Source-Channel Coding for Fading Channels with Correlated Interference

Ahmad Abou Saleh, Wai-Yip Chan, and Fady Alajaji

Abstract

We consider the problem of sending a Gaussian source over a fading channel with Gaussian interference known to the transmitter. We study joint source-channel coding schemes for the case of unequal bandwidth between the source and the channel and when the source and the interference are correlated. An outer bound on the system's distortion is first derived by assuming additional information at the decoder side. We then propose layered coding schemes based on proper combination of power splitting, bandwidth splitting, Wyner-Ziv and hybrid coding. More precisely, a hybrid layer, that uses the source and the interference, is concatenated (superimposed) with a purely digital layer to achieve bandwidth expansion (reduction). The achievable (square error) distortion region of these schemes under matched and mismatched noise levels is then analyzed. Numerical results show that the proposed schemes perform close to the best derived bound and to be resilient to channel noise mismatch. As an application of the proposed schemes, we derive both inner and outer bounds on the source-channel-state distortion region for the fading channel with correlated interference; the receiver, in this case, aims to jointly estimate both the source signal as well as the channel-state (interference).

Index Terms

Joint source-channel coding, distortion region, correlated interference, dirty paper coding, hybrid digital-analog coding, fading channels.

This work was supported in part by NSERC of Canada.

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).

I. Introduction

The traditional approach for analog source transmission in point-to-point communications systems is to employ separate source and channel coders. This separation is (asymptotically) optimal given unlimited delay and complexity in the coders [1]. There are, however, two disadvantages associated with digital transmission. One is the threshold effect: the system typically performs well at the design noise level, while its performance degrades drastically when the true noise level is higher than the design level. This effect is due to the quantizers sensitivity to channel errors and the eventual breakdown of the employed error correcting code at high noise levels (no matter how powerful it is). The other trait is the levelling-off effect: as the noise level decreases, the performance remains constant beyond a certain threshold. This is due to the non-recoverable distortion introduced by the quantizer which limits the system performance at low noise levels. Joint source-channel coding (JSCC) schemes are more robust to noise level mismatch than tandem systems which use separate source and channel coding. Analog JSCC schemes are studied in [2]–[11]. These schemes are based on the so-called direct source-channel mappings. A family of hybrid digital-analog (HDA) schemes are introduced in [12]–[14] to overcome the threshold and the levelling-off effects. In [15]–[17], HDA schemes are proposed for broadcast channels and Wyner-Ziv systems.

It is well known that for the problem of transmitting a Gaussian source over an additive white Gaussian noise (AWGN) channel with interference that is known to the transmitter, a tandem Costa coding scheme, which comprises an optimal source encoder followed by Costa's dirty paper channel code [18], is optimal in the absence of correlation between the source and the interference. In [19], the authors studied the same problem as in [18] and proposed an HDA scheme (for the matched bandwidth case) that is able to achieve the optimal performance (same as the tandem Costa scheme). In [20], the authors adapted the scheme proposed in [19] for the bandwidth reduction case. In [21], the authors proposed an HDA scheme for broadcasting correlated sources and showed that their scheme is optimal whenever the uncoded scheme of [22] is not. In [23], the authors studied HDA schemes for broadcasting correlated sources under mismatched source-channel bandwidth; in [24], the authors studied the same problem and proposed a tandem scheme based on successive coding. In [25], we derived inner and outer bounds on the system's distortion for the broadcast channel with correlated interference. Recently,

[26] studied a joint source channel coding scheme for transmitting analog Gaussian source over AWGN channel with interference known to the transmitter and correlated with the source. The authors proposed two schemes for the matched source-channel bandwidth; the first one is the superposition of the uncoded signal and a digital signal resulting from the concatenation of a Wyner-Ziv coder [27] and a Costa coder, while in the second scheme the digital part is replaced by an HDA part proposed in [19]. In [28], we consider the problem of [26] under bandwidth expansion; more precisely, we studied both low and high-delay JSCC schemes. The limiting case of this problem, where the source and the interference are fully correlated was studied in [29]; the authors showed that a purely analog scheme (uncoded) is optimal. Moreover, they also considered the problem of sending a digital (finite alphabet) source in the presence of interference where the interference is independent from the source. More precisely, the optimal tradeoff between the achievable rate for transmitting the digital source and the distortion in estimating the interference is studied; they showed that the optimal rate-state-distortion tradeoff is achieved by a coding scheme that uses a portion of the power to amplify the interference and uses the remaining power to transmit the digital source via Costa coding. In [30], the authors considered the same problem as in [29] but with imperfect knowledge of the interference at the transmitter side.

In this work, we study the reliable transmission of a memoryless Gaussian source over a Rayleigh fading channel with known correlated interference at the transmitter. More precisely, we consider equal and unequal source-channel bandwidths and analyze the achievable distortion region under matched and mismatched noise levels. We propose a layered scheme based on hybrid coding. One application of JSCC with correlated interference can be found in sensor network and cognitive radio channels where two nodes interfere with each other. One node transmits directly its signal; the other, however, is able to detect its neighbour node transmission and treat it as a correlated interference. In [31], we studied this problem under low-delay constraints; more specifically, we designed low-delay source-channel mappings based on joint optimization between the encoder and the decoder. One interesting application of this problem is to study the source-channel-state distortion region for the fading channel with correlated interference; in that case, the receiver side is interested in estimating both the source and the channel-state (interference). Inner and outer bounds on the source-interference distortion region are established. Our setting contains several interesting limiting cases. In the absence of fading and for the matched source-channel bandwidth, our system reverts to that of [26]; for the uncorrelated source-interference scenario without fading, our problem reduces to the one in [20] for the bandwidth reduction case. Moreover, the source-channel-state transmission scenario generalizes the setting in [29] to include fading and correlation between source and interference. The rest of the paper is organized as follows. In Section II, we present the problem formulation. In Section III, we derive an outer bound and introduce linear and tandem digital schemes. In Section IV, we derive inner bounds (achievable distortion region) under both matched and mismatched noise levels by proposing layered hybrid coding schemes. We extend these inner and outer bounds to the source-channel-state communication scenario in Section V. Finally, conclusions are drawn in Section VI.

Throughout the paper, we will use the following notation. Vectors are denoted by characters superscripted by their dimensions. For a given vector $X^N = (X(1), ..., X(N))^T$, we let $[X^N]_1^K$ and $[X^N]_{K+1}^N$ denote the sub-vectors $[X^N]_1^K \triangleq (X(1), ..., X(K))^T$ and $[X^N]_{K+1}^N \triangleq (X(K+1), ..., X(N))^T$, respectively, where $(\cdot)^T$ is the transpose operator. When there is no confusion, we also write $[X^N]_1^K$ as X^K. When all samples in a vector are independent and identically distributed (i.i.d.), we drop the indexing when referring to a sample in a vector (i.e., $X(i) = X$).

II. Problem Formulation and Main Contributions

We consider the transmission of a Gaussian source $V^K = (V(1), ..., V(K))^T \in \mathbb{R}^K$ over a Rayleigh fading channel in the presence of Gaussian interference $S^N \in \mathbb{R}^N$ known at the transmitter (see Fig. 1). The source vector V^K represents the first K samples of $V^{\max(K,N)}$; S^N is similarly defined. The source vector V^K, which is composed of i.i.d. samples, is transformed into an N dimensional channel input $X^N \in \mathbb{R}^N$ using a nonlinear mapping function, in general, $\alpha(.) : \mathbb{R}^K \times \mathbb{R}^N \to \mathbb{R}^N$. The received symbol is $Y^N = F^N(X^N + S^N) + W^N$, where addition and multiplication are component-wise, F^N represents an N-block Rayleigh fading that is independent of $(V^K; S^N; W^N)$ and known to the receiver side only, $X^N = \alpha(V^K, S^N)$, S^N is an i.i.d. Gaussian interference vector (with each sample $S \sim \mathcal{N}(0, \sigma_S^2)$) that is considered to be the output of a side channel with input $V^{\max(K,N)}$ as shown in Fig. 1, and each sample in the additive noise W^N is drawn from a Gaussian distribution ($W \sim \mathcal{N}(0, \sigma_W^2)$) independently from both the source and the interference. Unlike the typical dirty paper problem which assumes an AWGN channel with interference (that is uncorrelated to the source) [18], we consider a fading channel and assume that V^K and S^N are jointly Gaussian. Since the fading realization is known only at the receiver, we have partial knowledge of the actual interference $F^N S^N$ at the transmitter. In this work, we assume that only $V(i)$ and $S(i)$, $i = 1, ..., \min(K, N)$, are correlated according to the following covariance matrix

$$\Sigma_{VS} = \begin{bmatrix} \sigma_V^2 & \rho\sigma_V\sigma_S \\ \rho\sigma_V\sigma_S & \sigma_S^2 \end{bmatrix} \tag{1}$$

where σ_V^2, σ_S^2 are, respectively, the variance of the source and the interference, and ρ is the source-interference correlation coefficient. The system operates under an average power constraint P

$$\mathbb{E}[||\alpha(V^K, S^N)||^2]/N \leq P \tag{2}$$

where $\mathbb{E}[(\cdot)]$ denotes the expectation operator. The reconstructed signal is given by $\hat{V}^K = \gamma(Y^N, F^N)$, where the decoder is a mapping from $\mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^K$. The rate of the system is given by $r = \frac{N}{K}$ channel use/source symbol. When $r = 1$, the system has an equal-bandwidth between the source and the channel. For $r < 1$ $(r > 1)$, the system performs bandwidth reduction (expansion). According to the correlation model described above, note that for $r < 1$, the first N source samples $[V^K]_1^N$ and S^N are correlated via the covariance matrix in (1), while the remaining $K - N$ samples $[V^K]_{N+1}^K$ and S^N are independent. For $r > 1$, however, V^K and $[S^N]_1^K$ are correlated via the covariance matrix in (1), while V^K and $[S^N]_{K+1}^N$ are uncorrelated.

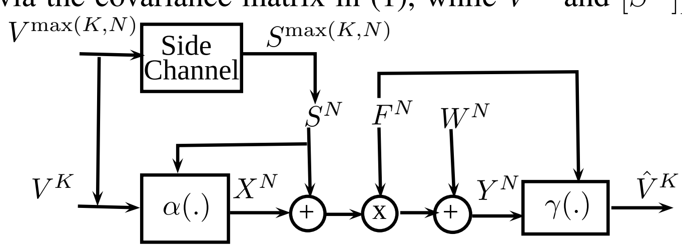

Fig. 1. A $K : N$ system structure over a fading channel with interference known at the transmitter side. The interference $S^{\max(K,N)}$ is assumed to be the output of a noisy side channel with input $V^{\max(K,N)}$. V^K represents the first K samples of $V^{\max(K,N)}$ (S^N is defined similarly). The fading coefficient is assumed to be known at the receiver side; the transmitter side, however, knows the fading distribution only.

In this paper, we aim to find a source-channel encoder α and decoder γ that minimize the mean square error (MSE) distortion $D = \mathbb{E}[||V^K - \hat{V}^K||^2]/K$ under the average power constraint in (2). For a particular coding scheme (α, γ), the performance is determined by the channel power constraint P, the fading distribution, the system rate r, and the incurred distortion D at the receiver. For a given power constraint P, fading distribution and rate r, the distortion region is defined as the closure of all distortions D_o for which (P, D_o) is achievable. A power-distortion pair is achievable if for any $\delta > 0$, there exist sufficiently large integers K and N with $N/K = r$, a pair of encoding and decoding functions (α, γ) satisfying (2), such that $D < D_o + \delta$. In this work, we analyze the distortion for equal and unequal bandwidths between the source and the channel with no constraint on the delay (i.e., both N and K tend to infinity with $\frac{N}{K} = r$ fixed). Our main contributions can be summarized as follows

- We derive inner and outer bounds for the system's distortion region for a Gaussian source over fading channel with correlated interference under equal and unequal source-channel bandwidths. The outer bounds are found by assuming full/partial knowledge of the interference at the decoder side. The inner bounds are derived by proposing hybrid coding schemes and analyzing their achievable distortion region. These schemes are based on proper combination of power splitting, bandwidth splitting, Wyner-Ziv and hybrid coding; a hybrid layer that uses the source and the interference is concatenated (superimposed) with a purely digital layer to achieve bandwidth expansion (reduction). Different from the problem considered in [26], we consider the case of fading and mismatch in the source-channel bandwidth. Our scheme offers better performance than the one in [26] under matched bandwidth (when accommodating the Costa coder in their scheme for fading channels). Moreover, our scheme is optimal when there is no fading and when the source-interference are either uncorrelated or fully correlated.
- As an application of the proposed schemes, we consider source-channel-state transmission over fading channels with correlated interference. In such case, the receiver aims to jointly estimate both the source signal as well as the channel-state. Inner and outer bounds are derived for this scenario. For the special case of uncorrelated source-interference over AWGN channels, we obtain the optimal source-channel-state distortion tradeoff; this result is analogous to the optimal rate-state distortion for the transmission of a finite discrete source over a Gaussian state interference derived in [29]. For correlated source-interference and fading channels, our inner bound performs close to the derived outer bound and outperforms the adapted scheme of [29].

III. Outer Bounds and Reference Systems

A. Outer Bounds

In [26] and [32], outer bounds on the achievable distortion were derived for point-to-point communication over Gaussian channel with correlated interference under matched bandwidth between the source and the channel. This was done by assuming full/partial knowledge of the interference at the decoder side. In this section, for the correlation model considered above, we derive outer bounds for the fading interference channel under unequal source-channel bandwidth. Since $S(i)$ and $V(i)$ are correlated for $i = 1, ..., \min(K, N)$, we have $S(i) = S_I(i) + S_D(i)$, with $S_D(i) = \frac{\rho\sigma_S}{\sigma_V}V(i)$ and $S_I \sim \mathcal{N}(0, (1-\rho^2)\sigma_S^2)$ are independent of each other. To derive an outer bound, we assume knowledge of both $(\tilde{S}^K, [S^N]_{K+1}^N)$ and F^N at the decoder side for the case of bandwidth expansion, where $\tilde{S}^K = \eta_1 S_I^K + \eta_2 S_D^K$ (the linear combination $\tilde{S}$ is motivated by [32]), and (η_1, η_2) is a pair of real parameters. For the bandwidth reduction case, we assume knowledge of $\tilde{S}^N$ and F^N at the decoder to derive a bound on the average distortion for the first N samples; the derivation of a bound on the average distortion for the remaining $K - N$ samples assumes knowledge of $[V^K]_1^N$ in addition to $\tilde{S}^N$.

Definition 1 *Let MSE$(Y;\tilde{S})$ be the distortion incurred from estimating Y based on $\tilde{S}$ using a linear minimum MSE estimator (LMMSE) denoted by $\gamma_{lmse}(\tilde{S}^K, f^K)$. This distortion, which is a function of η_1, η_2, $\mathbb{E}[XS_I]$ and $\mathbb{E}[XS_D]$, is given by MSE$(Y;\tilde{S}) = \mathbb{E}[(Y - \gamma_{lmse}(\tilde{S}^K, f^K))^2] = \left(\mathbb{E}[Y^2] - \frac{(\mathbb{E}[Y\tilde{S}])^2}{\mathbb{E}[\tilde{S}^2]}\right)$, where $\mathbb{E}[Y^2] = f^2(P + \sigma_S^2 + 2(\mathbb{E}[XS_I + XS_D])) + \sigma_W^2$, $\mathbb{E}[Y\tilde{S}] = f\big(\mathbb{E}[X(\eta_1 S_I + \eta_2 S_D)] + \mathbb{E}[\eta_1 S_I^2 + \eta_2 S_D^2]\big)$ and $\mathbb{E}[\tilde{S}^2] = \mathbb{E}[\eta_1^2 S_I^2 + \eta_2^2 S_D^2]$. These terms will be used in Lemmas 1 and 2.*

Lemma 1 *For a $K : N$ bandwidth expansion system with $N \geq K$ (the matched case is treated as a special case), the outer bound on the system's distortion D can be expressed as follows*

$$D \geq D_{ob} \triangleq \sup_{\eta_1,\eta_2} \inf_{\substack{X: \\ |\mathbb{E}[XS_I]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[S_I^2]} \\ |\mathbb{E}[XS_D]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[S_D^2]}}} \left\{ \frac{Var(V|\tilde{S})}{exp\left\{\mathbb{E}_F\left[\log\left(\left(\frac{MSE(Y;\tilde{S})}{\sigma_W^2}\right)\left(\frac{f^2P + \sigma_W^2}{\sigma_W^2}\right)^{r-1}\right)\right]\right\}} \right\} \qquad (3)$$

where $Var(V|\tilde{S}) = \sigma_V^2\left(1 - \frac{\eta_2^2\rho^2}{\eta_1^2(1-\rho^2) + \eta_2^2\rho^2}\right)$ is the variance of V given $\tilde{S}$.

Proof: For a $K : N$ system with $N \geq K$, we have the following

$$
\begin{aligned}
\frac{K}{2}\log\frac{\mathrm{Var}(V|\tilde{S})}{D} &\leq I(V^K;\hat{V}^K|\tilde{S}^K,[S^N]_{K+1}^N,F^N) \leq I(V^K;Y^N|\tilde{S}^K,[S^N]_{K+1}^N,F^N) \\
&= h(Y^N|\tilde{S}^K,[S^N]_{K+1}^N,F^N) - h(Y^N|V^K,S^N,F^N) \\
&\leq h(Y^K|\tilde{S}^K,F^K) + h([Y^N]_{K+1}^N\big|[S^N]_{K+1}^N,[F^N]_{K+1}^N) - h(Y^N|V^K,S^N,F^N) \\
&= \mathbb{E}_F\left[h(Y^K|\tilde{S}^K,f^k) + h([Y^N]_{K+1}^N\big|[S^N]_{K+1}^N,[f^n]_{k+1}^n)\right] - h(W^N) \\
&\leq \mathbb{E}_F\left[\frac{K}{2}\log 2\pi e(\mathrm{MSE}(Y;\tilde{S})) + \frac{N-K}{2}\log 2\pi e(f^2P+\sigma_W^2)\right] - \frac{N}{2}\log 2\pi e\sigma_W^2 \\
&= \mathbb{E}_F\left[\frac{K}{2}\log\left(\frac{\mathrm{MSE}(Y;\tilde{S})}{\sigma_W^2}\right) + \frac{N-K}{2}\log\left(\frac{f^2P+\sigma_W^2}{\sigma_W^2}\right)\right] \qquad (4)
\end{aligned}
$$

where we used $h(Y^K|\tilde{S}^K,f^K) \leq h(Y^K - \gamma_{\mathrm{lmse}}(\tilde{S}^K,f^K)) \leq \frac{K}{2}\log 2\pi e\left(\mathrm{MSE}(Y;\tilde{S})\right)$. By the Cauchy-Schwarz inequality, we have $|\mathbb{E}[XS_I]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[S_I^2]}$ and $|\mathbb{E}[XS_D]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[S_D^2]}$. For a given η_1 and η_2, we have to choose the highest value of $\mathrm{MSE}(Y;\tilde{S})$ over $\mathbb{E}[XS_D]$ and $\mathbb{E}[XS_I]$; then we need to maximize the right-hand side of (3) over η_1 and η_2. Note that most inequalities follow from rate-distortion theory, the data processing inequality and the facts that conditioning reduces differential entropy and that the Gaussian distribution maximizes differential entropy. ■

Lemma 2 *For $K : N$ bandwidth reduction ($K > N$), the outer bound on D is given by*

$$
D \geq D_{ob}(\xi^*) \triangleq \sup_{\eta_1,\eta_2}\inf_{\xi}\inf_{\substack{X:\\ |\mathbb{E}[XS_I]|\leq\sqrt{(1-\xi)P\mathbb{E}[S_I^2]}\\ |\mathbb{E}[XS_D]|\leq\sqrt{(1-\xi)P\mathbb{E}[S_D^2]}}}\left\{r\frac{\mathit{Var}(V|\tilde{S})}{\mathit{exp}\left\{\mathbb{E}_F\left[\log\left(\frac{\mathit{MSE}(Y;\tilde{S})}{\xi Pf^2+\sigma_W^2}\right)\right]\right\}} + (1-r)\frac{\sigma_V^2}{\mathit{exp}\left\{\mathbb{E}_F\left[\frac{N}{K-N}\log\left(\frac{\xi Pf^2+\sigma_W^2}{\sigma_W^2}\right)\right]\right\}}\right\} \qquad (5)
$$

where $\xi \in [0,\ 1]$.

Proof: We start by decomposing the average MSE distortion as follows

$$
\begin{aligned}
D &= \frac{1}{K}\mathbb{E}[||V^K-\hat{V}^K||^2] = \frac{1}{K}\left(\mathbb{E}[||V^N-\hat{V}^N||^2] + \mathbb{E}[||[V^K]_{N+1}^K - [\hat{V}^K]_{N+1}^K||^2]\right) \\
&= \frac{N}{K}\left(\frac{1}{N}\mathbb{E}[||V^N-\hat{V}^N||^2]\right) + \frac{K-N}{K}\left(\frac{1}{K-N}\mathbb{E}[||[V^K]_{N+1}^K - [\hat{V}^K]_{N+1}^K||^2]\right) \\
&= rD_1 + (1-r)D_2 \qquad (6)
\end{aligned}
$$

where D_1 and D_2 are the average distortion in reconstructing V^N and $[V^K]_{N+1}^K$, respectively. To find an outer bound on D, we derive bounds on both D_1 and D_2. To bound D_1, We can write the following expression

$$\begin{aligned}
\frac{N}{2}\log\frac{\text{Var}(V|\tilde{S})}{D_1} &\leq I(V^N;\hat{V}^N|\tilde{S}^N,F^N) \leq I(V^N;Y^N|\tilde{S}^N,F^N) \\
&= h(Y^N|\tilde{S}^N,F^N) - h(Y^N|\tilde{S}^N,V^N,F^N) \\
&= h(Y^N|\tilde{S}^N,F^N) - h(Y^N|S^N,V^N,F^N) \\
&\overset{(a)}{\leq} \mathbb{E}_F\left[\frac{N}{2}\log 2\pi e(\text{MSE}(Y;\tilde{S})) - \frac{N}{2}\log 2\pi e(\xi P f^2 + \sigma_W^2)\right] \\
&\leq \sup_{Y\in\mathcal{A}} \mathbb{E}_F\left[\frac{N}{2}\log\left(\frac{\text{MSE}(Y;\tilde{S})}{\xi P f^2+\sigma_W^2}\right)\right] \qquad (7)
\end{aligned}$$

where the set $\mathcal{A} = \{Y : h(Y^N|S^N,V^N,F^N) = \mathbb{E}_F\left[\frac{N}{2}\log 2\pi e(\xi P f^2 + \sigma_W^2)\right]\}$. Note that in (7)-$(a)$ we use the fact that $h(Y^N|S^N,V^N,F^N) = \mathbb{E}_F\left[\frac{N}{2}\log 2\pi e\left(\xi P f^2 + \sigma_W^2\right)\right]$, for some $\xi \in [0 \;\; 1]$. This can be shown by noting that the following inequality holds $\frac{N}{2}\log 2\pi e(\sigma_W^2) = h(W^N) \leq h(Y^N|S^N,V^N,F^N) \leq h(F^N X^N + W^N|F^N) = \mathbb{E}_F[\frac{N}{2}\log 2\pi e(P f^2 + \sigma_W^2)]$; as a result, there is a $\xi \in [0 \;\; 1]$ such that $h(Y^N|S^N,V^N,F^N) = \mathbb{E}_F\left[\frac{N}{2}\log 2\pi e\left(\xi P f^2 + \sigma_W^2\right)\right]$. Moreover in (7)-$(a)$, we used the fact that

$$\begin{aligned}
h(Y^N|\tilde{S}^N,F^N) &= \mathbb{E}_F[h(Y^N|\tilde{S}^N,f^n)] = \mathbb{E}_F[h(Y^N-\gamma_{lmse}(\tilde{S}^N,f^n)|\tilde{S}^N,f^n)] \\
&\leq \mathbb{E}_F[h(Y^N-\gamma_{lmse}(\tilde{S}^N,f^n))] \leq \frac{N}{2}\mathbb{E}_F[\log 2\pi e(\text{MSE}(Y;\tilde{S}))]. \qquad (8)
\end{aligned}$$

Similarly, to derive a bound on D_2, we have the following

$$\begin{aligned}
\frac{K-N}{2}\log\frac{\sigma_V^2}{D_2} &\leq I([V^K]_{N+1}^K;[\hat{V}^K]_{N+1}^K|S^N,V^N,F^N) \leq I([V^K]_{N+1}^K;Y^N|S^N,V^N,F^N) \\
&= h(Y^N|S^N,V^N,F^N) - h(Y^N|S^N,V^N,[V^K]_{N+1}^K,F^N) \\
&= h(Y^N|S^N,V^N,F^N) - h(Y^N|S^N,V^K,F^N) \\
&= \mathbb{E}\left[\frac{N}{2}\log\left(\frac{\xi P f^2+\sigma_W^2}{\sigma_W^2}\right)\right] \qquad (9)
\end{aligned}$$

where in the last equality, we used $h(Y^N|S^N,V^N,F^N) = \mathbb{E}_F\left[\frac{N}{2}\log 2\pi e\left(\xi P f^2 + \sigma_W^2\right)\right]$ as shown earlier. Note that since we do not know the value of ξ, the overall distortion has to be minimized over the parameter ξ. Now using (7) and (9) in (6), we have the following bound

$$D \geq \inf_{\xi}\inf_{Y\in\mathcal{A}}\left\{ r\frac{\text{Var}(V|\tilde{S})}{\exp\left\{\mathbb{E}_F\left[\log\left(\frac{\text{MSE}(Y;\tilde{S})}{\xi P f^2+\sigma_W^2}\right)\right]\right\}} + (1-r)\frac{\sigma_V^2}{\exp\left\{\mathbb{E}_F\left[\frac{N}{K-N}\log\left(\frac{\xi P f^2+\sigma_W^2}{\sigma_W^2}\right)\right]\right\}}\right\} \qquad (10)$$

where the $\sup$ in (7) is manifested as $\inf$ on the distortion. Note that the above sequence of inequalities in (8) becomes equalities when Y is conditionally Gaussian given F and when $Y - \gamma_{lmse}(\tilde{S}, f)$ and $\tilde{S}$ are jointly Gaussian and orthogonal to each other given F; this happens when X^* is jointly Gaussian with S, V and W given F. Hence, the $\sup$ in (7) happens when X^* is Gaussian. Now we write $X^* = N_\xi^* + X_\xi^*$, where $N_\xi^* \sim \mathcal{N}(0, \xi P)$ is independent of (V, S) and $X_\xi^* \sim \mathcal{N}(0, (1-\xi)P)$ is a function of (V, S). Note that X_ξ^* is independent of N_ξ^*. As a result , the equality $h(Y^N|S^N, V^N, F^N) = \mathbb{E}_F\left[\frac{N}{2}\log 2\pi e\left(\xi P f^2 + \sigma_W^2\right)\right]$ still holds and hence $Y^* \in \mathcal{A}$, $\mathbb{E}[Y^2] = f^2(P + \sigma_S^2 + 2(\mathbb{E}[X_\xi^* S_I + X_\xi^* S_D])) + \sigma_W^2$ and $\mathbb{E}[Y\tilde{S}] = f\big(\mathbb{E}[X_\xi^*(\eta_1 S_I + \eta_2 S_D)] + \mathbb{E}[\eta_1 S_I^2 + \eta_2 S_D^2]\big)$. By the Cauchy-Schwarz inequality, $|\mathbb{E}[X^* S_I]| = |\mathbb{E}[X_\xi^* S_I]| \leq \sqrt{\mathbb{E}[(X_\xi^*)^2]\mathbb{E}[S_I^2]}$ and $|\mathbb{E}[X^* S_D]| = |\mathbb{E}[X_\xi^* S_D]| \leq \sqrt{\mathbb{E}[(X_\xi^*)^2]\mathbb{E}[S_D^2]}$. Hence we maximize the value of MSE$(Y;\tilde{S})$ over X or equivalently over $\mathbb{E}[X_\xi^* S_I]$ and $\mathbb{E}[X_\xi^* S_D]$ satisfying the above constraints. Finally, the parameters η_1 and η_2 are chosen so that the right hand side of (5) is maximized. ■

B. Linear Scheme

In this section, we assume that the encoder transforms the K-dimensional signal V^K into an N-dimensional channel input X^N using a linear transformation according to

$$X^N = \alpha(V^K, S^N) = \mathbf{T}V^K + \mathbf{M}S^N \tag{11}$$

where $\mathbf{T}$ and $\mathbf{M}$ are $\mathbb{R}^{N\times K}$ and $\mathbb{R}^{N\times N}$ matrices, respectively. In such case, Y^N is conditionally Gaussian given F^N and the minimum MSE (MMSE) decoder is a linear estimator, with, $\hat{V}^K = \Sigma_{VY}\Sigma_Y^{-1}Y^N$, where $\Sigma_{VY} = \mathbb{E}\big[(V^K)(Y^N)^T\big]$ and $\Sigma_Y = \mathbb{E}\big[(Y^N)(Y^N)^T\big]$. The matrices $\mathbf{T}$ and $\mathbf{M}$ can be found (numerically) by minimizing the MSE distortion $D_{linear} = \mathbb{E}_F\left[\frac{1}{K}tr\left\{\sigma_V^2\mathbf{I}_{K\times K} - \Sigma_{VY}\Sigma_Y^{-1}\Sigma_{VY}^T\right\}\right]$ under the power constraint in (2), where $tr(.)$ is the trace operator and $\mathbf{I}_{K\times K}$ is a $K\times K$ identity matrix. Note that by setting $\mathbf{M}$ to be the zero matrix and $\mathbf{T} = \sqrt{P/\sigma_V^2}\mathbf{I}_{N\times K}$, the system reduces to the uncoded scheme. Focusing on the matched case $(K = N)$, we have the following lemma for finite block length K.

Lemma 3 *For the matched-bandwidth source-channel coding of a Gaussian source transmitted over an AWGN fading channel with correlated interference, the distortion lower bound for any linear scheme is achieved with single-letter linear codes.*

Proof: Recall that since V^K and S^K are correlated, we have $S^K = \frac{\rho\sigma_S}{\sigma_V} V^K + N_\rho^K$, where the samples in N_ρ^K are i.i.d. Gaussian with common variance $\sigma_S^2(1-\rho^2)$. As a result and using (11)

$$\begin{aligned} Y^K &= \mathbf{F}\left(\mathbf{T} + \frac{\rho\sigma_S}{\sigma_V}\mathbf{M} + \frac{\rho\sigma_S}{\sigma_V}\mathbf{I}_{K\times K}\right) V^K + \mathbf{F}\left(\mathbf{M} + \mathbf{I}_{K\times K}\right) N_\rho^K + W^K \\ &= \mathbf{F}\tilde{\mathbf{T}} V^K + \mathbf{F}\tilde{\mathbf{M}} N_\rho^K + W^K \end{aligned} \tag{12}$$

where $\mathbf{F} = diag(F^K)$ is a diagonal matrix that represents the fading channel, $\tilde{\mathbf{M}} = (\mathbf{M} + \mathbf{I}_{K\times K})$ and $\tilde{\mathbf{T}} = \left(\mathbf{T} + \frac{\rho\sigma_S}{\sigma_V}\mathbf{M} + \frac{\rho\sigma_S}{\sigma_V}\mathbf{I}_{K\times K}\right)$. After some manipulation, the distortion D_{linear} is given by

$$\begin{aligned} D_{linear} &= \frac{1}{K}\mathbb{E}_F\left[tr\left\{\left(\tilde{\mathbf{T}}^T\mathbf{F}^T[\sigma_S^2(1-\rho^2)\mathbf{F}\tilde{\mathbf{M}}\tilde{\mathbf{M}}^T\mathbf{F}^T + \sigma_W^2\mathbf{I}_{K\times K}]^{-1}\mathbf{F}\tilde{\mathbf{T}} + \sigma_V^{-2}\mathbf{I}_{K\times K}\right)^{-1}\right\}\right] \\ &= \frac{1}{K}\mathbb{E}_F\left[tr\left\{\left(\mathbf{Q}\mathbf{F}^T\mathbf{R}\mathbf{F} + \sigma_V^{-2}\mathbf{I}_{K\times K}\right)^{-1}\right\}\right] \end{aligned} \tag{13}$$

where we define $\mathbf{Q} = \tilde{\mathbf{T}}\tilde{\mathbf{T}}^T$, $\mathbf{R} = [\sigma_S^2(1-\rho^2)\mathbf{F}\tilde{\mathbf{M}}\tilde{\mathbf{M}}^T\mathbf{F}^T + \sigma_W^2\mathbf{I}_{K\times K}]^{-1}$ and use the fact that for any square matrices $\mathbf{A}$ and $\mathbf{B}$, $tr\left(\mathbf{I} + \mathbf{AB}\right)^{-1} = tr\left(\mathbf{I} + \mathbf{BA}\right)^{-1}$ [33]. Now by noting that for any positive-definite $K \times K$ square matrix $\mathbf{D}$, $tr(\mathbf{D}^{-1}) \geq \sum_{i=1}^K D_{ii}^{-1}$ [33], where D_{ii} denotes the diagonal elements in $\mathbf{D}$ and equality holds iff $\mathbf{D}$ is diagonal, we can write the following

$$D_{linear} \geq \frac{1}{K}\sum_{i=1}^{K}\frac{1}{Q_{ii}|F_{ii}|^2 R_{ii} + \sigma_V^{-2}}. \tag{14}$$

Equality in (14) holds iff $\mathbf{Q}$ and $\mathbf{R}$ are diagonal; hence the optimal solution gives a diagonal $\mathbf{T}$ and $\mathbf{M}$. Thus, any linear coding can be achieved in a scalar form without performance loss. ■

C. Tandem Digital Scheme

In [34], Gel'fand and Pinsker showed that the capacity of a point-to-point communication with side information (interference) known at the encoder side is given by

$$C = \max_{p(u,x|s)} I(U;Y) - I(U;S) \tag{15}$$

where the maximum is over all joint distributions of the form $p(s)p(u,x|s)p(y|x,s)$ and U denotes an auxiliary random variable. In [18], Costa showed that using $U = X + \alpha S$, with $\alpha = \frac{P}{P+\sigma_W^2}$ over AWGN channel with interference known at the transmitter, the achievable capacity is $C = \frac{1}{2}\log\left(1 + \frac{P}{\sigma_W^2}\right)$, which coincides with the capacity when both encoder and decoder know the interference S. As a result, this choice of U is optimal in terms of maximizing capacity. Next, we adapt the Costa scheme for the fading channel; we choose $U = X + \alpha S$ as above, where α is redesigned to fit our problem. Using (15) and by interpreting the fading F as a second channel output, an achievable rate R is given by

$$R = I(U;Y,F) - I(U;S) = I(U;Y|F) - I(U;S) \tag{16}$$

where we used the fact that $I(U;F) = 0$. After some manipulations, the rate R is

$$R = \mathbb{E}_F\left[\frac{1}{2}\log\left(\frac{P[f^2(P+\sigma_S^2)+\sigma_W^2]}{P\sigma_S^2 f^2(1-\alpha)^2+\sigma_W^2(P+\alpha^2\sigma_S^2)}\right)\right]. \tag{17}$$

To find α, we minimize the expected value of the denominator in (17) (i.e., $\mathbb{E}_F[P\sigma_S^2 f^2(1-\alpha)^2+\sigma_W^2(P+\alpha^2\sigma_S^2)]$). As a result, we choose $\alpha = \frac{P\mathbb{E}[f^2]}{P\mathbb{E}[f^2]+\sigma_W^2}$ for finite noise levels. Note that this choice of α is independent of S and depends on the second order statistics of the fading. In [35], the authors show that by choosing $\alpha = \frac{P}{P+\sigma_W^2}$, Costa coding maximizes the achievable rate for fading channels in the limits of both high and low noise levels.

The tandem scheme is based on the concatenation of an optimal source code and the adapted Costa coding (described above). The optimal source code quantizes the analog source with a rate close to that in (17), and the adapted Costa coder achieves a rate equal to (17). Hence, from the lossy JSCC theorem, the MSE distortion for a $K:N$ system can be expressed as follows

$$D_{tandem} = \frac{\sigma_V^2}{\exp\left\{\mathbb{E}_F\left[r\log\left(\frac{P[f^2(P+\sigma_S^2)+\sigma_W^2]}{P\sigma_S^2 f^2(1-\alpha)^2+\sigma_W^2(P+\alpha^2\sigma_S^2)}\right)\right]\right\}} \tag{18}$$

where $r = N/K$ is the system's rate. Note that the performance of this scheme does not improve when the noise level decreases (levelling-off effect) or in the presence of correlation between the source and the interference.

Remark 1 *For the AWGN channel, the distortion of the tandem scheme in* (18) *can be simplified as follows* $D_{tandem} = \sigma_V^2/(1+P/\sigma_W^2)^r$. *This can be shown by setting* $\alpha = P/(P+\sigma_W^2)$ *and cancelling out the expectation in* (18). *This scheme is optimal for the uncorrelated case* ($\rho = 0$)*.*

IV. Distortion Region for the Layered Schemes

In this section, we propose layered schemes based on Wyner-Ziv and HDA coding for transmitting a Gaussian source over a fading channel with correlated interference. These schemes require proper combination of power splitting, bandwidth splitting, rate splitting, Wyner-Ziv and HDA coding. A performance analysis in the presence of noise mismatch is also conducted.

A. Scheme 1: Layering Wyner-Ziv Costa and HDA for Bandwidth Expansion

This scheme comprises two layers that output X_1^K and X_2^{N-K}. The channel input is obtained by multiplexing (concatenating) the output codeword of both layers $X^N = [X_1^K \; X_2^{N-K}]$ as shown in Fig. 2. The first layer is composed of two sublayers that are superimposed to produce the first K samples of the channel input $X_1^K = X_a^K + X_d^K$. The first sublayer is purely analog and consumes an average power of P_a; the output of this sublayer is given by $X_a^K = \sqrt{a}(\beta_1 V^K + \beta_2 S^K)$, where $\beta_1, \beta_2 \in [-1 \; 1]$, $a = \frac{P_a}{\beta_1^2\sigma_V^2 + \beta_2^2\sigma_S^2 + 2\beta_1\beta_2\rho\sigma_V\sigma_S}$ with $0 \leq P_a \leq P$. The second sublayer, that outputs X_d^K and consumes the remaining power $P_d = P - P_a$, encodes the source V^K using a Wyner-Ziv coder followed by a (generalized) Costa coder. The Wyner-Ziv encoder, which uses the fact that an estimate of V^K can be obtained at the decoder side, forms a random variable T_1^K as follows

$$T_1^K = \alpha_{wz_1} V^K + B_1^K \tag{19}$$

where each sample in B_1^K is a zero mean i.i.d. Gaussian, α_{wz_1} and the variance of B_1 are defined later. The encoding process starts by generating a K-length i.i.d. Gaussian codebook $\mathcal{T}_1$ of size $2^{KI(T_1;V)}$ and randomly assigning the codewords into 2^{KR_1} bins with R_1 defined later. For each source realization V^K, the encoder searches for a codeword $T_1^K \in \mathcal{T}_1$ such that (V^K, T_1^K) are jointly typical. In the case of success, the Wyner-Ziv encoder transmits the bin index of this codeword using Costa coding. The Costa coder, which treats the analog sublayer X_a^K in addition to S^K as interference, forms the following auxiliary random variable $U_{c_1}^K = X_d^K + \alpha_{c_1}\check{S}^K$, where $\check{S}^K = (X_a^K + S^K)$, the samples in X_d^K are i.i.d. zero mean Gaussian with variance $P_d = P - P_a$ and $0 \leq \alpha_{c_1} \leq 1$ is a real parameter. Note that X_d^K is independent of V^K and S^K. The encoding process of the Costa coding can be summarized as follows

- *Codebook Generation:* Generate a K-length i.i.d. Gaussian codebook $\mathcal{U}_{c_1}$ with $2^{KI(U_{c_1};Y_1,F)}$ codewords, where Y_1^K is the first K samples of the received signal Y^N. Every codeword is generated following the random variable $U_{c_1}^K$ and uniformly distributed over 2^{KR_1} bins. The codebook is revealed to both encoder and decoder.
- *Encoding:* For a given bin index (the output of the Wyner-Ziv encoder), the Costa encoder searches for a codeword $U_{c_1}^K$ such that the bin index of $U_{c_1}^K$ is equal to the Wyner-Ziv output and $(U_{c_1}^K, \check{S}^K)$ are jointly typical. In the case of success, the Costa encoder outputs $X_d^K = U_{c_1}^K - \alpha_{c_1}\check{S}^K$. Otherwise, an encoding failure is declared. Note that the probability of encoder failure vanishes by using $R_1 = I(U_{c_1}; Y_1, F) - I(U_{c_1}; \check{S})$.

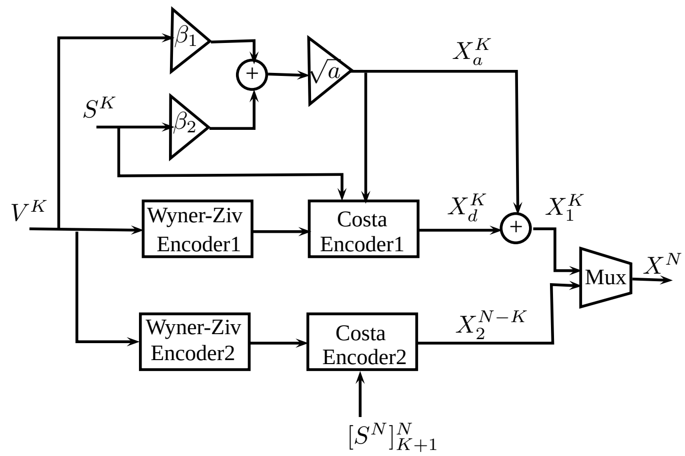

Fig. 2. Scheme 1 (bandwidth expansion) encoder structure.

The second layer, which outputs X_2^{N-K}, encodes V^K using a Wyner-Ziv with rate R_2 and a Costa coder that treats $[S^N]_{K+1}^N$ as interference. The Wyner-Ziv encoder, which uses the fact that an estimate of V^K is obtained from the first layer, forms the random variable T_2^K as follows

$$T_2^K = \alpha_{wz2} V^K + B_2^K \tag{20}$$

where the samples in B_2^K are i.i.d. and follow a zero mean Gaussian distribution, α_{wz2} and the variance of B_2 are defined later. The Costa coder forms the auxiliary random variable $U_{c_2}^{N-K} = X_2^{N-K} + \alpha_{c_2}[S^N]_{K+1}^N$, where the samples in X_2^{N-K} are i.i.d. zero mean Gaussian with variance P, and the real parameter α_{c_2} is defined later. The encoding process of the Wyner-Ziv and the Costa coder for the second layer is very similar to the one described for the first layer; hence, no details are provided.

At the receiver side, as shown in Fig. 3, from the first K components of the received signal $Y^N = [Y_1^K, Y_2^{N-K}] = F^N(X^N + S^N) + W^N$, where $Y_1^K = [Y^N]_1^K$ and $Y_2^{N-K} = [Y^N]_{K+1}^N$, the Costa decoder estimates the codeword $U_{c_1}^K$ by searching for a codeword $U_{c_1}^K$ such that $(U_{c_1}^K, Y_1^K, F^K)$ are jointly typical. By the result of Gelfand-Pinsker [34] (or Costa [18]) and by treating the fading coefficient F^K as a second channel output, the error probability of encoding and decoding the codeword $U_{c_1}^K$ vanishes as $K \to \infty$ if

$$\begin{aligned} R_1 &= I(U_{c_1}; Y_1, F) - I(U_{c_1}; \check{S}) = I(U_{c_1}; Y_1|F) - \big(h(U_{c_1}) - h(U_{c_1}|\check{S})\big) \\ &= h(U_{c_1}) + h(Y_1|F) - h(U_{c_1}, Y_1|F) - h(U_{c_1}) + h(U_{c_1}|\check{S}) \\ &= \mathbb{E}_F\left[\frac{1}{2}\log\left(\frac{P_d[f^2(P_d + \sigma_{\check{S}}^2) + \sigma_W^2]}{P_d\sigma_{\check{S}}^2 f^2(1-\alpha_{c_1})^2 + \sigma_W^2(P_d + \alpha_{c_1}^2\sigma_{\check{S}}^2)}\right)\right] \end{aligned} \tag{21}$$

where $\sigma_{\breve{S}}^2 = \mathbb{E}[(X_a + S)^2]$. We then obtain a linear MMSE estimate of V^K (based on Y_1^K and $U_{c_1}^K$), denoted by V_a^K. The distortion from estimating the source using V_a^K is given by

$$D_a = \mathbb{E}_F \left[\sigma_V^2 - \Gamma \Lambda^{-1} \Gamma^T \right] \tag{22}$$

where $\Lambda = \mathbb{E}[[U_{c_1} \quad Y_1]^T [U_{c_1} \quad Y_1]]$ is the covariance of $[U_{c_1} \quad Y_1]$ and $\Gamma = \mathbb{E}[V[U_{c_1} \quad Y_1]]$ is the correlation vector between V and $[U_{c_1} \quad Y_1]$. By using rate R_1 on the Wyner-Ziv encoder, the bin index of the Wyner-Ziv can be decoded correctly (with high probability). The Wyner-Ziv decoder then looks for a codeword T_1^K in this bin such that (T_1^K, V_a^K) are jointly typical (as $K \to \infty$, the probability of error in decoding T_1^K vanishes). A better estimate of V^K is then obtained based on V_a^K and the decoded codeword T_1^K. The distortion in the estimated source $\tilde{V}^K$ is then

$$\tilde{D} = \frac{D_a}{\exp\left\{ \mathbb{E}_F \left[\log \left(\frac{P_d [f^2 (P_d + \sigma_{\breve{S}}^2) + \sigma_W^2]}{P_d \sigma_{\breve{S}}^2 f^2 (1 - \alpha_{c_1})^2 + \sigma_W^2 (P_d + \alpha_{c_1}^2 \sigma_{\breve{S}}^2)} \right) \right] \right\}}. \tag{23}$$

Note that this distortion is equal to the distortion incurred when assuming that the side information V_a^K is also known at the transmitter side; this can be achieved by choosing $\alpha_{wz_1} = \sqrt{1 - \frac{\tilde{D}}{D_a}}$ and $B_1 \sim \mathcal{N}(0, \tilde{D})$ in (19) and using a linear MMSE estimator based on V_a^K and T_1^K. In contrast to the AWGN channel with correlated interference [26], a purely analog layer is not sufficient to accommodate for the correlation over AWGN fading channel with correlated interference; indeed using the knowledge of $U_{c_1}^K$ as a side information to obtain a better description of the Wyner-Ziv codewords T_1^K will achieve a better performance. From the last $N-K$ received symbols Y_2^{N-K},

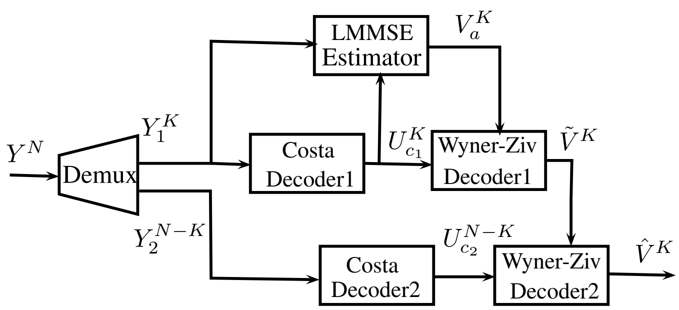

Fig. 3. Scheme 1 (bandwidth expansion) decoder structure.

the Costa decoder estimates the codeword $U_{c_2}^{N-K}$ by searching for a codeword $U_{c_2}^{N-K}$ such that $(U_{c_2}^{N-K}, Y_2^{N-K}, [F^N]_{K+1}^N)$ are jointly typical. The probability of error in encoding and decoding the codeword $U_{c_2}^{N-K}$ goes to zero by choosing

$$R_2 = I(U_{c_2}; Y_2, F) - I(U_{c_2}; S) = \mathbb{E}_F \left[\frac{1}{2} \log \left(\frac{P[f^2 (P + \sigma_S^2) + \sigma_W^2]}{P \sigma_S^2 f^2 (1 - \alpha_{c_2})^2 + \sigma_W^2 (P + \alpha_{c_2}^2 \sigma_S^2)} \right) \right] \tag{24}$$

where $\alpha_{c_2} = P\mathbb{E}[f^2]/(P\mathbb{E}[f^2]+\sigma_W^2)$ is found in a similar way as done in Sec. III-C. By using this rate, the Wyner-Ziv bin index can be decoded correctly (with high probability). The Wyner-Ziv decoder then looks for a codeword T_2^K in the decoded bin such that T_2^K and the side information from the first layer $\tilde{V}^K$ are jointly typical. A refined estimate of the source can be found using the side information $\tilde{V}^K$ and the decoded codeword T_2^K. The resulting distortion is then

$$D_{\text{Scheme 1}} = \inf_{\beta_1,\beta_2,P_a,\alpha_{c_1}} \left\{ \frac{\tilde{D}}{\exp\left\{\mathbb{E}_F\left[\log\left(\frac{P[f^2(P+\sigma_S^2)+\sigma_W^2]}{P\sigma_S^2 f^2(1-\alpha_{c_2})^2+\sigma_W^2(P+\alpha_{c_2}^2\sigma_S^2)}\right)^{r-1}\right]\right\}} \right\}. \tag{25}$$

Note that this distortion is equal to the distortion realized when assuming $\tilde{V}^K$ is also known at the transmitter side; this can be achieved using a linear MMSE estimator based on $[T_1 \;\; T_2 \;\; Y_1]$, and by setting $\alpha_{wz_2} = \sqrt{1 - \frac{D_{\text{Scheme 1}}}{\tilde{D}}}$ and $B_2 \sim \mathcal{N}(0, D_{\text{Scheme 1}})$ in (20).

Remark 2 *For AWGN channels with no fading, the same scheme can be used. In this case, the distortion from reconstructing the source can be expressed as follows*

$$D_{\textit{Scheme 1}} = \inf_{\beta_1,\beta_2,P_a} \left\{D_a/[(1+P/\sigma_W^2)^{r-1}(1+P_d/\sigma_W^2)]\right\}. \tag{26}$$

This distortion can be found by setting the fading coefficient $F=1$, $\alpha_{c_1} = P_d/(P_d+\sigma_W^2)$ and $\alpha_{c_2} = P/(P+\sigma_W^2)$ in (25)*. The distortion in* (26) *can be also achieved by replacing the sublayer that outputs X_d^K by an HDA Costa layer as we proposed in [28]. Note that using only Y_1^K as input to the LMMSE estimator in Fig. 3 is enough for the AWGN case. In such case, D_a in* (26) *can be simplified as follows*

$$D_a = \left(\sigma_V^2 - \frac{(\sqrt{a}\beta_1^2\sigma_V^2 + (\sqrt{a}\beta_2+1)\rho\sigma_V\sigma_S)^2}{P + (2\sqrt{a}\beta_2+1)\sigma_S^2 + 2\sqrt{a}\beta_1\rho\sigma_V\sigma_S + \sigma_W^2}\right). \tag{27}$$

Moreover, one can check that this scheme is optimal (for the AWGN channel) for $\rho = 0$ and $\rho = 1$. For $\rho = 0$, this happens by shutting down the analog sublayer (i.e., $P_a = 0$) in the scheme and using $(\eta_1 = 1, \eta_2 = 1)$ on the outer bound in (3)*. For the case of $\rho = 1$, the optimal power allocation for the scheme is $(P_a = P, P_d = 0)$. The resulting system's distortion can be shown to be equal to the outer bound in* (3) *for $(\eta_1 = 1, \eta_2 = 0)$.*

Scheme 1 under mismatch in noise levels: Next, we study the distortion of the proposed scheme in the presence of noise mismatch between the transmitter and the receiver. The actual channel noise power $\sigma_{W_a}^2$ is assumed to be lower than the design one σ_W^2 (i.e., $\sigma_{W_a}^2 < \sigma_W^2$).

Under such assumption, the Costa and Wyner-Ziv decoders are still able to decode correctly all codewords with low probability of error. After decoding T_1^K and T_2^K, a symbol-by-symbol linear MMSE estimator of V^K based on Y_1^K, T_1^K and T_2^K is calculated. Hence Scheme 1's distortion under noise mismatch is $D_{\text{(Scheme 1)-mis}} = \mathbb{E}_F\left[\sigma_V^2 - \Gamma^T\Lambda^{-1}\Gamma\right]$, where Λ is the covariance matrix of $[T_1 \;\; T_2 \;\; Y_1]$, and Γ is the correlation vector between V and $[T_1 \;\; T_2 \;\; Y_1]$. Note that $\sigma_{W_a}^2$ is used in the covariance matrix Λ instead of σ_W^2.

Remark 3 *When $\sigma_{W_a}^2 > \sigma_W^2$, all codewords cannot be decoded correctly at the receiver side; as a result we can only estimate the source vector V^K by applying a linear MMSE estimator based on the noisy received signal Y_1^K. The system's distortion in this case is given by*

$$D_{\textit{(Scheme 1)-mis}} = \mathbb{E}_F\left[\sigma_V^2 - \frac{f^2(\sqrt{a}\beta_1^2\sigma_V^2 + (\sqrt{a}\beta_2+1)\rho\sigma_V\sigma_S)^2}{f^2(P + (2\sqrt{a}\beta_2+1)\sigma_S^2 + 2\sqrt{a}\beta_1\rho\sigma_V\sigma_S) + \sigma_{W_a}^2}\right]. \tag{28}$$

B. Scheme 2: Layering Wyner-Ziv Costa and HDA for Bandwidth Reduction

In this section, we present a layered scheme for bandwidth reduction. This scheme comprises three layers that are superposed to produce the channel input $X^N = X_a^N + X_1^N + X_2^N$, where X_a^N, X_1^N and X_2^N denote the outputs of the first, second and third layers, respectively. The scheme's encoder structure is depicted in Fig. 4. Recall that we denote the first N samples of V^K by V^N and the last $K-N$ samples by $[V^K]_{N+1}^K$. The first layer is an analog layer that outputs $X_a^N = \sqrt{a}(\beta_1 V^N + \beta_2 S^N)$, a linear combination between the V^N and S^N, and consumes $P_a \le P$ as average power, where $\beta_1, \beta_2 \in [-1\ 1]$, and $a = \frac{P_a}{\beta_1^2\sigma_V^2 + \beta_2^2\sigma_S^2 + 2\beta_1\beta_2\rho\sigma_V\sigma_S}$ is a gain factor related to the power constraint P_a. The second layer, which operates on the first N samples of the source, encodes V^N using a Wyner-Ziv with rate R_1 followed by a Costa coder. The Wyner-Ziv encoder forms a random variable

$$T_1^N = \alpha_{wz_1} V^N + B_1^N \tag{29}$$

where the samples in B_1^N are i.i.d and follow a zero mean Gaussian distribution, the parameter α_{wz_1} and the variance of B_1 are related to the side information from the first layer and hence defined later. The Costa coder that treats both X_a^N and S^N as interference forms the following auxiliary random variable $U_{c_1}^N = X_1^N + \alpha_{c_1}\check{S}^N$, where the samples in X_1^N are i.i.d. zero mean Gaussian with variance $P_1 \le P - P_a$ and independent of the source and the interference, $\check{S}^N = X_a^N + S^N$ and $0 \le \alpha_{c_1} \le 1$ is a real parameter. The last layer encodes $[V^K]_{N+1}^K$ using an optimal source encoder with rate R_2 followed by a Costa coder. The Costa encoder, which treats the outputs of the first two layers (X_a^N, X_1^N) as well as S^N as known interference, forms the following auxiliary random variable $U_{c_2}^N = X_2^N + \alpha_{c_2}\tilde{S}^N$, where $\tilde{S}^N = (X_a^N + X_1^N + S^N)$, the samples in X_2^N are zero mean i.i.d. Gaussian with variance $P_2 = P - P_1 - P_a$ and $\alpha_{c_2} = P_2\mathbb{E}[f^2]/(P_2\mathbb{E}[f^2] + \sigma_W^2)$.

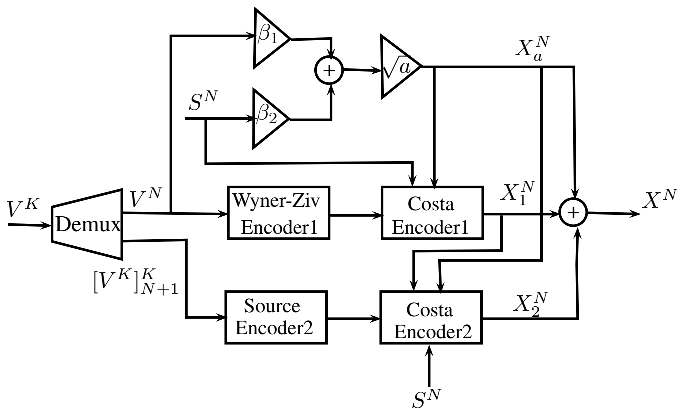

Fig. 4. Scheme 2 (bandwidth reduction) encoder structure.

At the receiver, as shown in Fig. 5, from the received signal Y^N the Costa decoder estimates $U_{c_1}^N$. By using a rate $R_1 = I(U_{c_1}; Y, F) - I(U_{c_1}; \check{S}) = \mathbb{E}_F\left[\frac{1}{2}\log\left(\frac{P_1[f^2(P_1+\sigma_{\check{S}}^2+P_2)+\sigma_W^2]}{P_1(\sigma_{\check{S}}^2)f^2(1-\alpha_{c_1})^2+(\sigma_W^2+f^2P_2)(P_1+\alpha_{c_1}^2\sigma_{\check{S}}^2)}\right)\right]$, where $\sigma_{\check{S}}^2 = \mathbb{E}[(X_a + S)^2]$, the Costa decoder (of the second layer) is able to estimate the codewords $U_{c_1}^N$ with vanishing error probability. We then obtain an estimate of V^N, denoted by V_a^N, using a linear MMSE estimator based on Y^N and $U_{c_1}^N$. The distortion from estimating V^N using V_a^N is then given by

$$D_a = \mathbb{E}_F\left[\sigma_V^2 - \Gamma\Lambda^{-1}\Gamma^T\right] \tag{30}$$

where Λ is the covariance of $[U_{c_1} \;\; Y]$ and Γ is the correlation vector between V and $[U_{c_1} \;\; Y]$. The Wyner-Ziv decoder (of the second layer) then looks for a codeword T_1^N such that (T_1^N, V_a^N) are jointly typical (as $N \to \infty$, the probability of error in decoding T_1^N vanishes). A better estimate of V^N is then obtained based on the side information V_a^N and the decoded codeword T_1^N. The distortion from reconstructing V^N is then given by

$$D_1 = \frac{D_a}{\exp\left(\mathbb{E}_F\left[\log\left(\frac{P_1[f^2(P_1+\sigma_{\check{S}}^2+P_2)+\sigma_W^2]}{P_1(\sigma_{\check{S}}^2)f^2(1-\alpha_{c_1})^2+(\sigma_W^2+f^2P_2)(P_1+\alpha_{c_1}^2\sigma_{\check{S}}^2)}\right)\right]\right)}. \tag{31}$$

Note that the distortion in (31) can be found by choosing $\alpha_{wz_1} = \sqrt{1 - \frac{D_1}{D_a}}$ and $B_1 \sim \mathcal{N}(0, D_1)$ in (29) and using a linear MMSE estimator based on V_a^N and T_1^N. To get an estimate of

$[V^K]_{N+1}^K$, we use a Costa decoder followed by a source decoder. Codewords of this layer can be decoded correctly (with high probability) by choosing the rate $R_2 = I(U_{c_2}; Y, F) - I(U_{c_2}; \tilde{S}) = \mathbb{E}_F \left[\frac{1}{2} \log \left(\frac{P_2[f^2(P_2+\sigma_{\tilde{S}}^2)+\sigma_W^2]}{P_2\sigma_{\tilde{S}}^2 f^2(1-\alpha_2)^2+\sigma_W^2(P_2+\alpha_2^2\sigma_{\tilde{S}}^2)} \right) \right]$, where $\sigma_{\tilde{S}}^2 = \mathbb{E}[(X_a + X_1 + S)^2]$. The distortion in reconstructing $[V^K]_{N+1}^K$ can be found by equating the rate-distortion function to the transmission rate R_2; this means that $\frac{K-N}{2} \log \frac{\sigma_V^2}{D_2} = (N)R_2$. As a result, the distortion in reconstructing $[V^K]_{N+1}^K$, denoted by D_2, is given by

$$D_2 = \frac{\sigma_V^2}{\exp \left\{ \mathbb{E}_F \left[\frac{r}{1-r} \log \left(\frac{P_2[f^2(P_2+\sigma_{\tilde{S}}^2)+\sigma_W^2]}{P_2\sigma_{\tilde{S}}^2 f^2(1-\alpha_2)^2+\sigma_W^2)(P_2+\alpha_2^2\sigma_{\tilde{S}}^2)} \right) \right] \right\}}. \tag{32}$$

Hence, the system's distortion is given by $D_{Scheme\ 2} = \inf_{\beta_1,\beta_2,P_a,P_1,\alpha_{c_1}} \left\{ rD_1 + (1-r)D_2 \right\}$.

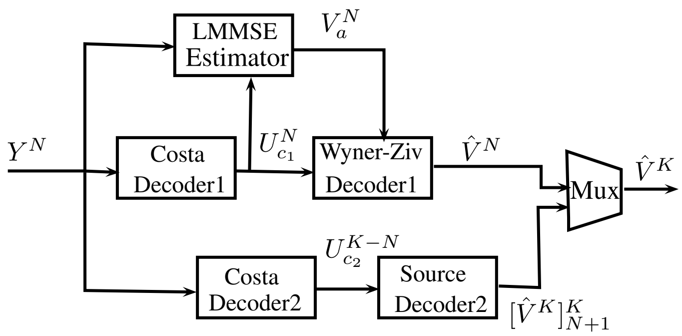

Fig. 5. Scheme 2 (bandwidth reduction) decoder structure.

Remark 4 *For the AWGN channel, the distortion D_1 and D_2 for the reduction case are*

$$D_1 = \frac{D_a}{1 + \frac{P_1}{P_2+\sigma_W^2}} \quad and \quad D_2 = \frac{\sigma_V^2}{\left(1 + \frac{P_2}{\sigma_W^2}\right)^{\frac{r}{1-r}}}. \tag{33}$$

Since for AWGN channel, the use of $U_{c_1}^N$ as input to the LMMSE estimator in Fig. 5 does not improve the performance, the distortion D_a admits a simplified expression as given in (27)*. The distortions in* (33) *can be derived by choosing $\alpha_{c_1} = \frac{P_1}{P_1+P_2+\sigma_W^2}$ and $\alpha_{c_2} = \frac{P_2}{P_2+\sigma_W^2}$. Note that this scheme is optimal for uncorrelated source-interference and for full correlation between the source and the interference. For the uncorrelated case, the analog layer is not needed ($P_a = 0$, $D_a = \sigma_V^2$) and the optimal power allocation between the two other layers can be derived by minimizing the resulting distortion with respect to P_1; the optimal power P_1 is $P_1^* = \sigma_W^2 \left[1 - \left(1 + \frac{P}{\sigma_W^2}\right)^{1-r} \right] + P$. For the case of full correlation between the (first N samples of the) source and the interference ($\rho = 1$), the second layer can be shut down ($P_1 = 0$) and the optimal P_a^* satisfies*

$$\sigma_W^2 \left(1 + \frac{\sigma_W}{\sqrt{P_a}}\right) \left(1 + \frac{P - P_a}{\sigma_W^2}\right)^{\frac{1}{1-r}} \left(P + \sigma_W^2 + \sqrt{P_a \sigma_V^2}\right) - \left(P + \sigma_W^2 + \sigma_V^2 + 2\sqrt{P_a \sigma_V^2}\right)^2 = 0.$$

Scheme 2 under mismatch in noise levels: We next examine the distortion of the proposed scheme in the presence of noise mismatch between the transmitter and the receiver. The actual channel noise power $\sigma_{W_a}^2$ is assumed to be lower than the design one σ_W^2 (i.e., $\sigma_{W_a}^2 < \sigma_W^2$). Under such assumption, the Costa and Wyner-Ziv decoders can decode all codewords with vanishing probability of error. The distortion in reconstructing $[V^K]_{N+1}^K$, D_{2-mis}, is hence the same as in the matched noise level case; and the distortion from reconstructing V^N is $D_{1-mis} = \mathbb{E}_F\left[\sigma_V^2 - \Gamma^T\Lambda^{-1}\Gamma\right]$, where Λ is the covariance matrix of $[T_1 \;\; Y]$, and Γ is the correlation vector between V and $[T_1 \;\; Y]$. As a result, the system's distortion is $D_{\text{(Scheme 2)-mis}} = rD_{1-mis} + (1-r)D_{2-mis}$. Note that $\sigma_{W_a}^2$ is used in Λ instead of σ_W^2 when computing D_{1-mis}.

Remark 5 *When $\sigma_{W_a}^2 > \sigma_W^2$, all codewords cannot be decoded correctly at the receiver side; as a result we can only estimate the source vector V^N by applying a linear MMSE estimator based on the noisy received signal Y^N. The system's distortion is then given by*

$$D_{\textit{(Scheme 2)-mis}} = r\mathbb{E}_F\left[\left(\sigma_V^2 - \frac{f^2(\sqrt{a}\beta_1^2\sigma_V^2 + (\sqrt{a}\beta_2+1)\rho\sigma_V\sigma_S)^2}{f^2(P + (2\sqrt{a}\beta_2+1)\sigma_S^2 + 2\sqrt{a}\beta_1\rho\sigma_V\sigma_S) + \sigma_{W_a}^2}\right)\right] + (1-r)\sigma_V^2.$$

C. Numerical Results

In this section, we assume an i.i.d. zero-mean Gaussian source with unitary variance that is transmitted over an AWGN Rayleigh fading channel with Gaussian interference. The interference power is $\sigma_S^2 = 1$, the power constraint is set to $P = 1$ and the Rayleigh fading has $\mathbb{E}[F^2] = 1$. To evaluate the performance, we consider the signal-to-distortion ratio (SDR $= \mathbb{E}[||V^K||^2]/\mathbb{E}[||V^K - \hat{V}^K||^2]$); the designed channel signal-to-noise ratio (CSNR $\triangleq \frac{P\mathbb{E}[F^2]}{\sigma_W^2}$) is set to 10 dB for all numerical results. Fig. 6, which considers the AWGN channel, shows the SDR performance versus the correlation coefficient ρ for bandwidth expansion $(r = 2)$ and matched noise levels between the transmitter and receiver. We note that the proposed scheme outperforms the tandem Costa reference scheme (described in Sec. III-C) and performs very close to the "best" derived outer bound for a wide range of correlation coefficients. Although not shown, the proposed scheme also outperforms significantly the linear scheme of Sec. III-B. For the limiting cases of $\rho = 0$ and 1, we can notice that the SDR performance of the proposed scheme coincides with the outer bound and hence is optimal. Figs. 7, 8 and 9 show the SDR performance versus ρ for the fading channel with interference under matched noise levels and for $r = 1, 2$ and $1/2$, respectively. As in the case of the AWGN channel, we remark that the proposed HDA schemes

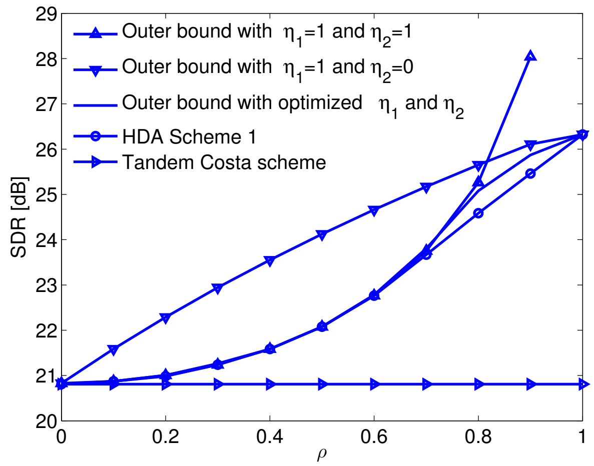

Fig. 6. Performance of HDA Scheme 1 ($r = 2$) over the AWGN channel under matched noise levels for different correlation coefficients. Tandem scheme and outer bounds on SDR are plotted. The graph is made for $P = 1$, $\sigma_S^2 = 1$ and CSNR=10 dB.

outperform the tandem Costa and the linear schemes and perform close to the best outer bound. In contrast to the AWGN case, the proposed scheme never coincides with the outer bound for finite noise levels; this can be explained from the fact that the (generalized) Costa and linear scheme are not optimal for the fading case. The sub-optimality (assuming that the outer bound is tight) of the generalized Costa coding comes from the form of the auxiliary random variable. We choose the same form as the one used for AWGN channels (a form linear in S); it remains unclear if such auxiliary random variable is optimal for fading channels. Note that using the result in [35], one can easily show that our schemes are optimal for $\rho = 0$ in the limits of high and low noise levels. As a result, the auxiliary random variable used for the Costa coder is optimal in the noise level limits.

Fig. 10 shows the SDR performance versus CSNR levels under mismatched noise levels. All schemes in Fig. 10 are designed for CSNR=10 dB, $r = 1$ and $\rho = 0.7$. The true CSNR varies between 0 and 35 dB. We observe that the proposed scheme is resilient to noise mismatch due to its hybrid digital-analog nature. As the correlation coefficient values decreases, the power allocated to the analog layer decreases. Hence, the SDR gap between the proposed and the tandem Costa scheme under mismatched noise levels decreases and the robustness (which is the trait of analog schemes) reduces.

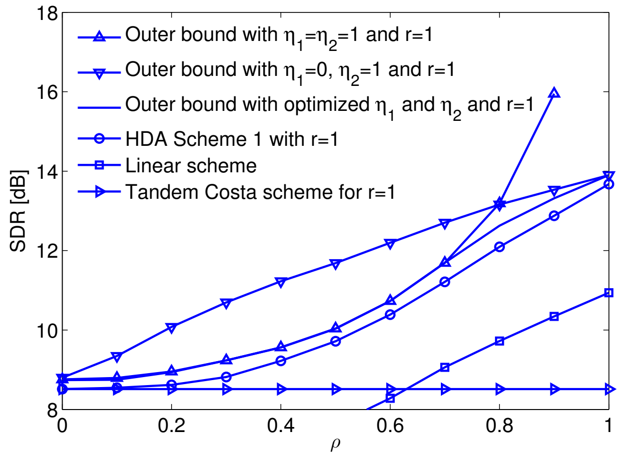

Fig. 7. Performance of Scheme 1 ($r = 1$) over the fading channel under matched noise levels for different correlation coefficient. Tandem, linear schemes and outer bounds on SDR are plotted. The graph is made for $P = 1$, $\sigma_S^2 = 1$, CSNR=10 dB and $\mathbb{E}[F^2] = 1$.

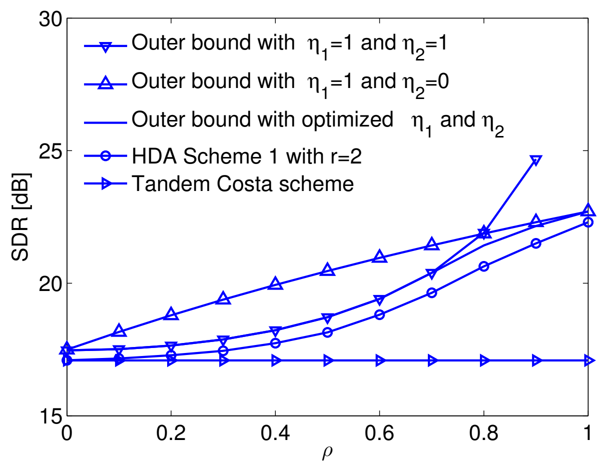

Fig. 8. Performance of Scheme 1 ($r = 2$) over the fading channel under matched noise levels for different correlation coefficient. Tandem scheme and outer bounds on SDR are plotted. The graph is given for $P = 1$, $\sigma_S^2 = 1$, CSNR=10 dB and $\mathbb{E}[F^2] = 1$.

V. JSCC for Source-Channel-State Transmission

As an application of the joint source-channel coding problem examined in this paper we consider the transmission of analog source-channel-state pairs over a fading channel with Gaussian state interference. We establish inner and outer bounds on the source-interference distortion for the fading channel. The only difference between this problem and that examined in the previous sections is that the decoder is also interested in estimating the interference S^N. For simplicity, we focus on the matched bandwidth case (i.e., $K = N$); the unequal source-channel bandwidth case can be treated in a similar way as in Section IV. We also assume that the decoder has knowledge of the fading. We denote the distortion from reconstructing the source and the interference by

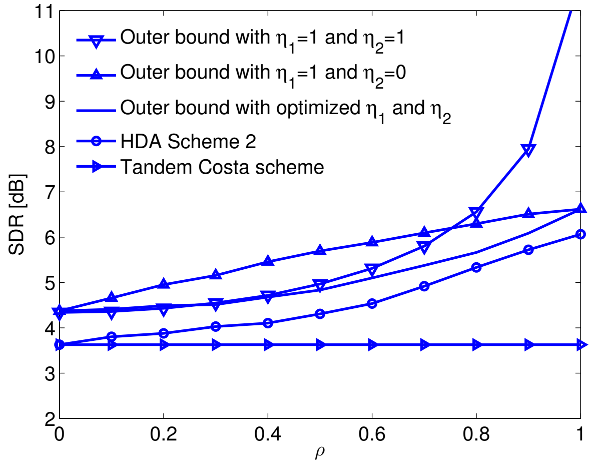

Fig. 9. Performance of Scheme 2 ($r = 1/2$) over the fading channel under matched noise levels for different ρ. Tandem scheme and outer bounds on SDR are plotted. The graph is given for $P = 1$, $\sigma_S^2 = 1$, CSNR=10 dB and $\mathbb{E}[F^2] = 1$.

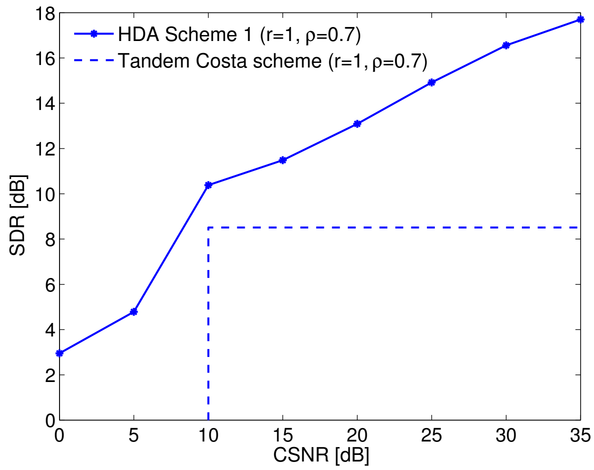

Fig. 10. Performance of Scheme 1 ($r = 1$) over the fading channel under mismatched noise levels. Tandem scheme, analog and upper bounds on SDR are plotted. The system is designed for $P = 1$, $\sigma_S^2 = 1$, CSNR=10 dB, $\rho = 0.7$ and $\mathbb{E}[F^2] = 1$.

$D_v = \frac{1}{K}\mathbb{E}[||V^K - \hat{V}^K||^2]$ and $D_s = \frac{1}{K}\mathbb{E}[||S^K - \hat{S}^K||^2]$, respectively. For a given power constraint P, a rate r and a Rayleigh fading channel, the distortion region is defined as the closure of all distortion pair $(\tilde{D}_v, \tilde{D}_s)$ for which $(P, \tilde{D}_v, \tilde{D}_s)$ is achievable, where a power-distortion triple is achievable if for any $\delta_v, \delta_s > 0$, there exist sufficiently large integers K and N with $N/K = r$, encoding and decoding functions satisfying (2), such that $D_v < \tilde{D}_v + \delta_v$ and $D_s < \tilde{D}_s + \delta_s$.

A. Outer Bound

Lemma 4 *For the matched bandwidth case, the outer bound on the distortion region* (D_v, D_s) *can be expressed as follows*

$$D_v \geq \frac{Var(V|S)}{\exp\left\{\mathbb{E}_F\left[\log\frac{\zeta P|f|^2+\sigma_W^2}{\sigma_W^2}\right]\right\}}, \quad D_s \geq \frac{\sigma_S^2}{\exp\left\{\mathbb{E}_F\left[\log\frac{|f|^2\left(P+\sigma_S^2+2\sqrt{(1-\zeta)P\sigma_S^2}\right)+\sigma_W^2}{\zeta P|f|^2+\sigma_W^2}\right]\right\}} \tag{34}$$

where Var$(V|S) = \sigma_V^2(1-\rho^2)$ *is the variance of* V *given* S *and* $0 \le \zeta \le 1$.

Proof: For the source distortion, we can write the following

$$\begin{aligned}
\frac{K}{2}\log\frac{\sigma_V^2}{D_v} &\overset{(a)}{\le} I(V^K;\hat{V}^K|F^K) \overset{(b)}{\le} I(V^K;\hat{V}^K|F^K) + I(V^K;S^K|\hat{V}^K,F^K) \\
&= I(V^K;\hat{V}^K,S^K|F^K) \overset{(c)}{=} I(V^K;S^K) + I(V^K;\hat{V}^K|S^K,F^K) \\
&\overset{(d)}{\le} \frac{K}{2}\log\frac{\sigma_V^2}{\text{Var}(V|S)} + I(V^K;Y^K|S^K,F^K) \\
&= \frac{K}{2}\log\frac{\sigma_V^2}{\text{Var}(V|S)} + h(Y^K|S^K,F^K) - h(W^K) \\
&\overset{(e)}{=} \frac{K}{2}\log\frac{\sigma_V^2}{\text{Var}(V|S)} + \frac{K}{2}\mathbb{E}_F\left[\log\frac{\zeta P|f|^2+\sigma_W^2}{\sigma_W^2}\right] \qquad (35)
\end{aligned}$$

where (a) follows from the rate-distortion theorem, (b) follows from the non-negativity of mutual information, (c) follows from the chain rule of mutual information and the fact that F^K is independent of (V^K,S^K), (d) holds by the data processing inequality and in (e) we used $h(Y^K|S^K,F^K) = \frac{K}{2}\mathbb{E}_F\left[\log\left(\zeta P|f|^2+\sigma_W^2\right)\right]$ for some $\zeta \in [0,\ 1]$; this can be proved from the fact that $\frac{K}{2}\log\sigma_W^2 = h(W^N) \le h(Y^K|S^K,F^K) \le h(F^K X^K + W^K|F^K) = \frac{K}{2}\mathbb{E}_F\left[\log\left(P|f|^2+\sigma_W^2\right)\right]$. Hence, there exists a $\zeta \in [0,\ 1]$ such that $h(Y^K|S^K,F^K) = \frac{K}{2}\mathbb{E}_F\left[\log\left(\zeta P|f|^2+\sigma_W^2\right)\right]$.

For the interference distortion, we have the following

$$\begin{aligned}
&\frac{K}{2}\log\frac{\sigma_S^2}{D_s} \overset{(a)}{\le} I(S^K;\hat{S}^K|F^K) \overset{(b)}{\le} I(S^K;Y^K|F^K) = h(Y^K|F^K) - h(Y^K|S^K,F^K) \\
&\overset{(c)}{\le} \sup_{X\in\mathcal{B}} \mathbb{E}_F\left[\frac{K}{2}\log 2\pi e(|f|^2(P+\sigma_S^2+2\mathbb{E}[SX])+\sigma_W^2)\right] \\
&-\mathbb{E}_F\left[\frac{K}{2}\log 2\pi e(\zeta P|f|^2+\sigma_W^2)\right] \\
&\overset{(d)}{=} \mathbb{E}_F\left[\frac{K}{2}\log\frac{|f|^2(P+\sigma_S^2+2\sqrt{(1-\zeta)P\sigma_S^2})+\sigma_W^2}{\zeta P|f|^2+\sigma_W^2}\right] \qquad (36)
\end{aligned}$$

where (a) follows from the rate-distortion theorem, (b) follows from data processing inequality for the mutual information, in (c) the set $\mathcal{B} = \left\{X : h(Y^K|S^K,F^K) = \mathbb{E}_F\left[\frac{K}{2}\log 2\pi e(\zeta P|f|^2+\sigma_W^2)\right]\right\}$ and the inequality in (c) holds from the fact that Gaussian maximizes differential entropy and $h(Y^K|S^K,F^K) = \frac{K}{2}\mathbb{E}_F\left[\log\left(\zeta P|f|^2+\sigma_W^2\right)\right]$ (as used in (35)). Note that the supremum over X in (c) happens when Y, S and W are jointly Gaussian given F (i.e., X^* is Gaussian). Now, we represent $X^* = N_\zeta^* + X_\zeta^*$, where $N_\zeta^* \sim \mathcal{N}(0,\zeta P)$ is independent of $X_\zeta^* \sim \mathcal{N}(0,(1-\zeta)P)$. Note that X_ζ^* is a function of S. Using this form of X^*, $h(Y^K|S^K,F^K) = \frac{K}{2}\mathbb{E}_F\left[\log\left(\zeta P|f|^2+\sigma_W^2\right)\right]$

still holds (i.e., $X^* \in \mathcal{B}$) and $\mathbb{E}[X^*S] = \mathbb{E}[X^*_\zeta S]$. Maximizing over X is equivalent to maximizing over $\mathbb{E}[XS]$; using Cauchy-Schwarz $\left(\mathbb{E}[X^*S] = \mathbb{E}[X^*_\zeta S] \leq \sqrt{\mathbb{E}[(X^*_\zeta)^2]\mathbb{E}[S^2]}\right)$ we get (d). ■

B. Proposed Hybrid Coding Scheme

The proposed scheme is composed of three layers as shown in Fig. 11. The first layer, which is purely analog, consumes an average power P_a and outputs a linear combination between the source and the interference $X_a^K = \sqrt{a_1}(\beta_{11}V^K + \beta_{12}S^K)$, where $\beta_{11}, \beta_{12} \in [-1 \quad 1]$ and $a_1 = \frac{P_a}{(\beta_{11}^2\sigma_V^2 + 2\beta_{11}\beta_{12}\rho\sigma_V\sigma_S + \beta_{12}^2\sigma_S^2)}$ is a gain factor related to power constraint P_a. The second layer employs a source-channel vector-quantizer (VQ) on the interference; the output of this layer is $X_q^K = \mu(S^K + U_q^K)$, where $\mu > 0$ is a gain related to the power constraint and samples in U_q^K follow a zero mean i.i.d. Gaussian that is independent of V and S and has a variance Q. A similar VQ encoder was used in [21] for the broadcast of bivariate sources and for the multiple access channel [36]. In what follows, we outline the encoding process of the VQ

- *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 the encoder and decoders.
- *Encoding:* The encoder searches for a codeword X_q^K in the codebook that is jointly typical with S^K. In case of success, the transmitter sends X_q^K.

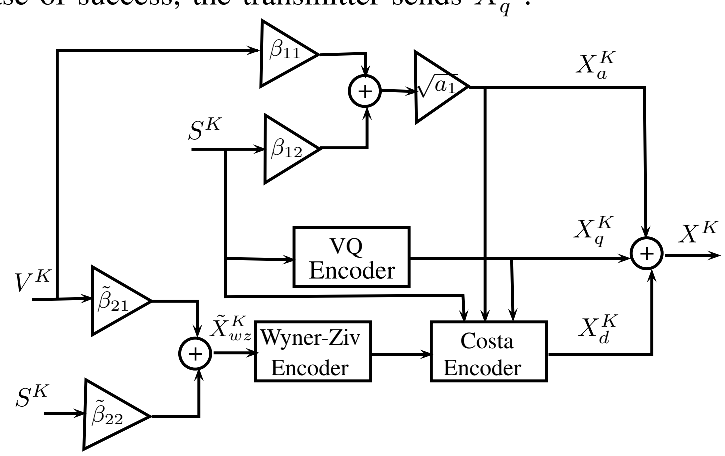

Fig. 11. Encoder structure.

The last layer encodes a linear combination between V^K and S^K, denoted by $\tilde{X}_{wz}^K$, using a Wyner-Ziv with rate R followed by a Costa coder. The Costa coder uses an average power of P_d and treats X_a^K, S^K and X_q^K as known interference. The linear combination $\tilde{X}_{wz}^K = \tilde{\beta}_{21}V^K + \tilde{\beta}_{22}S^K = \sqrt{a_2}(\beta_{21}V^K + \beta_{22}S^K)$, where $\beta_{21}, \beta_{22} \in [-1 \ 1]$ and $a_2 = \frac{P_d}{(\beta_{21}^2\sigma_V^2 + 2\beta_{21}\beta_{22}\rho\sigma_V\sigma_S + \beta_{22}^2\sigma_S^2)}$. The Wyner-Ziv encoder forms a random variable T^K as follows

$$T^K = \alpha_{wz}\tilde{X}_{wz}^K + B^K \tag{37}$$

where the samples in B^K are zero mean i.i.d. Gaussian, the parameter α_{wz} and the variance of B 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} bins with R 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. The Costa coder, that treats $\tilde{S}^K = X_a^K + X_q^K + S^K$ as known interference, forms the following auxiliary random variable $U_c^K = X_d^K + \alpha_c \tilde{S}^K$, where each sample in X_d^K is $\mathcal{N}(0, P_d)$ that is independent of the source and the interference and $0 \leq \alpha_c \leq 1$. The encoding process for the Costa coder can be described in a similar way as done before.

At the receiver side, as shown in Fig. 12, from the noisy received signal Y^K, the VQ decoder estimates X_q^K by searching for a codeword $X_q^K \in \mathcal{X}_q$ that is jointly typical with the received signal Y^K and F^K. Following the error analysis of [37], the error probability of decoding X_q^K goes to zero by choosing the rate R_q to satisfy $I(S; X_q) \leq R_q \leq I(X_q; Y, F)$, where

$$\begin{aligned}
I(S; X_q) &= h(X_q) - h(X_q|S) = \frac{1}{2}\log\frac{\sigma_S^2 + Q}{Q} \\
I(X_q; Y, F) &= I(X_q; F) + I(X_q; Y|F) = h(Y|F) - h(Y|X_q, F) \\
&= \mathbb{E}_F\left\{\frac{1}{2}\log 2\pi e\left(\mathbb{E}[Y^2]\right) - \frac{1}{2}\log 2\pi e\left(\mathbb{E}[Y^2] - \frac{\mathbb{E}[X_q Y]^2}{\mathbb{E}[X_q^2]}\right)\right\}.
\end{aligned} \tag{38}$$

The variance Q has to be chosen to satisfy the above rate constraint. Furthermore, to ensure the power constraint is satisfied we need μ to satisfy $P_a + \mu^2(\sigma_S^2 + Q) + 2\mu\mathbb{E}[SX_a] + P_d \leq P$. The Costa decoder then searches for a codeword U_c^K that is jointly typical with (Y^K, F^K). Since the received signal Y^K and the codewords X_q^K and U_c^K are correlated with $\tilde{X}_{wz}^K$, an LMMSE estimate of $\tilde{X}_{wz}^K$, denoted by $\hat{\tilde{X}}_{wz}^K$, can be obtained based on Y^K and the decoded codewords X_q^K and U_c^K. Mathematically, the estimate is given by $\hat{\tilde{X}}_{wz} = \Gamma_a \Lambda_a^{-1}[X_q \;\; U_c \;\; Y]^T$, where Λ_a is the covariance of $[X_q \;\; U_c \;\; Y]$ and Γ_a is the correlation vector between $\tilde{X}_{wz}$ and $[X_q \;\; U_c \;\; Y]$. The distortion in reconstructing $\tilde{X}_{wz}^K$ is then $D_a = \mathbb{E}_F\left[P_d - \Gamma_a \Lambda_a^{-1}\Gamma_a^T\right]$. Moreover, the Wyner-Ziv decoder estimates the codeword T^K by searching for a $T^K \in \mathcal{T}$ that is jointly typical with $\hat{\tilde{X}}_{wz}^K$. The error probability of decoding both codewords T^K and U_c^K vanishes as $K \to \infty$ if the coding rate of the Wyner-Ziv and the Costa coder is set to

$$R = \mathbb{E}_F \left[\frac{1}{2} \log \left(\frac{P_d[f^2(P_d + \sigma_{\tilde{S}}^2) + \sigma_W^2]}{P_d(\sigma_{\tilde{S}}^2) f^2 (1-\alpha_c)^2 + \sigma_W^2 (P_d + \alpha_c^2 \sigma_{\tilde{S}}^2)} \right) \right] \tag{39}$$

where $\sigma_{\tilde{S}}^2 = \mathbb{E}[(X_a + X_q + S)^2]$. A better estimate of $\tilde{X}_{wz}^K$ can be obtained by using the codeword T^K and $\hat{\tilde{X}}_{wz}^K$. The distortion in reconstructing $\tilde{X}_{wz}^K$ can be expressed as follows

$$\tilde{D} = \frac{D_a}{\exp\left\{ \mathbb{E}_F \left[\log \left(\frac{P_d[f^2(P_d + \sigma_{\tilde{S}}^2) + \sigma_W^2]}{P_d(\sigma_{\tilde{S}}^2) f^2 (1-\alpha_c)^2 + \sigma_W^2 (P_d + \alpha_c^2 \sigma_{\tilde{S}}^2)} \right) \right] \right\}}. \tag{40}$$

This distortion can be achieved using a linear MMSE estimate based on T^K, X_q^K and Y^K by choosing $\alpha_{wz} = \sqrt{1 - \frac{\tilde{D}}{D_a}}$ and $B \sim \mathcal{N}(0, \tilde{D})$ in (37).

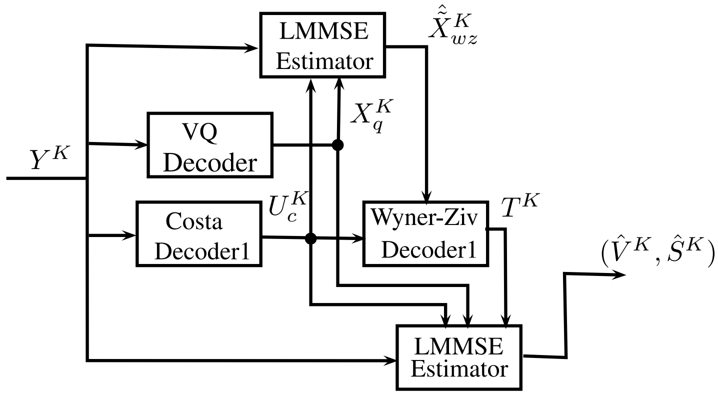

Fig. 12. Decoder structure.

After decoding T^K, X_q^K, a linear MMSE estimator is used to reconstruct the source and the interference signals. As a result, the distortion in decoding V^K and S^K are given as follows

$$D_v = \mathbb{E}_F \left[\sigma_V^2 - \Gamma_v \Lambda^{-1} \Gamma_v^T \right] \quad D_s = \mathbb{E}_F \left[\sigma_S^2 - \Gamma_s \Lambda^{-1} \Gamma_s^T \right] \tag{41}$$

where Λ is the covariance of $[X_q \;\; T \;\; Y]$, Γ_v is the correlation vector between V and $[X_q \;\; T \;\; Y]$ and Γ_s is the correlation vector between S and $[X_q \;\; T \;\; Y]$.

Remark 6 *Using a linear combination of the source and the interference $\tilde{X}_{wz}^K$ instead of just the source V^K as an input to the Wyner-Ziv encoder in Fig. 11 is shown to be beneficial in some parts of the source-interference distortion region. However, quantizing a linear combination of the source and the interference by the VQ encoder (instead of just S^K as done in Fig. 11) does not seem to give any improvement.*

Remark 7 *For the AWGN channel with $\rho = 0$, using the source itself instead of $\tilde{X}_{wz}^K$ as input to the Wyner-Ziv encoder, shutting down the second layer and setting $\beta_{11} = 0$ in X_a^K give the best* *possible performance; the inner bound in such case coincides with the outer bound, hence the scheme is optimal. This result is analogous to the optimality result of the rate-state-distortion for the transmission of a finite discrete source over a Gaussian state interference derived in [29].*

C. Numerical Results

We consider a source-interference pairs that are transmitted over a Rayleigh fading channel ($\mathbb{E}[F^2] = 1$) with Gaussian interference and power constraint $P = 1$; the CSNR level is set to 10 dB. For reference, we adapt the scheme of [29] to our scenario. Recall that the source and the interference are jointly Gaussian, hence $V^K = \rho \frac{\sigma_V}{\sigma_S} S^K + N_\rho^K$, where samples in N_ρ^K are i.i.d. Gaussian with variance $\sigma_{N_\rho}^2 = (1-\rho^2)\sigma_V^2$ and independent of S^K. Now if we quantize N_ρ^K into digital data, the setup becomes similar to the one considered in [29]; hence the encoding is done by allocating a portion of the power, denoted by P_s, to transmit S^K and the remaining power $P_d = (P - P_s)$ is used to communicate the digitized N_ρ^K using the (generalized) Costa coder. The received signal of such scheme is given by $Y^K = F^K \left(\sqrt{\frac{P_s}{\sigma_S^2}} S^K + X_d^K + S^K \right) + W^K$, where X_d^K denotes the output of the digital part that communicates N_ρ^K. An estimate of S^K is obtained by applying a LMMSE estimator on the received signal; the distortion from reconstructing V^K is equal to the sum of the distortions from estimating $\rho \frac{\sigma_V}{\sigma_S} S^K$ and N_ρ^K. Mathematically, the distortion region of such reference scheme can be expressed as follows

$$
\begin{aligned}
D_s &= \mathbb{E}_F \left[\sigma_S^2 - \frac{\mathbb{E}[SY]^2}{\mathbb{E}[Y^2]} \right] = \mathbb{E}_F \left[\sigma_S^2 - \frac{f^2(\sqrt{P_s}\sigma_S + \sigma_S^2)^2}{f^2(P + \sigma_S^2 + 2\sqrt{P_s}\sigma_S) + \sigma_W^2} \right], \\
D_v &= \rho^2 \frac{\sigma_V^2}{\sigma_S^2} D_s + \frac{\sigma_{N_\rho}^2}{\exp\left\{ \mathbb{E}_F \left[\log\left(\frac{P_d[f^2(P_d + \sigma_{\tilde{S}}^2) + \sigma_W^2]}{P_d(\sigma_{\tilde{S}}^2) f^2 (1-\alpha_c)^2 + \sigma_W^2 (P_d + \alpha_c^2 \sigma_{\tilde{S}}^2)} \right) \right] \right\}}
\end{aligned}
\tag{42}
$$

where $\sigma_{\tilde{S}}^2 = \mathbb{E}[(\sqrt{P_s/\sigma_S^2} S + S)^2]$ and the parameter α_c is related to the Costa coder. To evaluate the performance, we plot the outer bound (given by (34)) and the inner bounds (the achievable distortion region) of the proposed hybrid coding (given by (41)) and the adapted scheme of [29]. Fig. 13, which considers the AWGN channel, shows the distortion region of the source-interference pair for $\rho = 0.8$ and $\sigma_S^2 = 0.5$. We can notice that the hybrid coding scheme is very close to the outer bound and outperforms the scheme of [29]. Moreover, the use of the VQ layer is shown to be beneficial under certain system settings. Fig. 14, which considers the fading channel, shows the distortion region of the source-interference pair for $\rho = 0.8$ and $\sigma_S^2 = 1$. The hybrid coding scheme performs relatively close to the outer bound.

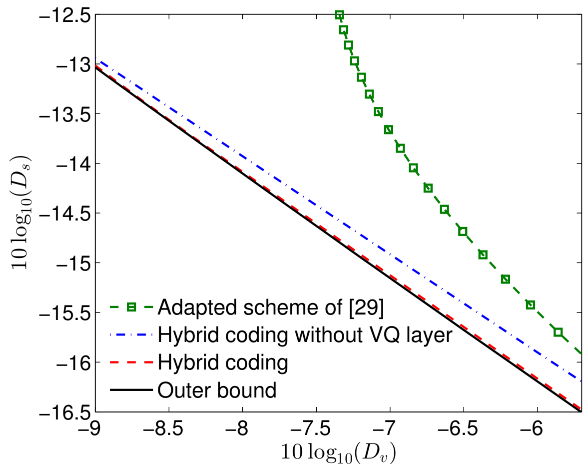

Fig. 13. Distortion region for hybrid coding scheme over the AWGN channel. This graph is made for $\sigma_V^2 = 1$, $P = 1$, $\sigma_S^2 = 0.5$ and $\rho = 0.8$.

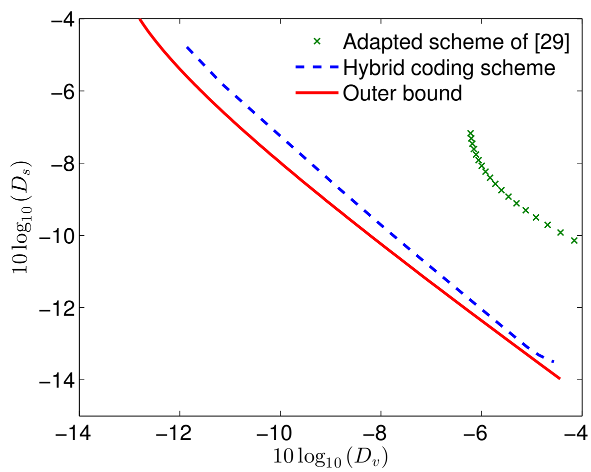

Fig. 14. Distortion region for hybrid coding scheme over the fading channel. This graph is made for $\sigma_V^2 = 1$, $P = 1$, $\sigma_S^2 = 1$, $\rho = 0.8$ and $\mathbb{E}[F^2] = 1$.

VI. Summary and Conclusions

In this paper, we considered the problem of reliable transmission of a Gaussian sources over Rayleigh fading channels with correlated interference under unequal source-channel bandwidth. Inner and outer bounds on the system's distortion are derived. The outer bound is derived by assuming additional knowledge at the decoder side; while the inner bound is found by analyzing the achievable distortion region of the proposed hybrid coding scheme. Numerical results show that the proposed schemes perform close to the derived outer bound and to be robust to channel noise mismatch. As an application of the proposed schemes, we derive inner and outer bounds on the source-channel-state distortion region for the fading interference channel; in this case, the receiver is interested in estimating both source and interference. Our setting contains several interesting limiting cases. In the absence of fading and/or correlation and for some source-channel bandwidths, our setting resorts to the scenarios considered in [20], [26], [29].

Acknowledgment

The authors sincerely thank the anonymous reviewers for their helpful comments.

References

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

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

[3] 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.

[4] 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.

[5] 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. Comm.*, vol. 60, no. 10, pp. 3091–3102, Oct 2012.

[6] 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.

[7] E. Akyol, K. Rose, and T. Ramstad, "Optimal mappings for joint source channel coding," in *IEEE Inform. Theory Works.*, Cairo, Egypt, 2010.

[8] E. Akyol, K. Rose, and K. Viswanatha, "On zero-delay source-channel coding," *IEEE Trans. Inform. Theory*, in review.

[9] M. Kleiner and B. Rimoldi, "Asymptotically optimal joint source-channel coding with minimal delay," in *Proc. IEEE Global Telecommunications Conference*, Honolulu, HI, 2009.

[10] X. Chen and E. Tuncel, "Zero-delay joint source-channel coding using hybrid digital-analog schemes in the Wyner-Ziv setting," *IEEE Trans. Communications*, vol. 62, no. 2, pp. 726–735, Feb 2014.

[11] A. Abou Saleh, F. Alajaji, and W.-Y. Chan, "Power-constrained bandwidth-reduction source-channel mappings for fading channels," in *Proc. of the 26th Bien. Symp. on Comm.*, Kingston, ON, May 2012.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

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

[19] 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.

[20] M. Varasteh and H. Behroozi, “Optimal HDA schemes for transmission of a Gaussian source over a Gaussian channel with bandwidth compression in the presence of an interference,” *IEEE Trans. Signal Processing*, vol. 60, no. 4, pp. 2081–2085, April 2012.

[21] 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.

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

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

[24] 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.

[25] 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.

[26] 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.

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

[28] A. Abou Saleh, W.-Y. Chan, and F. Alajaji, “Low and high-delay source-channel coding with bandwidth expansion and correlated interference,” in *Proc. of the 13th Canadian Workshop on Information Theory*, Toronto, ON, June 2013.

[29] 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.

[30] 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.

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

[32] 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.

[33] K. M. Abadir and J. R. Magnus, *Matrix Algebra*. New York: cambridge University Pressr, 2005.

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

[35] W. Zhang, S. Kotagiri, and J. Laneman, “Writing on dirty paper with resizing and its application to quasi-static fading broadcast channels,” in *Proc. IEEE Int. Symp. on Inform. Theory*, Nice, France, June 2007.

[36] 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.

[37] S. H. Lim, P. Minero, and Y.-H. Kim, “Lossy communication of correlated sources over multiple access channels,” in *Proceedings of the 48th Annual Allerton Conf. on Comm., Control and Computing*, Monticello, IL, Sep 2010.