Lossy Transmission over Slow-Fading AWGN Channels: a Comparison of Progressive, Superposition and Hybrid Approaches

Stefania Sesia
Philips Semiconductors
Sophia Antipolis, France
Email: Stefania.Sesia@philips.com

Giuseppe Caire
Eurecom Institut
Sophia Antipolis, France
Email: Giuseppe.Caire@eurecom.fr

Guillaume Vivier
Centre de Recherche Motorola Labs
Gif-sur-Yvette, France
Email: Guillaume.Vivier@motorola.com

***Abstract*— We consider some known techniques for transmitting an analog source under the quadratic distortion criterion over a slowly-varying fading AWGN channel. We examine the case where the source is actually *compressed*, i.e., the source bandwidth is larger than the channel bandwidth. This makes the problem non-trivial. We optimize progressive transmission, superposition coding and a analog-digital hybrid approach in order to achieve minimal average distortion. Then, we compare the resulting optimized systems in terms of average and instantaneous reconstructed signal to noise ratio.**

I. Introduction

Modern telecommunications often involve the transmission of *analog* sources over digital channels. Paramount examples are digital TV and audio broadcasting (DTV, DAB) and transmission of pictures over wireless radio channels in 3G mobile phones. In this case, bit-error probability at the output of the channel decoder is no longer a good measure of performance. On the contrary, the end-to-end distortion is more representative of the quality of transmission. In some lucky sporadic cases, such as a Gaussian source over a Gaussian channel with the same bandwidth, it is well-known that "analog transmission"[1] is optimal [1]. Nevertheless, the main interest in using digital techniques is to improve spectral efficiency, i.e., to maximize the number of source samples per channel use. In order to achieve high spectral efficiency, the source must be compressed. Digital systems can be optimized to achieve asymptotically optimal performance at a given target SNR^*, but they perform poorly for SNR below this target and they do not take advantage of better channel conditions when the actual SNR is above the target SNR. This is called "threshold effect". On the contrary, analog schemes improve gradually their reconstruction signal to noise ratio as the channel signal to noise ratio improves. Traditionally, digital schemes are designed based on Shannon's separation principle that states that no loss in performance is incurred when designing source and channel coding scheme separately [2]. However, this does not take into consideration complexity and delay and it does not hold for non-ergodic scenario such as the slowly fading AWGN considered here.

In this paper we consider the simplest possible scenario of this kind, which is, nevertheless, not yet fully solved. We consider a Gaussian i.i.d. source with bandwidth W_s that has to be transmitted over a band-limited channel with bandwidth W_c under the end-to-end quadratic distortion criterion. As motivated before, we assume spectral efficiency $\eta \triangleq W_s/W_c > 1$. We consider the so-called block-fading (BF) AWGN channel, for which the channel gain is random but constant over the duration of a codeword. The coding block length is assumed large enough such that any rate below the instantaneous channel capacity for the given fading realization can be decoded with negligible probability of error, while any rate above the instantaneous channel capacity yields probability of error close to 1. The BF-AWGN channel is a useful mathematical abstraction that models very slowly-varying fading channels as for example stationary terminals such as TV receivers, or the path loss determined by the distance to the base station in a mobile cellular communications. In these cases, the fading changes much more slowly that the coding delay and the channel behaves non-ergodically (see [3] for a thorough discussion). Moreover, the BF-AWGN channel under the assumption made here that the transmitter is not informed of the channel fading (but it knows its statistics), may also model a Gaussian broadcast channel [4] with $K \to \infty$ users, such that the empirical distribution of the users SNRs converges almost everywhere to the fading cdf. For this scenario, we optimize and compare three well-known strategies: progressive transmission, superposition coding and a hybrid digital/analog scheme (HDA). We optimize these schemes by minimizing the average end-to-end distortion for given transmit power Γ and fading statistics (assuming a *continuous* pdf $f_A(z)$ and cdf $F_A(z)$).

"Progressive transmission" is an approach widely used in practice. The source is split into independent streams mapped onto channel codewords with different coding rate and possibly different energy per channel symbol. The codewords are sent through the channel in sequence (time-sharing). In [5]

[1]By analog transmission we indicate that the source is sent directly through the channel up to some scaling (in order to meet the transmit power constraint). This can be seen as regular analog AM.

the optimization analysis is carried out for Binary Symmetric Channel (BSC) and Binary Erasure Channel (BEC). Using this principle, [6] characterizes an achievable average distortion region for the broadcasting of a common source. The splitting of the source is represented by an ideal successive refinement source encoder [7] that provides independent levels of information.

A *broadcast approach* to the BF-AWGN channel was proposed and analyzed in [8] in order to maximize the average transmission rate. It consists of splitting the information message into $L \to \infty$ parallel streams and mapping each stream onto a layer of a superposition coding scheme. Each layer is modulated with a power level and optimized such that the average successfully received rate is maximized. Following the approach of [8] yields unmanageable expressions due to the fact that the distortion is a non-linear function of the rate and the elegant solution of [8] does not apply. A more practical approach for the multiresolution/superposition framework is given in [9], where the authors give an algorithm that jointly optimizes the multiresolution source codebook, the (multilevel) modulation constellation and the decoding strategy by minimizing the overall distortion.

Finally, in [10] some hybrid digital/analog schemes based on bandwidth and power splitting are proposed. These schemes generally outperform the superposition or time-sharing approach in terms of the achievable distortion region for a two-user broadcast channel, thanks to the presence of the analog signal. In particular, for the case $\eta > 1$, [10] proposes a scheme formed by two layers, analog and digital, modulated with different power levels and superimposed. While the authors optimize the system in order to achieve asymptotically optimal performance for a particular target SNR, here, we optimize the scheme in order to minimize the average distortion for a given fading distribution.

II. Progressive Transmission

A multiresolution source encoder splits the source into L independent layers of information. Each level has a source coding rate r_s bits/source sample and it is mapped onto a codeword belonging to a channel code $\mathcal{C}_i^l$, modulated with a different power coefficient. We consider ideal source and channel codes characterized by their rate-SNR pair (r_c, τ) (channel codes) and their rate-distortion pair (r_s, D) (source codes). Under the hypothesis of a Gaussian source and quadratic distortion measure, the source is successively refinable, i.e the distortion-rate function is achieved at each level $D_\ell = 2^{2r_s\ell}$ and $D_0 = 1$. Codeword i has coding rate $r_i = \frac{k}{n_i}$, where k is the number of bits output by the i-th layer of the multiresolution source coder and n_i is the blocklength and the SNR threshold is τ_i. Figure 1 illustrates this scheme. The spectral efficiency is given by

$$\eta = \frac{1}{r_s \sum_{i=1}^{L} \frac{1}{r_i}} \quad \Rightarrow \quad \sum_{i=1}^{L} \frac{1}{r_i} = \frac{1}{\eta r_s}$$

Define γ_ℓ as the energy per channel symbol used to transmit the $\ell-$th codeword. The total power Γ can be computed as

$$\Gamma = \sum_{\ell=1}^{L} \frac{n_\ell}{\sum_{i=1}^{L} n_i} \gamma_\ell = \eta r_s \sum_{\ell=1}^{L} \frac{\gamma_\ell}{r_\ell}$$

Define a set of fading thresholds $0 < a_1 < \cdots < a_L$ (where $a_{L+1} = \infty$) such that layers up to ℓ can be decoded if $A \in [a_\ell, a_{\ell+1}]$. Assuming ideal Gaussian channel codes, the ℓ-th code rate is given by $r_\ell = \log_2(1 + a_\ell \gamma_\ell)$. Call now $z_\ell = \frac{1}{r_\ell}$ and $y_\ell = \frac{\gamma_\ell}{r_\ell}$. It follows that

$$a_\ell = \frac{\left(2^{1/z_\ell} - 1\right) z_\ell}{y_\ell} \tag{1}$$

We wish to solve the following minimization problem

$$\min_{\mathbf{z},\mathbf{y}} D_{\text{ave}} \quad s.t. \ \sum_{i=1}^{L} z_\ell = \frac{1}{\eta r_s}, ; \quad \sum_{i=1}^{L} y_\ell = \frac{\Gamma}{\eta r_s} \tag{2}$$

where $D_{\text{ave}}(r_s, \mathbf{z}, \mathbf{y})$ has the following expression

$$D_{\text{ave}}(r_s, \mathbf{z}, \mathbf{y}) = F_A(a_1) + \sum_{\ell=1}^{L} D_\ell \left(F_A(a_{\ell+1}) - F_A(a_\ell)\right) \tag{3}$$

with a_ℓ defined in (1) and r_s is fixed.

The associated Lagrangian functional Φ is given by

$$\Phi = D_{\text{ave}}(r_s, \mathbf{z}, \mathbf{y}) + \lambda \sum_{i=1}^{L} z_i + \rho \sum_{i=1}^{L} y_i$$

For the Kuhn Tucker's conditions it follows that the partial derivative with respect to z_ℓ and y_ℓ has to be greater or equal to 0,

$$\frac{\partial \Phi}{\partial z_\ell} = \Delta D_\ell f_A(a_\ell) \left(\frac{\left(2^{1/z_\ell} - 1\right) z_\ell - 2^{1/z_\ell} \ln 2}{z_\ell y_\ell} \right) + \lambda \geq 0 \tag{4}$$

$$\frac{\partial \Phi}{\partial y_\ell} = \Delta D_\ell f_A(a_\ell) \left(- \frac{z_\ell \left(2^{1/z_\ell} - 1\right)}{y_\ell^2} \right) + \rho \geq 0 \tag{5}$$

where $\Delta D_\ell = D_{\ell-1} - D_\ell$. y_ℓ can be found as a function of ρ, λ and z_ℓ

$$y_\ell = \frac{1}{\mu} \frac{\left(2^{1/z_\ell} - 1\right) z_\ell^2}{2^{1/z_\ell} \ln 2 - \left(2^{1/z_\ell} - 1\right) z_\ell} \tag{6}$$

where μ is defined as $\mu \triangleq \frac{\rho}{\lambda}$. The value z_ℓ is then obtained as

$$-\Delta D_\ell g\left(z_\ell, \mu\right) + \lambda = 0 \tag{7}$$

where $g\left(z_\ell, \mu\right)$ is defined as

$$g\left(z_\ell, \mu\right) \triangleq f_A \left(\mu \frac{2^{1/z_\ell} \ln 2 - \left(2^{1/z_\ell} - 1\right) z_\ell}{z_\ell} \right) \cdot \left(\mu \frac{\left(2^{1/z_\ell} \ln 2 - \left(2^{1/z_\ell} - 1\right) z_\ell\right)^2}{z_\ell^3 \left(2^{1/z_\ell} - 1\right)} \right) \tag{8}$$

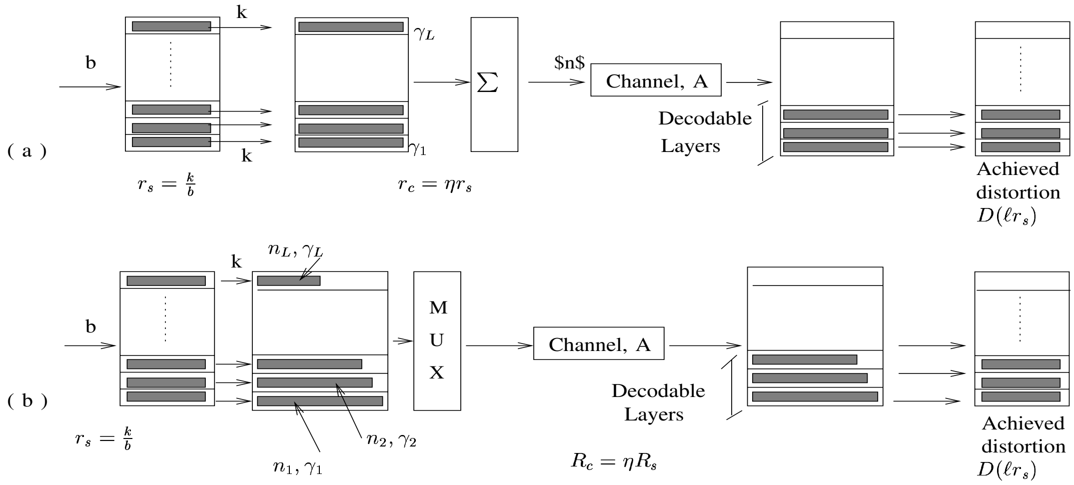

Fig. 1. (a) Progressive transmission-based scheme, (b) Superposition-based scheme.

III. Superposition-based Transmission Strategies

We consider now a superposition-based approach where each level is mapped onto a codeword of "a basic channel code" $\mathcal{C}'$ modulated at different power levels. The codewords are superimposed and sent through the channel. Each layer has source coding rate r_s bits/source sample and channel coding rate r_c bit/channel uses, so that $\eta = r_c/r_s$. Figure 1 shows the block diagram of the superposition scheme. The SNR threshold of the mother code $\mathcal{C}'$ is $\tau = 2^{2r_c} - 1 = 2^{2\eta r_s} - 1$. The transmitted superposition codeword is given by $\mathbf{x} = \sum_{\ell=1}^{L} \sqrt{\gamma_\ell}\mathbf{c}'_\ell$ where γ_ℓ and $\mathbf{c}'_\ell$ are the power level and the codeword of $\mathcal{C}'$ associated to level ℓ, respectively. Define again the set of fading thresholds $0 < a_1 < \cdots < a_L$ (where $a_{L+1} = \infty$) such that layers up to ℓ can be decoded if $A \in [a_\ell, a_{\ell+1}]$. The condition for successive decodability of the superposition code up to layer ℓ is given by

$$\frac{a_\ell \gamma_\ell}{1 + a_\ell \sum_{j=\ell+1}^{L} \gamma_j} \geq \tau \tag{9}$$

The levels a_ℓ are uniquely defined by the power levels γ_ℓ by imposing the constraint (9) with equality,

$$a_\ell = \frac{\tau}{\gamma_\ell - \tau \sum_{j=\ell+1}^{L} \gamma_j} \tag{10}$$

Conversely, the γ_ℓ's can be expressed in terms of the a_ℓ's by solving the triangular linear system $a_\ell\gamma_\ell - \tau a_\ell \sum_{j=\ell+1}^{L} \gamma_j = \tau$ for all $\ell = 1, \ldots, L$ which yields

$$\gamma_\ell = \tau x_\ell + \tau^2 x_{\ell+1} + \sum_{j=\ell+2}^{L} \tau^2 x_j \left(1+\tau\right)^{j-\ell-1} \tag{11}$$

where $x_\ell \triangleq \frac{1}{a_\ell}$. We wish to minimize the average distortion defined in (3) with respect to $\{\gamma_1, \ldots, \gamma_L\}$ subject to the constraint $\sum_\ell \gamma_\ell = \Gamma$. The associated Lagrangian functional is

$$\Phi = D_{\text{av}}(r_s, \gamma_1, \ldots, \gamma_L) + \lambda \sum_{\ell=1}^{L} \gamma_\ell \tag{12}$$

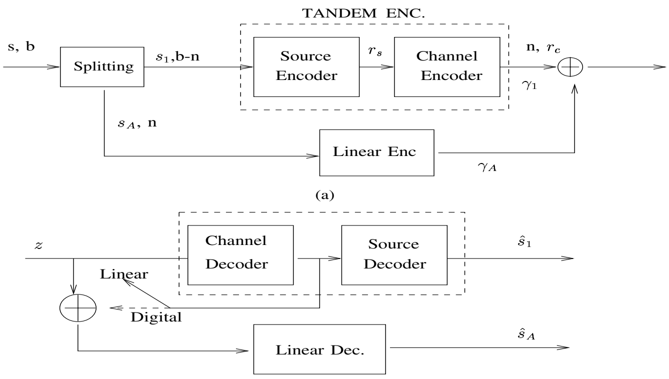

Fig. 2. Hybrid digital-analog scheme.

Kuhn-Tucker's conditions yield

$$\frac{\partial \Phi}{\partial \gamma_\ell} = \sum_{j=1}^{\ell-1} \Delta D_j \frac{1}{x_j^2} f_A\left(\frac{1}{x_j}\right) - \Delta D_\ell \frac{1}{\tau}\frac{1}{x_\ell^2} f_A\left(\frac{1}{x_\ell}\right) + \lambda \geq 0 \tag{13}$$

With the substitution $w_\ell = \frac{1}{x_\ell^2} f_A\left(\frac{1}{x_\ell}\right)$, the system given by $\frac{\partial\Phi}{\partial\gamma_\ell} = 0$ is linear and lower triangular, and the solution is given by

$$w_\ell = \frac{\lambda\tau\left(1+\tau\right)^{\ell-1}}{D_{\ell-1} - D_\ell} = \lambda \mathcal{G}_\ell \tag{14}$$

where we have defined $\mathcal{G}_\ell \triangleq \frac{\tau(1+\tau)^{\ell-1}}{D_{\ell-1}-D_\ell}$.

Finally, the derivative is greater or equal to zero if

$$-\frac{\frac{1}{x_\ell^2} f_A\left(\frac{1}{x_\ell}\right)}{\mathcal{G}_\ell} + \lambda \geq 0 \tag{15}$$

IV. Hybrid Analog/Digital Schemes

The scheme of the encoder and decoder is shown in figure 2. The bandwidth of the source W_s is split so that $W_{s,A} = W_c$ dimensions are uncoded and $W_s - W_{s,A}$ are sent to the tandem encoder with rate $\eta - 1$ source symbol per channel use. 'Tandem encoder' refers to source/channel codes independently designed [10]. The two signals are modulated by a different power coefficient, superimposed and sent through the channel. In [10] the system is designed to give asymptotically optimal performances for a particular value of $\mathsf{SNR} = \mathsf{SNR}^{\star}$. This is called a 'nearly robust' scheme. This yields a particular power splitting between the digital and the analog part. Here, we find that the power allocation that minimizes the average distortion does not coincides with that of [10], and does not yield an optimal distortion point at any SNR. The "analog code" is a linear coder/decoder with coefficients that minimize the mean squared error.

In [10] the authors consider a matched tandem encoder. The encoder is said to be matched if the channel input is a scaled version of the first n components of the source code reconstruction point, where n is the blocklength. The advantage is that, when the SNR is low, the tandem decoder is replaced by a linear decoder that estimates the source symbols. In the following we derive the optimality conditions for both the matched and non-matched tandem encoder.

Let us consider first the non-matched scheme. This scheme has only one digital layer so we will define only one fading threshold a_1 s.t. $0 < a_1 < +\infty$. Call γ_1 the power allocated to the digital layer and γ_A the power allocated to the analog layer, such that $\gamma_1 + \gamma_A = \Gamma$. The SNR threshold of the channel code is $\tau = 2^{2r_c} - 1 = 2^{2(\eta-1)r_s} - 1$ and the condition for successful decodability is given by

$$\frac{a_1\gamma_1}{1 + a_1\gamma_A} \geq \tau \tag{16}$$

where the analog layer is treated as additional noise by the digital decoder. By imposing the equality in (16) and substituting the total power constraint, we obtain $\gamma_A = \frac{\Gamma}{(1+\tau)} - \frac{\tau}{a_1(1+\tau)}$. Note that for a given fading value a, if $a < a_1$ the digital signal cannot be decoded correctly, i.e the output of the channel encoder acts as noise for the linear estimator of the analog layer. The distortion due to the analog layer is given by

$$D_A(r_s, a_1) = \begin{cases} \frac{1}{1+a\gamma_A} & \text{if } a \geq a_1 \\ 1 - \frac{a\gamma_A}{1+a\Gamma} & \text{else} \end{cases} \tag{17}$$

The average distortion can be written as

$$D_{\text{ave}}(r_s, a_1) = \frac{\eta - 1}{\eta}\left(D_0 F_A(a_1) + D_1(1 - F_A(a_1))\right) + \frac{1}{\eta} \left[F_A(a_1) - \int_0^{a_1} \frac{a\gamma_A f_A(a)da}{1 + a\Gamma} + \int_{a_1}^{\infty} \frac{f_A(a)da}{1 + a\gamma_A}\right] \tag{18}$$

where $D_0 = 1$. The result of the unconstrained minimization of (18) is obtained by finding a_1 solution of

$$f_A(a_1)(1 - D_1\frac{\eta - 1}{\eta}) - \frac{\gamma_A a_1 f_A(a_1)}{\eta(1 + a_1\Gamma)} - \frac{f_A(a_1)}{\eta(1 + a_1\gamma_A)} + \frac{\tau}{\eta(1+\tau)a_1^2}\left[\int_0^{a_1} \frac{af_A(a)da}{\Gamma(1 + a\Gamma)} - \int_{a_1}^{\infty} \frac{af_A(a)da}{(1 + a\gamma_A)^2}\right] = 0 \tag{19}$$

The matched tandem encoder is used to lower the distortion when $a < a_1$, i.e. when the digital decoder can not recover the information with arbitrarily small error probability. It this case, it can be shown [10] that the distortion of the digital part is given by

$$D_D(r_s, a_1) = \begin{cases} D_1 + \frac{1 - D_1}{1 + \frac{a\gamma_1}{1 + a\gamma_A}} & \text{if } a < a_1 \\ D_1 & \text{if } a \geq a_1 \end{cases} \tag{20}$$

The distortion due to the analog layer is still given by (17), and the average distortion is given by (18) where, now, D_0 is equal to the right end side of (20). The value of a_1 that yields minimum average distortion can be found by solving the following equation

$$f_A(a_1)(1 - \frac{\eta - 1}{\eta}D_1) - \frac{a_1 f_A(a_1)}{1 + a_1\Gamma}\mathcal{P}(\gamma_A) - \frac{f_A(a_1)}{\eta(1 + a_1\gamma_A)} - \mathcal{P}'(\gamma_A)\int_0^{a_1} \frac{af_A(a)da}{1 + a\Gamma} - \frac{\tau}{\eta a_1^2(1+\tau)}\int_{a_1}^{\infty} \frac{af_A(a)da}{(1 + a\gamma_A)^2} = 0 \tag{21}$$

where $\mathcal{P}(\gamma_A) \triangleq \frac{1}{\eta}\gamma_A + (1-D)\frac{\eta-1}{\eta}(\Gamma - \gamma_A)$, and $\mathcal{P}' = \frac{\partial \mathcal{P}(\gamma_A)}{\partial \gamma_A}$.

V. Results

We consider Rayleigh fading so that the pdf of the fading power gain in given by $f_A(a) = \exp(-a)$. The spectral efficiency η is fixed to 1.5 source symbol per channel use. Note that in all the above systems r_s is left as a design parameter and numerical optimization w.r.t r_s is further carried out. For the superposition strategy by letting L arbitrarily large with r_s arbitrarily small our numerical computable solution will approach arbitrarily closely the optimal solution of [8] when the average distortion is minimized instead of maximizing the average rate. For the progressive transmission approach, the optimal performances are, as well, obtained for $r_s \to \infty$. For the HDA scheme however the optimal r_s is a fixed non vanishing value $r_s^{\star}$. For comparison, a system based on the separation principle [11] that transmits a single layer is considered. This can be regarded as the baseline system representative of today's technology, such as terrestrial and satellite DTV, or DAB. We consider the minimization of average distortion with respect to the source coding rate r_s. The average distortion of the one-layer digital system is given by

$$D_{\text{sep}} = F_A\left(\frac{2^{\eta r_s} - 1}{\Gamma}\right) + 2^{-2r_s}\left[1 - F_A\left(\frac{2^{\eta r_s} - 1}{\Gamma}\right)\right] \tag{22}$$

For Rayleigh fading, the optimal value $r_s^{\star}$ is given by the solution of

$$\frac{\eta}{\Gamma}x^{\eta} - \frac{1}{x^2}\left(2 + \frac{\eta}{\Gamma}x^{\eta}\right) = 0$$

with $x = 2^{r_s^{\star}}$. Figure 3 shows the results in terms of RSNR defined as $10\log_{10}\frac{1}{D_{\text{opt}}}$ where D_{opt} is the optimal average distortion vs the average channel signal to noise ratio Γ. In the figure the average performances of progressive, superposition, and the non-matched/matched HDA ('HDA_D' in the figure) schemes are compared. Also plotted are the schemes based on the separation theorem and the nearly robust HDA of [10] ('HDA_{NR}' in the figure). For the progressive-based and superposition-based schemes, when vanishing r_s is optimal we plot the average performances for a r_s "small enough", $r_s = 1/20$. For practically small rate the gap from the optimal performances becomes negligible. For fixed r_s the solution of the minimization problem for the progressive transmission and the superposition approach gives also the optimal number of layer. Reducing r_s increases the optimal number layer and since a small enough value of r_s let achieve optimal performances, practically there is no need to make too many layers. For the separated approach and for the HDA the value of r_s is fixed to $r_s^\star$. As expected, that the matched/non-matched HDA_D outperform all the fully digital schemes, thanks to the presence of the analog layer. Finally, note that there is practically no difference between the separation theorem based approach and the progressive transmission in terms of average distortion. Figure 4 shows the performances in terms of RSNR vs instantaneous SNR,i.e. $a\Gamma$, for average SNR set to $\Gamma = 20$dB. In this figure we add, as a comparison, the Shannon limit given by $D_{Sh} = \frac{1}{(1+a\Gamma)^{1/\eta}}$. As announced before the enhancement of the performances due to the matched encoder is small and concentrated in a range $\mathsf{SNR} < \mathsf{SNR}^\star = a_1\Gamma$ because in that range the tandem decoder becomes a linear decoder that estimates the symbols. Moreover the construction of such codes is not straightforward. The HDA_D schemes and the superposition scheme gives smooth performances for a wide range of SNR providing more graceful degradation under mismatched channel condition compared to the separation theorem based scheme and the progressive approach . Finally, most of the gain in RSNR of the HDA is due to the presence of the linear encoder. A key issue is the practical construction of source/channel codes for the hybrid schemes. Of course, a joint source-channel approach has to be considered in order to design a scheme robust to channel errors and that performs as close as possible to the theoretical limits. Results on code construction for the HDA scheme can be found in [12], [13].

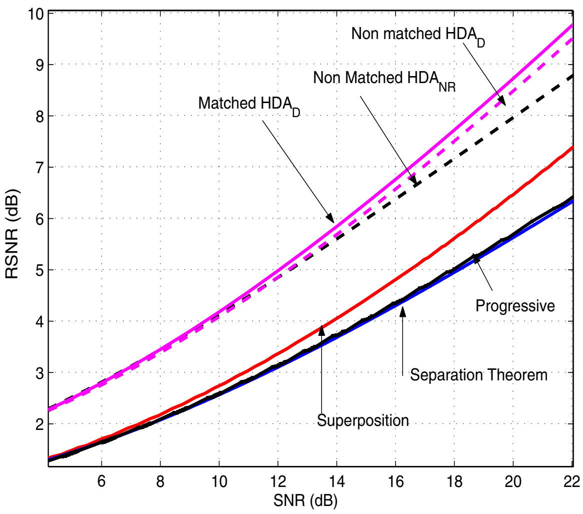

Fig. 3. RSNR vs channel average SNR (Γ).

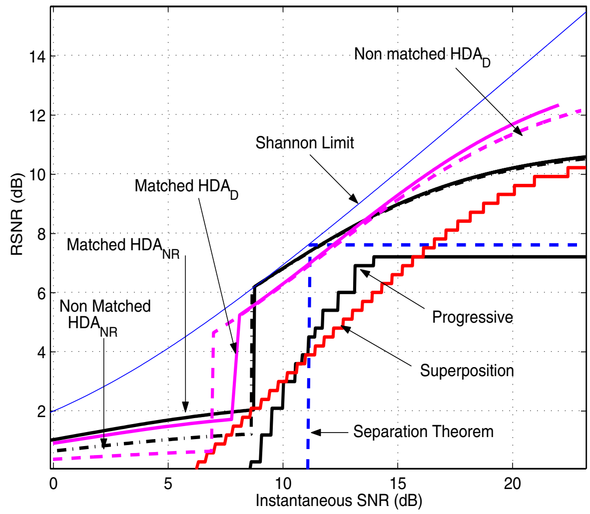

Fig. 4. RSNR vs channel instantaneous SNR for $\Gamma = 20$ dB.

ACKNOWLEDGMENT

This work has been done under a CIFRE project supported by Motorola Labs, Paris, France.

REFERENCES

[1] M. Gastpar, B. Rimoldi and M. Vetterli, "To code or not to code," in *IEEE Information Theory Symposium on Information Theory*, p. 236, June 2000.

[2] C. E. Shannon, "Communication in the presence of noise," *in Proc. IRE.*, Vol. 37, pp. 10–21, January 1949.

[3] E. Biglieri, J. Proakis and S. Shamai, "Fading channels: information theoretic and communication aspects," *IEEE Transactions on Information Theory*, Vol. 44, pp. 2619–2692, October 1996.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & sons, 1991.

[5] V. Chande and N. Favardin, "Progressive transmission of images over memoryless noisy channels,"*IEEE Journal on Selected Areas in Communications*, Vol. 18, pp. 850–860, June 2000.

[6] N. Sarshar and X. Wu, "Broadcasting with fidelity criteria," in *Information Theory Workshop (ITW)*, October 24-29 2004.

[7] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice Hall, 1971

[8] S. Shamai and A. Steiner, "A Broadcast approach for a single-user slowly fading MIMO channel," *IEEE Transactions on Information Theory*, Vol. 49, pp. 2617–2635, October 2003.

[9] I. Kozintsev and K. Ramchandran, "Robust image transmission over energy-constrained time varying channels using multiresolution joint source-channel coding," *IEEE Transactions on Signal Processing* Vol. 46, pp. 1012–1026, April 1998.

[10] U. Mittal and N. Phamdo, "Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications," *IEEE Transactions on Information Theory*, Vol. 48, pp. 1082–1102, May 2002.

[11] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Transactions on Information Theory*, Vol. 43, pp. 1412–1424, September 1997.

[12] S. Sesia, *Advanced Coding Techniques for Multicasting in Wireless Communications,* PhD thesis, Ecole Nationale Supérieure des Télécommunications (ENST), June 2005.

[13] S. Sesia, G. Caire, "Superposition vs progressive vs hybrid approaches for lossy transmission over BF-channels and code construction," *To be submitted to IEEE Trans. on Wireless*, 2005.