}[XS_I],\mathbb{E}[XS_D]:X\in\mathcal{A}} MSE(Y_2; \tilde{V}_1, \tilde{S})}, \\ D_{s_i} &\geq \frac{\sigma_S^2(\eta_1 P + \sigma_{W_i}^2)}{P + \sigma_S^2 + \sigma_{W_i}^2 + 2\sqrt{\sigma_S^2(1-\eta_1)P}}, i = 1, 2, \\ D_{v_1} &\geq \frac{Var(V_1|\tilde{S})(\eta_2 P + \sigma_{W_1}^2)}{\sup_{\mathbb{E}[XS_I],\mathbb{E}[XS_D]:X\in\mathcal{A}} MSE(Y_1; \tilde{S})} \end{aligned} \tag{6}$$

where $Var(V_2|\tilde{V}_1, \tilde{S}) = \sigma_{V_2}^2 - \Gamma_a \Lambda_{\tilde{V}_1\tilde{S}}^{-1} \Gamma_a^T$ is the variance of V_2 given $(\tilde{V}_1, \tilde{S})$ with Γ_a being the correlation vector between V_2 and $[\tilde{V}_1 \quad \tilde{S}]$, $Var(V_1|\tilde{S}) = \sigma_{V_1}^2\left(1 - \frac{\beta_2^2\rho_{V_1S}^2}{\beta_1^2(1-\rho_{V_1S}^2)+\beta_2^2\rho_{V_1S}^2}\right)$ is the variance of V_1 given $\tilde{S}$, $\eta_1 \in [0, 1]$ and $\eta_2 \in [0, \eta_1]$. Note that we need to optimize the outer bound over the parameters β_1, β_2 and ζ.

Proof: To find a lower bound on D_{v_2}, we assume knowledge of $\tilde{V}_1^K$ and $\tilde{S}^K$ at the strong user (user 2). As a result, we can write the following

$$\begin{aligned} &\frac{K}{2}\log\frac{\text{Var}(V_2|\tilde{V}_1, \tilde{S})}{D_{v_2}} \overset{(a)}{\leq} I(V_2^K; \hat{V}_2^K|\tilde{V}_1^K, \tilde{S}^K) \\ &\overset{(b)}{\leq} I(V_2^K; Y_2^K|\tilde{V}_1^K, \tilde{S}^K) \\ &= h(Y_2^K|\tilde{V}_1^K, \tilde{S}^K) - h(Y_2^K|\tilde{V}_1^K, V_2^K, \tilde{S}^K) \\ &\overset{(c)}{=} h\left(Y_2^K - \gamma_{lmse}(Y_2^K|\tilde{V}_1^K, \tilde{S}^K)|\tilde{V}_1^K, \tilde{S}^K\right) \\ &-h(Y_2^K|\tilde{V}_1^K, V_2^K, \tilde{S}^K) \\ &\overset{(d)}{\leq} h\left(Y_2^K - \gamma_{lmse}(Y_2^K|\tilde{V}_1^K, \tilde{S}^K)\right) \\ &-h(Y_2^K|\tilde{V}_1^K, V_2^K, \tilde{S}^K, V_1^K) \\ &= h\left(Y_2^K - \gamma_{lmse}(Y_2^K|\tilde{V}_1^K, \tilde{S}^K)\right) - h(Y_2^K|V_2^K, S^K, V_1^K) \\ &\overset{(e)}{\leq} \sup_{\mathbb{E}[XS_I],\mathbb{E}[XS_D]:X\in\mathcal{A}} \frac{K}{2}\log 2\pi e\text{MSE}(Y_2; \tilde{V}_1, \tilde{S}) \\ &-\frac{K}{2}\log 2\pi e(\sigma_{W_2}^2) \\ &= \frac{K}{2}\log\frac{\sup_{\mathbb{E}[XS_I],\mathbb{E}[XS_D]:X\in\mathcal{A}}\text{MSE}(Y_2; \tilde{V}_1, \tilde{S})}{\sigma_{W_2}^2} \end{aligned} \tag{7}$$

where (a) follows from the rate-distortion theorem, (b) holds by the data processing inequality, (c) uses the fact that differential entropy is invariant under translations and γ_{lmse} is a symbol-by-symbol LMMSE estimate, (d) follows from the fact that conditioning reduces differential entropy and (e) uses the fact that the Gaussian distribution maximizes differential entropy; the set $\mathcal{A}$ is as given in Definition 1. The reason for restricting the search space of the supremum over the set $\mathcal{A}$ is detailed later; solving the two constraints of set $\mathcal{A}$, as given in (5), leads to the following equations

$$\begin{aligned} &\mathbb{E}[XS]^2 = \sigma_S^2(1-\eta_1)P, \\ &\frac{\mathbb{E}[XV_1]^2}{\sigma_{V_1}^2} + \frac{\mathbb{E}[XS_I]^2}{\mathbb{E}[S_I^2]} = (1-\eta_2)P. \end{aligned} \tag{8}$$

Note that the constraint on the conditional differential entropy $h(Y_2^K|S^K)$ of the set $\mathcal{A}$ uses the fact that we can bound $h(Y_2^K|S^K)$ as follows

$$\begin{aligned} \frac{K}{2}\log 2\pi e\sigma_{W_2}^2 &\leq h(Y_2^K|S^K) \leq h(X^K + W_2^K) \\ &\leq \frac{K}{2}\log 2\pi e(P + \sigma_{W_2}^2). \end{aligned} \tag{9}$$

Hence, there exists a parameter $\eta_1 \in [0,\ 1]$ such that $h(Y_2^K|S^K) = \frac{K}{2}\log 2\pi e(\eta_1 P + \sigma_{W_2}^2)$. Similarly, the constraint on $h(Y_2^N|V_1^K, S^K)$ of the set $\mathcal{A}$ can be obtained by noting that

$$\begin{aligned} \frac{K}{2}\log 2\pi e\sigma_{W_2}^2 &\leq h(Y_2^N|V_1^K, S^K) \leq h(Y_2^K|S^K) \\ &= \frac{K}{2}\log 2\pi e(\eta_1 P + \sigma_{W_2}^2). \end{aligned} \tag{10}$$

As a result, there exists a parameter $\eta_2 \in [0,\ \eta_1]$ such that $h(Y_2^N|V_1^K, S^K) = \frac{K}{2}\log 2\pi e(\eta_2 P + \sigma_{W_2}^2)$. These two constraints (or equalities) are used later to bound the distortion from reconstructing the source V_1 and the interference S at the weak and the strong user, hence the reason for restricting the supremum search space over these two constraints in (7).

To get a bound on D_{s_2}, we can write the following

$$\begin{aligned}
&\frac{K}{2}\log\frac{\sigma_S^2}{D_{s_2}} \overset{(a)}{\leq} I(S^K;\hat{S}_2^K) \overset{(b)}{\leq} I(S^K;Y_2^K)\\
&= h(Y_2^K) - h(Y_2^K|S^K)\\
&\overset{(c)}{\leq} \sup_{\mathbb{E}[XS]:X\in\mathcal{A}} \left\{\frac{K}{2}\log 2\pi e(P+\sigma_S^2+\sigma_{W_2}^2+2\mathbb{E}[XS])\right\}\\
&-\frac{K}{2}\log 2\pi e(\eta_1 P+\sigma_{W_2}^2)\\
&= \frac{K}{2}\log\frac{P+\sigma_S^2+\sigma_{W_2}^2+2\sqrt{\sigma_S^2(1-\eta_1)P}}{\eta_1 P+\sigma_{W_2}^2} \qquad (11)
\end{aligned}$$

where (a) follows from the rate-distortion theorem, (b) holds by the data processing inequality, (c) follows from the fact that the Gaussian distribution maximizes differential entropy and by using the constraint in (8). Note that the last inequality in (11) uses the fact that there exists an $\eta_1 \in [0\ \ 1]$ such that $h(Y_2^K|S^K) = \frac{K}{2}\log 2\pi e(\eta_1 P+\sigma_{W_2}^2)$ as shown previously using (9).

To get a bound on the distortion from estimating V_1^K, we assume knowledge of $\tilde{S}^K$ at the weak user. As a result, we can write the following

$$\begin{aligned}
&\frac{K}{2}\log\frac{\text{Var}(V_1|\tilde{S})}{D_{v_1}} \overset{(a)}{\leq} I(V_1^K;\hat{V}_1^K|\tilde{S}^K) \overset{(b)}{\leq} I(V_1^K;Y_1^K|\tilde{S}^K)\\
&= h(Y_1^K|\tilde{S}^K) - h(Y_1^K|V_1^K,\tilde{S}^K)\\
&= h(Y_1^K|\tilde{S}^K) - h(Y_1^K|V_1^K,S^K)\\
&\overset{(c)}{\leq} h(Y_1^K-\gamma_{lmse}(Y_1^K|\tilde{S}^K)) - h(Y_1^K|V_1^K,S^K)\\
&\overset{(d)}{\leq} \sup_{\mathbb{E}[XS_I],\mathbb{E}[XS_D]:X\in\mathcal{A}} \frac{K}{2}\log 2\pi e\text{MSE}(Y_1;\tilde{S})\\
&-\frac{K}{2}\log 2\pi e(\eta_2 P+\sigma_{W_1}^2)\\
&= \frac{K}{2}\log\frac{\sup_{\mathbb{E}[XS_I],\mathbb{E}[XS_D]:X\in\mathcal{A}}(\text{MSE}(Y_1;\tilde{S}))}{\eta_2 P+\sigma_{W_1}^2} \qquad (12)
\end{aligned}$$

where (a) follows from the rate-distortion theorem, (b) holds by the data processing theorem, (c) follows from the fact that conditioning reduces differential entropy and $\gamma_{lmse}(Y_1^K|\tilde{S}^K)$ is the symbol-by-symbol LMMSE estimate of Y_1^K based on $\tilde{S}^K$, and in (d) we use the facts that the Gaussian distribution maximizes differential entropy and that $h(Y_1^K|V_1^K,S^K) \geq \frac{K}{2}\log 2\pi e(\eta_2 P+\sigma_{W_1}^2)$; this can be proved by noting that $Y_1^K = Y_2^K + Z^K$ with $Z \sim \mathcal{N}(0,\sigma_Z^2 = \sigma_{W_1}^2-\sigma_{W_2}^2)$ is independent of Y_2 (degraded broadcast channel) and using the entropy power inequality as follows

$$\begin{aligned}
2^{\frac{2}{K}h(Y_1^K|V_1^K,S^K)} &\geq 2^{\frac{2}{K}h(Y_2^K|V_1^K,S^K)} + 2^{\frac{2}{K}h(Z^K|V_1^K,S^K)}\\
&= 2\pi e(\eta_2 P+\sigma_{W_2}^2) + 2\pi e(\sigma_Z^2). \qquad (13)
\end{aligned}$$

After some manipulation, (13) can be written as $h(Y_1^K|V_1^K,S^K) \geq \frac{K}{2}\log 2\pi e(\eta_2 P+\sigma_{W_1}^2)$.

To get a bound on D_{s_1}, we first write the following

$$\begin{aligned}
&\frac{K}{2}\log\frac{\sigma_S^2}{D_{s_1}} \overset{(a)}{\leq} I(S^K;\hat{S}_1^K) \overset{(b)}{\leq} I(S^K;Y_1^K)\\
&= h(Y_1^K) - h(Y_1^K|S^K)\\
&\overset{(c)}{\leq} \sup_{\mathbb{E}[XS]:X\in\mathcal{A}} \left\{\frac{K}{2}\log 2\pi e(P+\sigma_S^2+\sigma_{W_1}^2+2\mathbb{E}[XS])\right\}\\
&-\frac{K}{2}\log 2\pi e(\eta_1 P+\sigma_{W_1}^2)\\
&= \frac{K}{2}\log\frac{P+\sigma_S^2+\sigma_{W_1}^2+2\sqrt{\sigma_S^2(1-\eta_1)P}}{\eta_1 P+\sigma_{W_2}^2} \qquad (14)
\end{aligned}$$

where (a) follows from the rate-distortion theorem, (b) holds by the data processing theorem and (c) holds since the Gaussian distribution maximizes differential entropy; we also use in (c) the fact that $h(Y_1^K|S^K) \geq \frac{K}{2}\log 2\pi e(\eta_1 P+\sigma_{W_1}^2)$ due to the entropy power inequality and since $Y_1^K = Y_2^K + Z^K$. ■

Remark 1 *Recall that in our problem, the interference and the source pair are considered as correlated measurements collected by a sensor node. Interference is called as such due to the relay setup considered in Fig. 1. Hence, as already stated, in this work we present results related to estimating both source pair and interference. Moreover, our problem reduces to several problems considered in the literature under certain system settings. The derived outer bound given by* (6) *can indeed be simplified to get outer bounds for those special cases. More specifically, by properly choosing the system parameters, our bound in* (6) *recovers the distortion bounds of the special cases considered in [6], [31], [32], [40]. Furthermore, in comparison with the bounds derived in [1], [5], our bound is observed numerically to be tighter.*[1] *We herein specify the conditions under which the bound in* (6) *reduces to existing bounds.*

- *In [6], the authors consider the transmission of both source and interference over single-user channels where the source and the interference are independent from each other. By focusing on (D_{v_1}, D_s) in our bound and setting $\eta_1 = \eta_2 = 0$, $\beta_1 = \beta_2 = 1$ and $Y_1 = Y_2$ in* (6)*, we get the optimal bound derived in [6]. Note that in [7] the authors assume noisy knowledge of the interference; this is the main difference from the problem setup in [6].*
- *In [31], [32], the authors consider the broadcasting of bivariate Gaussian sources over Gaussian channels; no interference is assumed. The bound on the source reconstruction can be derived by setting $\eta_1 = 1$ and $\tilde{S} = \emptyset$ in* (6)*. Note that this bound is optimal over certain regions.*
- *In [40], the authors consider the transmission of the source over single-user channels in the presence of interference that is correlated to the source. Focusing on D_{v_1} and setting $\eta_2 = 0$ and $Y_1 = Y_2$ in* (6) *yield the bound derived in [40].*

[1]For example, see Fig. 6 for a comparison with the bound in [1].

IV. Distortion Inner Bound

In this section, we present a distortion inner bound (achievable distortion region) on $(D_{v_1}, D_{v_2}, D_{s_1}, D_{s_2})$ by proposing a hybrid scheme that uses superposition, Wyner-Ziv and Costa coding.

A. Hybrid Scheme

As shown from the encoder structure in Fig. 3, this scheme has four layers that are merged to output X^K. The first layer which outputs $X_a^K = \sqrt{a}(\alpha_{11}V_1^K + \alpha_{12}V_2^K + \alpha_{13}S^K)$, a linear combination (LC) of the sources and the interference, is meant for all receivers and benefit from the correlation between the sources and the interference, where parameters $\alpha_{11}, \alpha_{12}, \alpha_{13} \in [-1,\ 1]$ and $a = P_a/(\alpha_{11}^2\sigma_{V_1}^2 + \alpha_{12}^2\sigma_{V_2}^2 + \alpha_{13}^2\sigma_S^2 + 2\alpha_{11}\alpha_{12}\rho_{V_1V_2}\sigma_{V_1}\sigma_{V_2} + 2\alpha_{11}\alpha_{13}\rho_{V_1S}\sigma_{V_1}\sigma_S + 2\alpha_{12}\alpha_{13}\rho_{V_2S}\sigma_{V_2}\sigma_S)$ is a gain factor related to power constraint P_a. The second layer which outputs X_q^K employs a source-channel quantizer on the source V_1^K; the output of this layer is given by $X_q^K = \mu(V_1^K + U_q^K)$, where $\mu \geq 0$ is a gain factor related to the power constraint and each sample in U_q^K follows a zero mean i.i.d. Gaussian that is independent of the sources and the interference and has a variance Q. A similar vector quantizer (VQ) encoder was used in [31] for the broadcast of bivariate sources and in [41] for the multiple access channel. In what follows, we summarize the encoding process of the VQ layer

- *Codebook Generation:* Generate a K-length i.i.d. Gaussian codebook $\mathcal{X}_q$ with 2^{KR_q} codewords with R_q defined later. Every codeword is generated following the random variable X_q^K; this codebook is revealed to both encoder and decoders.
- *Encoding:* The encoder searches for a codeword X_q^K in the codebook that is jointly typical with V_1^K. In case of success, the transmitter sends X_q^K.

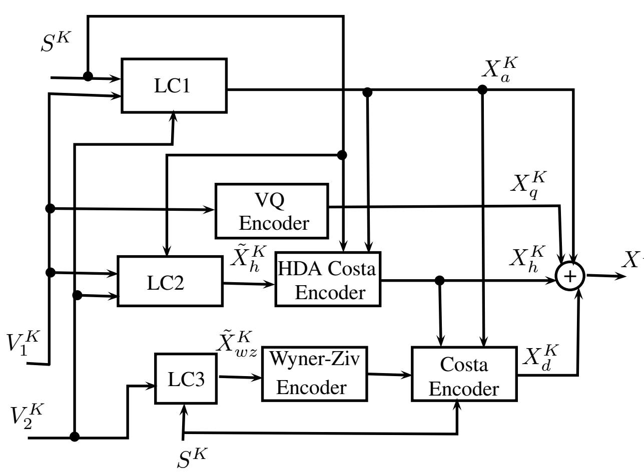

Fig. 3. Encoder structure. LC represents the linear combination operation. X_a^K, X_q^K, X_h^K and X_d^K are the output of the first, second, third and fourth layers, respectively.

The third layer first forms a linear combination of the sources and the interference $\tilde{X}_h^K = \alpha_{31}V_1^K + \alpha_{32}V_2^K + \alpha_{33}S^K$, where α_{31}, α_{32} and $\alpha_{33} \in [-1,\ 1]$; this linear combination $\tilde{X}_h^K$ is then encoded using an HDA Costa encoder [42] which treats X_a^K and S^K as known interference and is meant for both users. The HDA Costa encoder, which uses an average power of P_h, forms a codebook $\mathcal{U}_h$ with codeword length K and 2^{KR_h} codewords, where R_h is defined later. Each codeword follows the auxiliary random variable

$$U_h^K = X_h^K + \alpha_h S'^K + \kappa_h \tilde{X}_h^K \tag{15}$$

where $S'^K = (X_a^K + S^K)$ is the interference, the samples in X_h^K are i.i.d. Gaussian with variance P_h, $\alpha_h = P_h/(P - \mathbb{E}[(X_a + X_q)^2] + \sigma_{W_1}^2)$ and κ_h is defined later. Motivated by the work of Costa [4], the coefficient α_h is chosen such that $\alpha_h(X_h + X_d + W_1)$ is the MMSE estimator of X_h given $(X_h + X_d + W_1)$, where X_d is the output of the last layer in Fig. 3. The fourth layer is purely digital and is meant for the strong user only. This layer starts by forming a linear combination of the source V_2^K and the interference S^K denoted by $\tilde{X}_{wz}^K = \alpha_{41}V_2^K + \alpha_{42}S^K$, where α_{41} and $\alpha_{42} \in [-1,\ 1]$. The output of the linear combination LC3, $\tilde{X}_{wz}^K$, is then encoded using a Wyner-Ziv [38] with rate R_d followed by a Costa coder. The Costa coder uses an average power of P_d and treats X_a^K, X_h^K and S^K as known interference. The Wyner-Ziv encoder forms an auxiliary random variable as follows

$$T^K = \alpha_{wz}\tilde{X}_{wz} + H^K \tag{16}$$

where the samples in H^K are zero mean i.i.d. Gaussian, the parameter α_{wz} and the variance of H are defined later. The encoding process of the Wyner-Ziv starts by generating a K length i.i.d. Gaussian codebook $\mathcal{T}$ of size $2^{KI(T;\tilde{X}_{wz})}$ and randomly assigning the codewords into 2^{KR_d} bins with R_d defined later. For each realization $\tilde{X}_{wz}^K$, the Wyner-Ziv encoder searches for a codeword $T^K \in \mathcal{T}$ such that $(\tilde{X}_{wz}^K, T^K)$ are jointly typical. In the case of success, the Wyner-Ziv encoder transmits the bin index of this codeword using Costa coding with rate $R_d = \frac{1}{2}\log\left(1 + \frac{P_d}{\sigma_{W_2}^2}\right)$. The Costa coder that treats $\tilde{S}^K = X_a^K + X_h^K + S^K$ as known interference, forms the following auxiliary random variable

$$U_d^K = X_d^K + \alpha_c \tilde{S}^K \tag{17}$$

where each sample in X_d^K follows $\mathcal{N}(0, P_d)$ that is independent of the sources and the interference and $\alpha_c = P_d/(P_d + \sigma_{W_2}^2)$. In a nutshell, the intuition of using such layered scheme is to benefit from the correlation between the sources and the interference; this is manifested by using a linear layer and a VQ layer which are beneficial for transmitting over broadcast channels. The third layer, on the other hand, is needed to "mitigate" the channel interference. The last layer which has a similar effect as the third layer is used to refine the estimates at the strong user.

As shown from the receiver structure of the weak user in Fig. 4, the VQ decoder estimates X_q^K by searching for a codeword $\hat{X}_q^K \in \mathcal{X}_q$ that is jointly typical with the received signal Y_1^K. Following the error analysis in [31], the error probability of decoding X_q^K goes to zero as $K \to \infty$ by choosing the rate R_q to satisfy the following constraint $I(V_1; X_q) \leq R_q \leq I(X_q; Y_1)$. The variance Q of the random variable U_q has to be chosen to satisfy the above rate constraint. Furthermore, to ensure the power constraint we need

μ to satisfy $P_a+\mu^2(\sigma^2_{V_1}+Q)+2\mu\mathbb{E}[V_1X_a]+P_h+P_d \leq P$. The HDA Costa decoder then searches for a codeword $\hat{U}_h^K$ that is jointly typical with $\tilde{Y}_1^K = Y_1^K - \hat{X}_q^K$ and $\hat{X}_q^K$. The parameter κ_h in (15) has to be chosen to satisfy the rate constraint $I(U_h; S', \tilde{X}_h) \leq R_h \leq I(U_h; \tilde{Y}_1, X_q)$. Using the above constraint on the rate R_h, the error of probability of decoding U_h^K goes to zero as $K \to \infty$. We then employ a linear MMSE estimator based on $\tilde{Y}_1^K, \hat{X}_q^K$ and $\hat{U}_h^K$ to obtain an estimate of V_1^K and S^K at the weak user. Mathematically, the estimate of the source and the interference are given by $\hat{V}_1^K = \Gamma_{v_1}\Lambda_1^{-1}[\hat{X}_q \;\; \hat{U}_h \;\; \tilde{Y}_1]$ and $\hat{S}_1^K = \Gamma_{s_1}\Lambda_1^{-1}[\hat{X}_q \;\; \hat{U}_h \;\; \tilde{Y}_1]$, where Γ_{v_1} is the correlation vector between V_1 and $[\hat{X}_q \;\; \hat{U}_h \;\; \tilde{Y}_1]$, Γ_{s_1} is the correlation vector between S and $[\hat{X}_q \;\; \hat{U}_h \;\; \tilde{Y}_1]$ and Λ_1 is the covariance matrix of $[\hat{X}_q \;\; \hat{U}_h \;\; \tilde{Y}_1]$. As a result, the achievable distortion at the weak user can be expressed as follows

$$\begin{aligned} D_{v_1}^{(Hybrid)} &= \sigma^2_{V_1} - \Gamma_{v_1}\Lambda_1^{-1}\Gamma_{v_1}^T \\ D_{s_1}^{(Hybrid)} &= \sigma^2_S - \Gamma_{s_1}\Lambda_1^{-1}\Gamma_{s_1}^T. \end{aligned} \tag{18}$$

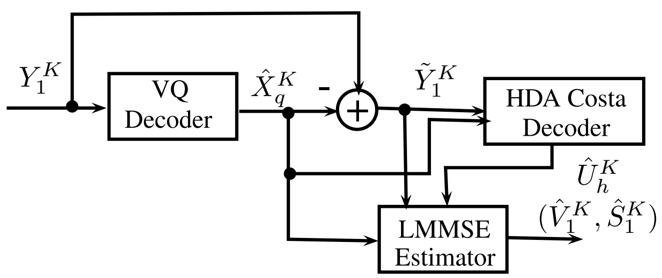

Fig. 4. Decoder structure of the weak user.

As shown from the receiver structure in Fig. 5, the strong user, that is able to decode all codewords used by the weak user, can refine the estimates of its source V_2^K and interference S^K using Wyner-Ziv decoder. First an estimate of $\tilde{X}_{wz}^K$, denoted by $\hat{\tilde{X}}_{wz}^K$, is obtained using an LMMSE estimator based on the decoded codewords $\hat{X}_q^K, \hat{U}_h^K$ and $\tilde{Y}_2^K$, where $\tilde{Y}_2^K = Y_2^K - \hat{X}_q^K$. The distortion from reconstruction $\tilde{X}_{wz}^K$ can be expressed as follows

$$\tilde{D}_{\tilde{x}_{wz}} = \mathbb{E}[\tilde{X}_{wz}^2] - \Gamma_{\tilde{x}_{wz}}\Lambda_{\tilde{x}_{wz}}^{-1}\Gamma_{\tilde{x}_{wz}}^T \tag{19}$$

where $\Gamma_{\tilde{x}_{wz}}$ is the correlation vector between $\tilde{X}_{wz}$ and $[\hat{X}_q \;\; \hat{U}_h \;\; \tilde{Y}_2]$ and $\Lambda_{\tilde{x}_{wz}}$ is the covariance matrix of $[\hat{X}_q \;\; \hat{U}_h \;\; \tilde{Y}_2]$. The Wyner-Ziv decoder then estimates the codeword T^K by searching for a $\hat{T}^K \in \mathcal{T}$ that is jointly typical with $\hat{\tilde{X}}_{wz}^K$. The error probability of decoding T^K vanishes as $K \to \infty$ with the chosen rate R_d. Note that a better estimate of $\tilde{X}_{wz}$ can be obtained using the Wyner-Ziv codeword. The distortion from reconstructing X_{wz}^K is given by

$$D_{\tilde{x}_{wz}} = \frac{\tilde{D}_{\tilde{x}_{wz}}}{1+\frac{P_d}{\sigma^2_{W_2}}}. \tag{20}$$

This distortion can be achieved using a linear MMSE estimator based on $\hat{X}_q^K, \hat{U}_h^K, \hat{T}^K, \tilde{Y}_2^K$ and by choosing $\alpha_{wz} = \sqrt{1-\frac{D_{\tilde{x}_{wz}}}{\tilde{D}_{\tilde{x}_{wz}}}}$ and $H \sim \mathcal{N}(0, D_{\tilde{x}_{wz}})$ in (16). To get an estimate of V_2^K and S^K at the strong user, a linear MMSE estimator is then used based on the decoded codewords $\hat{X}_q^K, \hat{U}_h^K, \hat{T}^K$ and $\tilde{Y}_2^K$. The distortion from reconstructing V_2^K and S^K at the strong user can then be expressed as follows

$$D_{v_2} = \sigma^2_{V_2} - \Gamma_{v_2}\Lambda_2^{-1}\Gamma_{v_2}^T, \;\; D_{s_2} = \sigma^2_S - \Gamma_{s_2}\Lambda_2^{-1}\Gamma_{s_2}^T \tag{21}$$

where Γ_{v_2} is the correlation vector between V_2 and $[\hat{X}_q \;\; \hat{U}_h \;\; \hat{T}^K \;\; \tilde{Y}_2]$, Γ_{s_2} is the correlation vector between S and $[\hat{X}_q \;\; \hat{U}_h \;\; \hat{T}^K \;\; \tilde{Y}_2]$ and Λ_2 is the covariance matrix of $[\hat{X}_q \;\; \hat{U}_h \;\; \hat{T}^K \;\; \tilde{Y}_2]$. After some manipulations, the distortions in (21) can be simplified as follows

$$\begin{aligned} D_{v_2}^{(Hybrid)} &= \sigma^2_{V_2}\left[1-\rho^2_{V_2\tilde{X}_{wz}}\left(1-\frac{D_{\tilde{x}_{wz}}}{\mathbb{E}[\tilde{X}_{wz}^2]}\right)\right] \\ D_{s_2}^{(Hybrid)} &= \sigma^2_S\left[1-\rho^2_{S\tilde{X}_{wz}}\left(1-\frac{D_{\tilde{x}_{wz}}}{\mathbb{E}[\tilde{X}_{wz}^2]}\right)\right] \end{aligned} \tag{22}$$

where $\rho_{V_2\tilde{X}_{wz}}$ is the correlation coefficient between V_2 and $\tilde{X}_{wz}$ and $\rho_{S\tilde{X}_{wz}}$ is the correlation coefficient between S and $\tilde{X}_{wz}$.

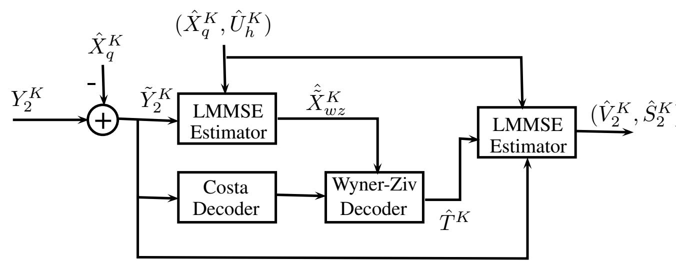

Fig. 5. Decoder structure of the strong user. Note that codewords $\hat{X}_q^K$ and $\hat{U}_h^K$ are estimated in a similar way as done at the weak user.

Remark 2 *By numerically calculating the distortions in* (18) *and* (22)*, we noticed that including the knowledge of* S^K *(in the form of linear combination of the sources and the interference) as an input to the HDA Costa and Wyner-Ziv coders is not beneficial in most scenarios.*

Remark 3 *Note that the distortion analysis in this section is not affected by restricting the linear combination parameters to* $[-1,\ 1]$*; these parameters can be assumed to take on any real values. However, from our numerical results, we did not notice any performance loss due to restraining the values of these parameters to* $[-1,\ 1]$.

Remark 4 *Note that our scheme resorts to the optimal schemes proposed in [6], [31] for the joint source-interference recovery over point-to-point Gaussian channels and for the transmission of bivariate Gaussian sources over degraded broadcast Gaussian channels, respectively. Moreover, we now use only one purely digital layer (Wyner-Ziv followed by Costa coding) as opposed to the scheme we proposed in [1].*

V. Low Delay Analog Codes

In this section, we propose low delay coding schemes (with $K = 1$) based on analog mappings. The objective is to design the encoder and decoder structures in order to minimize the overall weighted MSE distortion

$$D = \sum_{i=1}^{2} (\theta_{v_i} D_{v_i} + \theta_{s_i} D_{s_i}) \tag{23}$$

subject to the power constraint in (3) and $\sum_{i=1}^{2}(\theta_{v_i}+\theta_{s_i}) = 1$, where D_{v_i} and D_{s_i} are defined in (4). Note that θ_{v_i} and θ_{s_i} are set by the system designer and dictate the operational region of the system.

A. Linear Scheme

We first assume that the encoder transforms the source pair (V_1, V_2) and the interference S into a channel input X using a linear transformation according to

$$X = \alpha(V_1, V_2, S) = \sqrt{a_1}(\alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 S) \tag{24}$$

where $\alpha_1, \alpha_2, \alpha_3 \in [-1,\ 1]$ and $a_1 = P/(\alpha_1^2\sigma_{V_1}^2 + \alpha_2^2\sigma_{V_2}^2 + \alpha_3^2\sigma_S^2 + 2\alpha_1\alpha_2\rho_{V_1V_2}\sigma_{V_1}\sigma_{V_2} + 2\alpha_1\alpha_3\rho_{V_1S}\sigma_{V_1}\sigma_S + 2\alpha_2\alpha_3\rho_{V_2S}\sigma_{V_2}\sigma_S)$ is a gain factor related to power constraint P. In such case, Y_i is Gaussian and the MMSE decoders for reconstructing V_i and S are linear estimators. The reconstructed signals at user i $(i = 1, 2)$ are then found as follows

$$\hat{V}_i = \frac{\mathbb{E}[V_i Y_i]}{\mathbb{E}[Y_i^2]} Y_i, \quad \hat{S}_i = \frac{\mathbb{E}[S Y_i]}{\mathbb{E}[Y_i^2]} Y_i. \tag{25}$$

The resulting distortions (defined in (4)) from reconstructing the source and the interference at user i can then be expressed as follows

$$\begin{aligned} D_{v_i}^{(Linear)} &= \sigma_{V_i}^2 - \frac{\mathbb{E}[V_i Y_i]^2}{\mathbb{E}[Y_i^2]} \\ D_{s_i}^{(Linear)} &= \sigma_S^2 - \frac{\mathbb{E}[S Y_i]^2}{\mathbb{E}[Y_i^2]}. \end{aligned} \tag{26}$$

Note that the parameters $(\alpha_1, \alpha_2, \alpha_3)$ are found by minimizing the overall distortion $D^{(Linear)} = \sum_{i=1}^{2}(\theta_{v_i} D_{v_i}^{(Linear)} + \theta_{s_i} D_{s_i}^{(Linear)})$.

Remark 5 *The optimal tradeoff between distortion quadruple $(D_{v_1}, D_{v_2}, D_{s_1}, D_{s_2})$ for any linear scheme is achieved with single-letter linear codes (i.e., in a scalar form). Hence, there is no gain from using a higher dimensional linear transformation.*

Remark 6 *For the case of $\rho_{V_1V_2} = \rho_{V_1S} = \rho_{V_2S} = 1$ or when the users are only interested in the interference, the linear scheme (in this case, a scaled version of the interference $X^K = \sqrt{P/\sigma_S^2} S^K$) is optimal.*

B. Parametric Mapping

In this section, we propose a layered scheme based on linear coding and sawtooth mapping. Sawtooth mapping has been used for the relay channels in [43], the multiple access channels [21], the Gaussian broadcast channels [44] and for the dirty paper coding problem [24], [25].

1) System Structure: The proposed scheme is composed of two superposed layers and outputs

$$X = c(X_1 + X_2) \tag{27}$$

where c is a gain factor related to the power constraint (defined later). The first layer, which outputs $X_1 = \sqrt{P_s/\sigma_S^2} S$, simply scales the interference S, where $P_s \leq P$ represents the power consumed by this layer. The second layer, starts by forming a linear combination of the sources (V_1, V_2) and the interference S; this is given by $X_a = \alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 S$, where α_1, α_2 and α_3 are real parameters. We then use a sawtooth mapping $\mathcal{S}(\cdot)$ on X_a to output X_2 as follows

$$X_2 = \mathcal{S}(X_a) = (X_a - 2\Lambda m) \text{ for } X_a \in \big[\Lambda(2m-1), \Lambda(2m+1)\big) \tag{28}$$

where m is an integer and Λ is a nonnegative parameter dependent on the channel condition. The gain factor c in (27) is given by $c = \sqrt{P/(P_s + \mathbb{E}[X_2^2] + 2\sqrt{P_s/\sigma_S^2}\mathbb{E}[SX_2])}$, where $\mathbb{E}[X_2^2]$ can be written as follows

$$\mathbb{E}[X_a^2] + \sum_m -4\Lambda m \underbrace{\int_{D_m} x_a p(x_a) dx_a}_{I_1} + 4\Lambda^2 m^2 \underbrace{\int_{D_m} p(x_a) dx_a}_{I_2} \tag{29}$$

and $\mathbb{E}[SX_2]$ is given by

$$\mathbb{E}[SX_a] - \sum_m 2\Lambda m \int\int_{D_m} s p(x_a|s) p(s) dx_a ds \tag{30}$$

where $D_m = \big[\Lambda(2m-1), \Lambda(2m+1)\big)$ is the m^{th} domain region of $\mathcal{S}(\cdot)$, $p(\cdot)$ denotes a probability density function (pdf) and $p(\cdot|\cdot)$ is a conditional pdf. Note that the integrals in (29) can be simplified as follows

$$\begin{aligned} I_1 &= \frac{\sqrt{\mathbb{E}[X_a^2]}}{\sqrt{2\pi}} \Bigg[-\text{exp}\left(\frac{-(\Lambda(2m+1))^2}{2\mathbb{E}[X_a^2]}\right) \\ &\quad + \text{exp}\left(\frac{-(\Lambda(2m-1))^2}{2\mathbb{E}[X_a^2]}\right) \Bigg], \\ I_2 &= \frac{1}{2}\left[\text{erf}\left(\frac{\Lambda(2m+1)}{\sqrt{2\mathbb{E}[X_a^2]}}\right) - \text{erf}\left(\frac{\Lambda(2m-1)}{\sqrt{2\mathbb{E}[X_a^2]}}\right) \right] \end{aligned} \tag{31}$$

where erf$(\cdot)$ is the Gaussian error function. At the decoder side, to obtain an estimate of the source and the interference at each user i, we use the optimal MMSE estimator ($\hat{V}_i = \mathbb{E}[V_i|Y_i], \hat{S}_i = \mathbb{E}[S|Y_i]$). The use of an optimal decoder comes at the expense of computational and design complexity. To lower the design complexity, we resort to a suboptimal method for choosing the system parameters as described next.

2) System Optimization: We herein optimize the system parameters by minimizing an upper bound on the system's distortion. To get a closed form expression on the system's distortion upper bound, we propose the use of a suboptimal decoder. Let us first note that the sawtooth mapping, which uses the symmetric modulo function (28) over the interval $[-\Lambda\ \Lambda]$, can be written as $\mathcal{S}(X_a) = X_a \ mod\ \Lambda$. To reconstruct the interference, we simply use an LMMSE estimator based on the received signal Y_i. The distortion from reconstructing

S at each user is given by

$$D_{s_i}^{(Parametric)} = \sigma_S^2 - \frac{\mathbb{E}[SY_i]^2}{\mathbb{E}[Y_i^2]}$$
$$= \sigma_S^2 - \frac{(c(\mathbb{E}[S(X_a \ mod \ \Lambda)] + \sqrt{P_s}\sigma_S) + \sigma_S^2)^2}{P + \sigma_S^2 + \sigma_{W_i}^2 + 2c(\sqrt{P_s}\sigma_S + \mathbb{E}[S(X_a \ mod \ \Lambda)])}$$

where $\mathbb{E}[S(X_a \ mod \ \Lambda)]$ can be written as in (30).

To get an estimate of the source V_i, we first use a modulo function on the received signal Y_i and then apply an LMMSE estimator. More precisely, we first obtain

$$\tilde{Y}_i = (Y_i/c) \ mod \ \Lambda$$
$$= \underbrace{\left(\alpha_1 V_1 + \alpha_2 V_2 + (\alpha_3 + \sqrt{\frac{P_s}{\sigma_S^2}} + \frac{1}{c})S + \frac{W_i}{c}\right)}_{Z_i} mod \ \Lambda \tag{32}$$

where the last equality follows from the fact that the modulo operation satisfies the "distributive law" (i.e., $[x \ mod \ \Lambda + y] \ mod \ \Lambda = [x + y] \ mod \ \Lambda$). We then decode V_i using an LMMSE estimator based on $\tilde{Y}_i$. The resulting distortion is

$$D_{v_i}^{(Parametric)} = \sigma_{V_i}^2 - \frac{\mathbb{E}[V_i(Z_i \ mod \ \Lambda)]^2}{\mathbb{E}[(Z_i \ mod \ \Lambda)^2]} \tag{33}$$

where Z_i is defined in (32), $\mathbb{E}[(Z_i \ mod \ \Lambda)^2]$ is given by

$$\mathbb{E}[Z_i^2] - 4\Lambda \sum_m m \int_{\tilde{D}_m} z_i p(z_i) dz_i + 4\Lambda^2 \sum_m m^2 \int_{D_m} p(z_i) dz_i \tag{34}$$

and $\mathbb{E}[V_i(Z_i \ mod \ \Lambda)]$ can be expressed as follows

$$\mathbb{E}[V_i Z_i] - \sum_m 2\Lambda m \int_{\tilde{D}_m} \iint v_i p(v_i|s) p(z_i|v_i, s) p(s) dv_i ds dz_i \tag{35}$$

where $\tilde{D}_m = \big[\Lambda(2m-1), \Lambda(2m+1)\big)$ is the m^{th} domain region of $\mathcal{S}(\cdot)$. Note that integrals in (34) can be simplified in a similar way as in (31) and distributions in (35) are Gaussian. The upper bound on the system's distortion D_{upper} for parametric mapping is then given by

$$D_{upper} = \sum_{i=1}^{2} \left[\theta_{v_i} D_{v_i}^{(Parametric)} + \theta_{s_i} D_{s_i}^{(Parametric)}\right]. \tag{36}$$

Note that $(\alpha_1, \alpha_2, \alpha_3, P_s, \Lambda)$ are found by minimizing D_{upper}, θ_{v_i} and θ_{s_i} are set by the system designer.

C. Nonparametric Mappings

We next present a scheme based on joint optimization between the encoder and the decoders through an iterative algorithm. Closed form expressions for $\alpha(\cdot)$, $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ that minimize the distortion D (given in (23)) may not exist; this makes the optimization of the encoder and decoders difficult. The rest of this section is dedicated to the design of the source-channel mapping $\alpha(V_1, V_2, S)$ and the decoders $\gamma_i(Y_i)$, for $i = 1, \ 2$. Using the Lagrange multiplier method, the constrained minimization of the MSE distortion D subject to the power constraint in (3) can be recast into an unconstrained minimization via the Lagrange cost function $J(\alpha, \gamma_1, \gamma_2)$

$$J = \sum_{i=1}^{2} \theta_{v_i} \mathbb{E}\left[\left(V_i - \gamma_i^{(v)}(Y_i)\right)^2\right] + \theta_{s_i} \mathbb{E}\left[\left(S - \gamma_i^{(s)}(Y_i)\right)^2\right] + \lambda \mathbb{E}[\alpha(V_1, V_2, S)^2] \tag{37}$$

where λ is used to control the average power. For a given λ, if the solution of the unconstrained minimization fulfills the average power constraint in (3), the obtained solution is also proven to solve the constrained problem [45]. The above unconstrained minimization is still hard to solve due to interdependencies between the optimized components. To solve this, we proceed in a way similar to classical design problems [46] by deriving necessary conditions for optimality. This is done by determining the optimal encoder α given the decoders (γ_1, γ_2), and vice versa.

1) Necessary Conditions for Optimality: The optimal encoder mapping α^* (assuming (γ_1, γ_2) are fixed) is given by

$$\alpha^* = \arg\min_{\alpha} \Big\{ \sum_{i=1}^{2} \theta_{v_i} \mathbb{E}[\big(V_i - \hat{V}_i\big)^2] + \theta_{s_i} \mathbb{E}[\big(S - \hat{S}_i\big)^2] + \lambda \mathbb{E}[\alpha(V_1, V_2, S)^2] \Big\}. \tag{38}$$

Using Bayes' rule, the distortion $\mathbb{E}[\big(V_i - \hat{V}_i\big)^2]$ is given by

$$\iiiint p(v_1, v_2, s) p(y_i|\alpha(v_1, v_2, s), s) (v_i - \hat{v}_i)^2 dv_1 dv_2 ds dy_i. \tag{39}$$

Similarly, the distortion $\mathbb{E}[\big(S - \hat{S}_i\big)^2]$ can be expressed as follows

$$\iiiint p(v_1, v_2, s) p(y_i|\alpha(v_1, v_2, s), s) (s - \hat{s}_i)^2 dv_1 dv_2 ds dy_i. \tag{40}$$

The average consumed power is given by

$$P = \iiint p(v_1, v_2, s) \alpha(v_1, v_2, s)^2 dv_1 dv_2 ds. \tag{41}$$

Since $p(v_1, v_2, s)$ in (39), (40) and (41) is nonnegative, the encoder α can be optimized "pointwise" for each v_1, v_2 and s according to

$$\alpha^*(v_1, v_2, s) = \arg\min_{x \in \mathbb{R}} \Big\{ \sum_{i=1}^{2} \int p(y_i|x, s) [\theta_{v_i}(v_i - \hat{v}_i)^2 + \theta_{s_i}(s - \hat{s}_i)^2] dy_i + \lambda x^2 \Big\}. \tag{42}$$

Thus, (42) is a necessary condition for an optimal encoder.

On the receiver side, the optimal decoder $\gamma_i^{(v)}$ in the MSE sense (assuming α is fixed) at user i is given by $\mathbb{E}[V_i|y_i]$ as follows

$$\gamma_i^{(v)*}(y_i) = \frac{\iiint v_i p(y_i|v_1, v_2, s) p(v_1, v_2, s) dv_1 dv_2 ds}{\iiint p(y_i|v_1, v_2, s) p(v_1, v_2, s) dv_1 dv_2 ds}. \tag{43}$$

Similarly, the optimal decoder $\gamma_i^{(s)}$ in the MSE sense (assuming α is fixed) at user i is given by $\mathbb{E}[S|y_i]$ as follows

$$\gamma_i^{(s)*}(y_i) = \frac{\iiint s p(y_i|v_1, v_2, s) p(v_1, v_2, s) dv_1 dv_2 ds}{\iiint p(y_i|v_1, v_2, s) p(v_1, v_2, s) dv_1 dv_2 ds}. \tag{44}$$

2) Design Algorithm: Using the above necessary conditions for optimality, we optimize α, $\gamma_1 = (\gamma_1^{(v)}, \gamma_1^{(s)})$ and $\gamma_2 = (\gamma_2^{(v)}, \gamma_2^{(s)})$ via an iterative process based on (42), (43) and (44). The update equations (42), (43) and (44) yield a lower distortion at each iteration step; Thus, with a finite amount of training data, convergence is guaranteed. The main problem with such iterative process is that the final solution depends on the choice of the initial mapping in the algorithm and convergence to the global optimum is not ensured. The design Algorithm 1 is as follows:

1) Choose some initial mapping for the encoder α.
2) Find the optimal decoder (γ_1, γ_2) according to (43) and (44).
3) Set the iteration index $i = 0$ and the cost $J^{(0)} = \infty$.
4) Set $i = i + 1$.
5) Find the optimal mapping α according to (42).
6) Find the optimal decoder (γ_1, γ_2) according to (43) and (44).
7) Evaluate the cost function $J^{(i)}$. If the relative improvement of $\frac{J^{(i-1)} - J^{(i)}}{J^{(i-1)}} < \epsilon$ or $i > I_{max}$, stop iterating. Else go to step 4.

Algorithm 1 is nested inside a "bracketing" Lagrange multiplier search. We first set $\lambda = \lambda_0$. If the designed α produces $\mathbb{E}[\alpha(V_1, V_2, S)^2] > P$, λ_0 is increased; else λ_0 is decreased. The search ends if $\mathbb{E}[\alpha(V_1, V_2, S)^2]$ is close enough to but less than P. For initialization of the algorithm, we use the parametric mapping scheme proposed in the previous subsection.

3) Implementation Aspects: For the implementation of (42), (43) and (44), some modifications are required. Since it is intractable to evaluate the formulas for all real-valued (V_1, V_2, S), we form as in [47] a set of triplets $(\mathcal{V}_1, \mathcal{V}_2, \mathcal{S})$ composed of samples drawn from $p(v_1, v_2, s)$. Since the channel input and output spaces are real valued, we discretize them using a pulse amplitude modulation alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively, in each direction. We use

$$\mathcal{X} = \mathcal{Y} = \left\{ -d\frac{L-1}{2}, -d\frac{L-3}{2}, ..., d\frac{L-3}{2}, d\frac{L-1}{2} \right\} \tag{45}$$

where d and L are the resolution and the cardinality of the set, respectively. In our simulations, we use 10^5 triplets $(\mathcal{V}_1, \mathcal{V}_2, \mathcal{S})$, $\epsilon = 10^{-3}$, $I_{max} = 15$, $L = 700$ and $d = 12/(L-1)$. The discretized version of (42) which is used in the implementation of the design algorithm is expressed as follows

$$\alpha^*(v_1, v_2, s) = \arg\min_{x \in \mathcal{X}} \left\{ \sum_{i=1}^{2} \sum_{y_i \in \mathcal{Y}} P(y_i | x, s) \right. \\ \left. [\theta_{v_i}(v_i - \hat{v}_i)^2 + \theta_{s_i}(s - \hat{s}_i)^2] + \lambda x^2 \right\}. \tag{46}$$

where $P(\cdot)$ denotes the probability mass function. Note that the discretized versions of (43) and (44) can be written similarly.

VI. Numerical Results

In this section we assume that the source pairs with variance $\sigma_{V_1}^2 = \sigma_{V_2}^2 = 1$ are broadcasted to two users that are disturbed with common interference with variance $\sigma_S^2 = 1$. The system's average power is set to $P = 1$.

A. Special Cases

Since the quadruple distortion region is difficult to visualize, we next consider two special cases: 1) the transmission of bivariate source over a Gaussian degraded broadcast channel in the presence of interference, referred to as "Case 1" and 2) the joint transmission of Gaussian source and interference over a point-to-point Gaussian channel, referred to as "Case 2". To evaluate the performance of these two scenarios, we plot the inner bound (achievable distortion region of the proposed hybrid scheme as given by (18) and (22)) and the outer bound (derived in Theorem 1).

1) Case 1: Fig. 6 shows the distortion region (D_{v_1}, D_{v_2}) when users are only interested in estimating the source pair and not the interference. We can notice that the inner bound achieved using the HDA scheme is relatively close to the 'best' outer bound; moreover, the derived outer bound of Theorem 1 improves on our previous bound derived in [1]. Part of this improvement is related to restricting the search space over which the supremum is applied (refer to Theorem 1); the remaining gain is from assuming a partial (noisy) knowledge of the source at the strong user instead of full knowledge when deriving the bound. From Fig. 6, we can see that the gap between the inner and the outer bounds decreases for high distortion level D_{v_1}; this is because for high distortion level D_{v_1}, the system behave similar to a point-to-point communication.

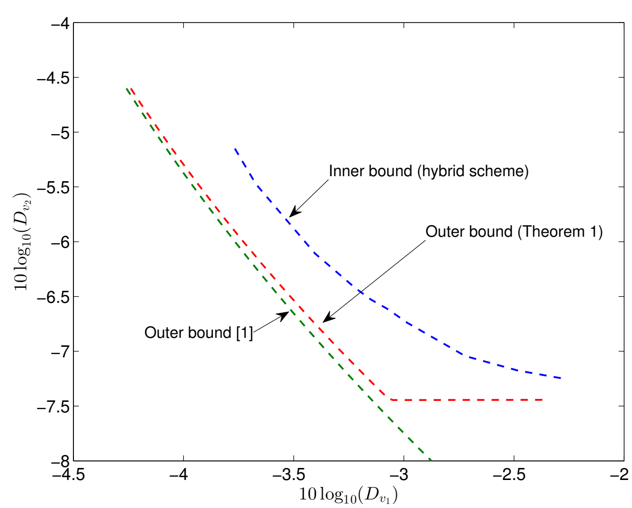

Fig. 6. Case 1: Distortion regions (D_{v_1}, D_{v_2}) for hybrid scheme for $\rho_{V_1V_2} = 0.8, \rho_{V_1S} = \rho_{V_2S} = 0.5, \sigma_{W_1}^2 = 0$ dB, $\sigma_{W_2}^2 = -5$ dB and $P = 1$. The inner bound is plotted using (18) and (22).

2) Case 2: Fig. 7 focuses on the joint source-interference recovery over a single-user channel. We can notice that the derived outer bound is tighter than the one derived previously in [2]. Moreover, the gap between the inner and outer bounds decreases for high distortion levels on D_{s_2}; in such case, our system behave as if we are only interested in estimating the source V_2. Note that we were only able to notice some gap

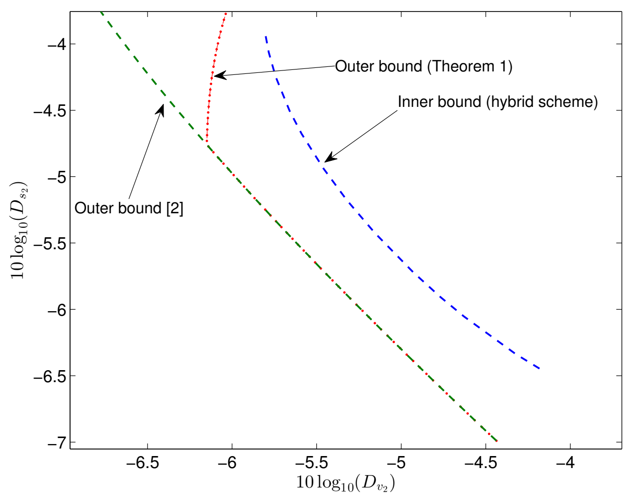

Fig. 7. Case 2: Distortion regions (D_{v_2}, D_{s_2}) for hybrid scheme for $\rho_{V_1V_2} = 1, \rho_{V_1S} = \rho_{V_2S} = 0.8, \sigma^2_{W_1} = \sigma^2_{W_2} = 0$ dB and $P = 1$. The inner bound is plotted using (22).

between the inner and the outer bounds for high noise levels ($\sigma^2_{W_1} \geq 0$ dB). For low noise levels ($\sigma^2_{W_1} < 0$ dB), the inner and the outer bounds overlap for most of the region; this is illustrated in Fig. 8.

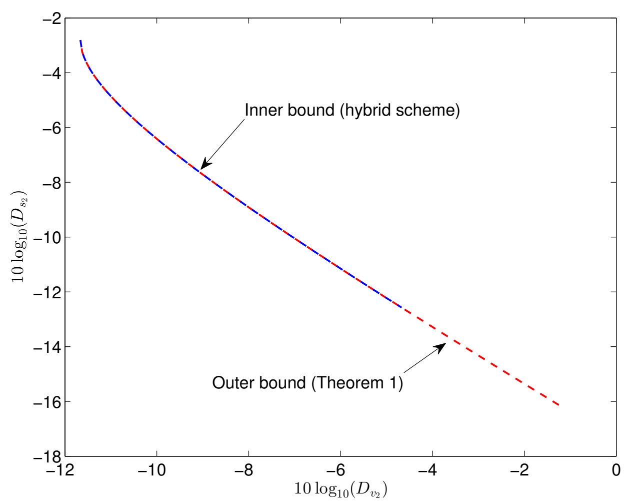

Fig. 8. Case 2: Distortion regions (D_{v_2}, D_{s_2}) for hybrid scheme for $\rho_{V_1V_2} = 1, \rho_{V_1S} = \rho_{V_2S} = 0.5, \sigma^2_{W_1} = \sigma^2_{W_2} = -10$ dB and $P = 1$. The inner bound is plotted using (22).

B. General Case

Figs. 9 and 10 show the performance of the proposed low-delay schemes relative to the theoretical bounds as a function of correlation values. we assume that the weights in the distortion measure D in (23) are set to $1/4$ (i.e., $\theta_{v_i} = \theta_{s_i} = 1/4$). As we can notice, the nonparametric and parametric mappings outperform the linear scheme. It is worth mentioning that the information-theoretic bounds are asymptotic in the sense of requiring infinite source and coding block lengths, hence the gap to the proposed low delay scheme is not surprising. Note that the performance of the parametric mapping is not shown in Fig. 10; this is due to the fact that for high noise levels, our parametric mapping behave similar to the linear one which usually tends to have a good performance relative to other low delay schemes. Moreover, the gap between the inner and the outer bounds is relatively small (~ 0.8 dB).

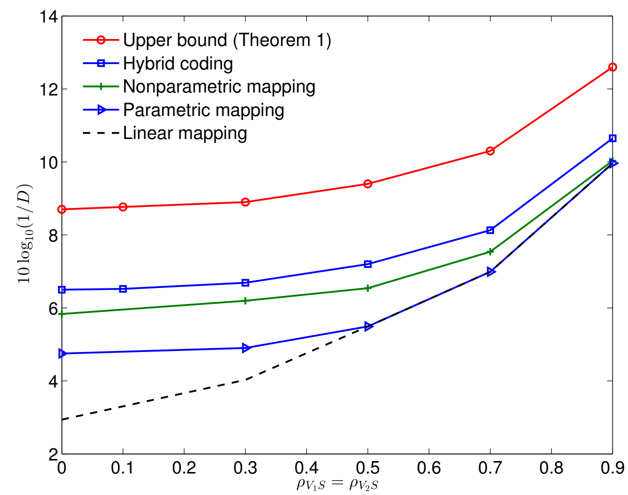

Fig. 9. General case: Performance of low delay coding versus ρ_{V_1S} for $\rho_{V_1V_2} = 0.7, \sigma^2_{W_1} = -10$ dB, $\sigma^2_{W_2} = -15$ dB and $P = 1$. The performance of hybrid coding is based on (18) and (22).

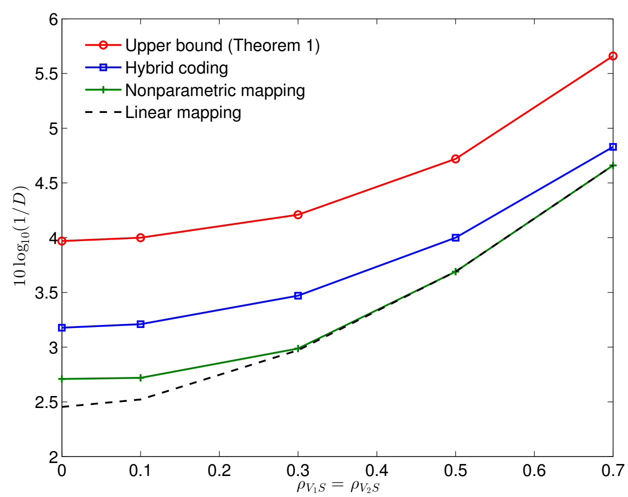

Fig. 10. General case: Performance of low delay coding versus ρ_{V_1S} for $\rho_{V_1V_2} = 0.2, \sigma^2_{W_1} = 0$ dB, $\sigma^2_{W_2} = -5$ dB and $P = 1$. The performance of hybrid coding is based on (18) and (22).

Fig. 11 shows the decoder structure of the nonparametric mapping for two different noise levels. Fig. 12 shows the encoder-decoder structure for the special case (Case 2) of joint transmission of Gaussian source and interference over point-to-point Gaussian channels (i.e., $\theta_{v_1} = \theta_{s_1} = 0$ and $\theta_{v_2} =$

$\theta_{s_2} = 1/2$). It is clear that the encoder and decoder mappings comprise a piecewise nonlinear function that combines hard and soft decision signalling. The proposed parametric mapping uses such combination; this explains the good performance achieved using parametric mapping. There is always a gain from using the nonparametric mapping; this is due to the fact that the nonparametric mapping has a higher degree of freedom in placing points in space without being restrained to a specific structure. Such gain comes at the expense of higher storage and offline design complexity.

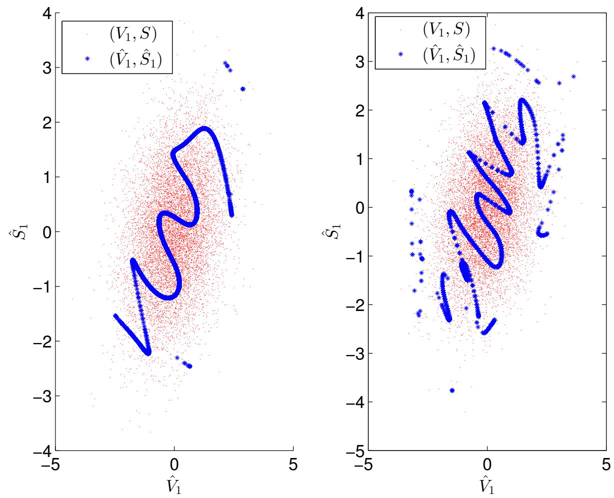

Fig. 11. General case: Decoder mappings structure optimized using Algorithm 1 for $\rho_{V_1V_2} = 0.7$, $\rho_{V_1S} = \rho_{V_2S} = 0.3$, $\sigma_{V_1}^2 = \sigma_{V_2}^2 = \sigma_S^2 = 1$ and $P = 1$. In the figure to the left we assume $(\sigma_{W_1}^2 = -10, \sigma_{W_2}^2 = -15)$ dB while the one to the right has $(\sigma_{W_1}^2 = -20, \sigma_{W_2}^2 = -25)$ dB. Note that the asterisks show the reconstructed $(\hat{V}_1, \hat{S}_1)$ and the small dots are samples from the distribution of (V_1, S).

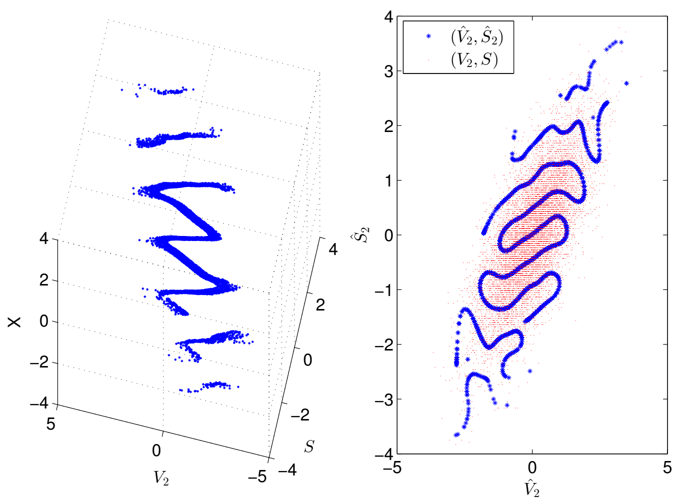

Fig. 12. Case 2: Encoder (left) and its corresponding decoder (right) mappings optimized using Algorithm 1 for $\sigma_{W_2}^2 = -25$ dB, $\rho_{V_2S} = 0.7$, $\sigma_{V_2} = \sigma_S = 1$ and $P = 1$; parametric mapping is used for the initialization of Algorithm 1. In the figure to the right, the asterisks show the reconstructed $(\hat{V}_2, \hat{S}_2)$ and the small dots are samples from the distribution of (V_2, S).

VII. Conclusions

In this paper, we considered the problem of reliable transmission of bivariate Gaussian source and interference recovery over the two-user Gaussian degraded channel in the presence of interference that is known non-causally to the transmitter and correlated to the sources. Information-theoretic outer and inner bounds using ideas from rate-distortion theory and hybrid digital-analog coding are derived. Low delay and low complexity codes based on analog transmission are then proposed. More precisely, parametric mappings based on linear and sawtooth curves are studied; nonparametric mappings based on joint optimization between the encoder and the decoder are designed using an iterative algorithm. Our setting contains several interesting limiting cases; the derived bounds resort to the optimal ones in [6], [31] and are tighter than the ones derived in [1] and [2].

References

[1] A. Abou Saleh, F. Alajaji, and W.-Y. Chan, "Hybrid digital-analog coding for interference broadcast channels," in *Proc. IEEE Int. Symp. on Inform. Theory*, Istanbul, Turkey, July 2013.

[2] A. Abou Saleh, W.-Y. Chan, and F. Alajaji, "Source-channel coding for fading channels with correlated interference," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3997–4011, Nov 2014.

[3] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Inform. Contr.*, vol. 9, no. 1, pp. 19–31, 1980.

[4] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[5] Y.-C. Huang and K. R. Narayanan, "Joint source-channel coding with correlated interference," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1315–1327, May 2012.

[6] A. Sutivong, M. Chiang, T. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1486–1495, Apr 2005.

[7] B. Bandemer, C. Tian, and S. Shamai, "Gaussian state amplification with noisy state observations," in *Proc. IEEE Int. Symp. on Inform. Theory*, Istanbul, Turkey, July 2013.

[8] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[9] U. Mittal and N. Phamdo, "Hybrid digital analog (HDA) joint source channel codes for broadcasting and robust communications," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1082–1102, May 2002.

[10] M. Skoglund, N. Phamdo, and F. Alajaji, "Hybrid digital-analog source-channel coding for bandwidth compression/expansion," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3757–3763, Aug 2006.

[11] Y. Wang, F. Alajaji, and T. Linder, "Hybrid digital-analog coding with bandwidth compression for Gaussian source-channel pairs," *IEEE Trans. Communications*, vol. 57, no. 4, pp. 997–1012, Apr 2009.

[12] J. Nayak, E. Tuncel, and D. Gunduz, "Wyner-Ziv coding over broadcast channels: digital schemes," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1782–1799, Apr 2010.

[13] Y. Gao and E. Tuncel, "Wyner-Ziv coding over broadcast channels: Hybrid digital/analog schemes," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp. 5660–5672, Sept 2011.

[14] E. Koken and E. Tuncel, "Gaussian HDA coding with bandwidth expansion and side information at the decoder," in *Proc. IEEE Int. Symp. on Inform. Theory*, Istanbul, Turkey, July 2013.

[15] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 86, no. 2, pp. 447–457, 1949.

[16] V. A. Kotel'nikov, *The Theory of Optimum Noise Immunity*. New York: McGraw-Hill, 1959.

[17] T. A. Ramstad, "Shannon mappings for robust communication," *Telektronikk*, vol. 98, no. 1, pp. 114–128, 2002.

[18] F. Hekland, P. A. Floor, and T. A. Ramstad, "Shannon-Kotel'nikov mappings in joint source-channel coding," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 94–105, Jan 2009.

[19] Y. Hu, J. Garcia-Frias, and M. Lamarca, "Analog joint source channel coding using non-linear mappings and MMSE decoding," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3016–3026, Nov 2011.

[20] E. Akyol, K. Rose, K. Viswanatha, and T. Ramstad, “On zero-delay source-channel coding,” *IEEE Trans. Inform. Theory*, vol. 60, no. 12, pp. 7473–7489, Dec 2014.

[21] P. A. Floor, A. N. Kim, N. Wernersson, T. Ramstad, M. Skoglund, and I. Balasingham, “Zero-delay joint source-channel coding for a bivariate Gaussian on a Gaussian MAC,” *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 3091–3102, Oct 2012.

[22] T. Cui, T. Ho, and J. Kliewer, “Memoryless relay strategies for two-way relay channels,” *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3132–3143, Oct 2009.

[23] A. Abou Saleh, F. Alajaji, and W.-Y. Chan, “Low-latency source-channel coding for fading channels with correlated interference,” *IEEE Wireless Commun. Letter*, vol. 3, no. 2, pp. 137–140, April 2014.

[24] ——, “Analog coding for Gaussian source and state interference estimation,” in *Proc. IEEE Int. Workshop on Signal Processing Advances in Wireless Communications*, Stockholm, Sweden, July 2015.

[25] M. Varasteh, D. Gunduz, and E. Tuncel, “Zero-delay joint source-channel coding in the presence of interference known at the encoder,” in *Proc. IEEE Int. Conf. on Communications*, London, UK, July 2015.

[26] M. Gastpar, B. Rimoldi, and M. Vetterli, “To code, or not to code: lossy source-channel communication revisited,” *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[27] V. M. Prabhakaran, R. Puri, and K. Ramachandran, “Hybrid analog-digital strategies for source-channel broadcast,” in *Proc. 43rd Allerton Conf. Communication, Control and Computing*, Allerton, IL, Sep 2005.

[28] Z. Reznic, M. Feder, and R. Zamir, “Distortion bounds for broadcasting with bandwidth expansion,” *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3778–3788, Aug 2006.

[29] C. Tian, S. N. Diggavi, and S. Shamai, “Approximate characterizations for the Gaussian broadcasting distortion region,” in *Proc. IEEE Int. Symp. on Inform. Theory*, Seoul, Korea, July 2009.

[30] S. Bross, A. Lapidoth, and S. Tinguely, “Broadcasting correlated Gaussians,” *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3057–3068, July 2010.

[31] C. Tian, S. N. Diggavi, and S. Shamai, “The achievable distortion region of sending a bivariate Gaussian source on the Gaussian broadcast channel,” *IEEE Trans. Inform. Theory*, vol. 57, no. 10, pp. 6419–6427, Oct 2011.

[32] H. Behroozi, F. Alajaji, and T. Linder, “On the performance of hybrid digital-analog coding for broadcasting correlated Gaussian sources,” *IEEE Trans. Communications*, vol. 59, no. 12, pp. 3335–3342, Dec 2011.

[33] Y. Gao and E. Tuncel, “Separate source-channel coding for transmitting correlated Gaussian sources over degraded broadcast channels,” *IEEE Trans. Inform. Theory*, vol. 59, no. 6, pp. 3619–3634, June 2013.

[34] L. Song, J. Chen, and C. Tian, “Broadcasting correlated vector Gaussians,” *IEEE Trans. Inform. Theory*, vol. 61, no. 5, pp. 2465–2477, May 2015.

[35] W. Liu and B. Chen, “Message transmission and state estimation over Gaussian broadcast channels,” in *Proc. 43rd Annual Conference on Information Sciences and Systems*, Baltimore, MD, March 2009.

[36] Y. Zhao and B. Chen, “Capacity theorems for multi-functioning radios,” in *Proc. IEEE Int. Symp. on Inform. Theory*, Honolulu, HI, July 2014.

[37] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2012.

[38] A. D. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. Inform. Theory*, vol. 22, pp. 1–10, Jan 1976.

[39] R. Soundararajan and S. Vishwanath, “Hybrid coding for Gaussian broadcast channels with Gaussian sources,” in *Proc. IEEE Int. Symp. on Inform. Theory*, Seoul, Korea, July 2009.

[40] Y.-K. Chia, R. Soundararajan, and T. Weissman, “Estimation with a helper who knows the interference,” in *Proc. IEEE Int. Symp. on Inform. Theory*, Cambridge, MA, July 2012.

[41] A. Lapidoth and S. Tinguely, “Sending a bivariate Gaussian over a Gaussian MAC,” *IEEE Trans. Information Theory*, vol. 56, no. 6, pp. 2714–2752, 2010.

[42] M. P. Wilson, K. R. Narayanan, and G. Caire, “Joint source-channel coding with side information using hybrid digital analog codes,” *IEEE Trans. Inform. Theory*, vol. 56, no. 10, pp. 4922–4940, Oct 2010.

[43] S. Yao, M. Khormuji, and M. Skoglund, “Sawtooth relaying,” *IEEE Commun. Letters*, vol. 12, no. 9, pp. 612–614, Sep 2008.

[44] M. Hassanin, J. Garcia-Frias, and L. Castedo, “Analog joint source channel coding for Gaussian broadcast channels,” in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Florence, Italy, May 2014.

[45] H. Everett III, “Generalized Lagrange multiplier method for solving problems of optimum allocation of resources,” *Operations Research*, vol. 11, no. 3, pp. 399–417, 1963.

[46] N. Farvardin and V. Vaishampayan, “On the performance and complexity of channel-optimized vector quantizers,” *IEEE Trans. Inform. Theory*, vol. 37, pp. 155–159, Jan 1991.

[47] J. (Karlsson) Kron and M. Skoglund, “Optimized low-delay source-channel-relay mappings,” *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1397–1404, May 2010.