# SCALAR SCHEME FOR MULTIPLE USER INFORMATION EMBEDDING

*Abdellatif Zaidi, Pablo Piantanida and Pierre Duhamel*

{zaidi,piantanida,pierre.duhamel}@lss.supelec.fr
Laboratoire des Signaux et Systèmes LSS/CNRS
Plateau de Moulon, 3 rue Joliot-Curie - 91190 Gif sur Yvette - FRANCE

**ABSTRACT**

Multiple watermarking is concerned with embedding several messages into the same host signal, with different robustness and transparency requirements. This paper proposes two implementable scalar schemes for multiple user "Dirty Paper Coding". The first -straightforward- approach consists in an independent superposition of two scalar Dirty Paper Coding schemes. The second consists in a joint design of a scalar Dirty Paper Coding. This joint approach is based on the ideal DPC scheme for Broadcast Channels with noncausal side information known to the transmitter. For this purpose, the "Scalar Costa Scheme" that has been originally conceived for one user is extended to two users. Performance evaluations, including Bit Error Rates and Capacity region curves are provided for both methods, illustrating the improvements brought by a joint design.

## 1. INTRODUCTION

Consider the problem of communicating over a Gaussian channel corrupted by an additive Gaussian interfering signal that is noncausally known only to the transmitter. This variation of the conventional additive white Gaussian noise (AWGN) channel is commonly known as *channel with side information at the encoder*. The state $S$ is a random Gaussian variable with power $Q$ which is independent of the gaussian noise $Z$. The channel input is the index $W \in \{1, \ldots, \mathcal{M}\}$ and its output is $Y^n = X^n + S^n + Z^n$, where $\mathcal{M}$ is the greatest integer smaller than or equal to $2^{nR}$ and $R$ is the rate in bit per transmission. The capacity expression of this single-user channel with random parameters has been derived by Gel'fand and Pinsker in [1]. They have shown that the capacity of such a channel $\{p(y|x,s)\}$ with side information $S$ noncausally available at the transmitter is

$$C = \sup_{p(u,x|s)} \big\{ I(U;Y) - I(U;S) \big\}. \qquad (1)$$

$U$ is an auxiliary random variable chosen so that $U \to (X,S) \to Y$ form a Markov Chain and $p(u,x|s) = \delta\big(x - f(u,s)\big)p(u|s)$.

In his "Writing on Dirty Paper" [2], Costa applied this result to the case of an AWGN channel corrupted by an additive white Gaussian interfering signal $S$. He showed that choosing $U = X + \alpha S$ with an appropriate value for $\alpha$ ($\alpha^* = P/(P+N)$, $N$ being the AWG noise) allows one to achieve the same capacity as if the interfering signal $S$ was not present, i.e. $\mathcal{C}_{DPC} = \frac{1}{2}\log(1+\frac{P}{N})$. This result has been then extended to non-Gaussian interfering signals [3] and non-stationary/non-ergodic Gaussian interference [4]. Costa's scheme for Dirty Paper Coding (DPC) is commonly known as the Ideal Costa Scheme (ICS). The term "Ideal" refers to an optimal random method for codebook generation and random binning coding. However, in order to attain the full capacity region, the number of codebook entries must be so huge that it is unfeasible in practical communication systems. Therefore, several suboptimal but low-complexity schemes have been proposed. In [5], G. Wornell designed a channel encoder based on a set of dithered quantizers. Whereas in [6], Eggers proposed a practical approach where the random codebook $U$ is chosen to be a concatenation of scalar uniform quantizers. This sample-wise (scalar) encoding and decoding procedure has been denoted as *Scalar Costa Scheme* (SCS). Using an appropriate parameter $\alpha$, SCS has been shown to have a better capacity and lower Bit Error Rates (BER) than other methods, including Quantization Index Modulation [5] and the traditional Spread-Spectrum. Further, the SCS performs close to the ideal DPC.

In this paper we propose two practical coding and decoding schemes for implementing a two-users Dirty Paper Coding. A straightforward generalization of SCS to an arbitrary number of users would consist in a multiple scalar DPC, independently superimposed. Performances of this method are analysed and we then show that this approach can be improved by a *joint coding*. This necessitates a connection with the physically degraded Gaussian Broadcast Channel (GBC) [7] with state information noncausally known to the transmitter. Therefore, a jointly scalar DPC ("Joint SDPC") scheme is derived from the ideal DPC for Broadcast Channels published in [8] which is, like the ICS, too complex to be implemented in real applications.

The paper is organised as follows: Section 2 presents the watermarking problem and the Scalar Costa Scheme (SCS). In section 3, a mathematical model for simultaneous watermarking, "Double SDPC" and its corresponding achievable rates are presented. Section 4 proposes a parallel between simultaneous watermarking and the GBC with side information to the encoder from which a Jointly SDPC is obtained. Finally, performances of these scalar approaches are compared to the ideal DPC for Broadcast Channel and implications are discussed.

## 2. WATERMARKING AS COMMUNICATION WITH SIDE INFORMATION

Digital watermarking can be considered as a communication problem, where a message $W \in \{1, \ldots, \mathcal{M}\}$ has to be sent to a receiver through some channel (*the watermark channel*), assumed to be Gaussian. $W$ is encoded into a code $X$ called the watermark which is then embedded into the host signal $S$ (referred as *the*

The authors would like to thank Rémy Boyer and Samson Lasaulce for great help during this work, and the SDMO RNRTproject for funding

*cover signal*) thus forming, the watermarked data $S + X$. The watermark is usually embedded without introducing perceptible distorsions to the host signal. This is called the *transparency* requirement. *The robustness* requirement, as for it, refers to degradations the watermark should survive. The resulting transmission channel

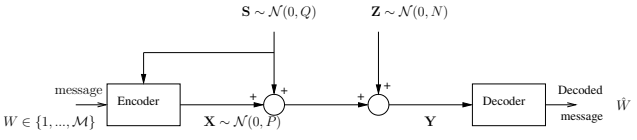


**Fig. 1**. *Watermarking viewed as a communication channel.*

in Fig. 1 is that described in "Writing on Dirty Paper". The corresponding channel capacity is the same as if the interfering signal $S$ was not present. However, for a practical implementation it is not possible to use the ideal DPC scheme. Instead of the ideal coding, in [6] Eggers proposed a suboptimal practical version of DPC, known as (SCS), outlined below.

### 2.1. The Scalar Costa Scheme (SCS)

Dirty Paper Coding is based on a random binning argument: given a realization of $S^n$, the encoder looks for a sequence $U^n$ such that $(U^n, S^n)$ is jointly typical. A geometrical interpretation is searching for a sequence $q = \frac{U}{\alpha} - S$, which is nearly orthogonal to $S$ and then transmit $X = \alpha q$. Consequently, the "search" could be viewed as a quantization of $S$, using a quantizer where every centroid is derived from the codebook entries via $U/\alpha$. This interpretation was used by Eggers [6] in order to propose a practical coding based on scalar quantizers at both encoders and decoders. The watermark signal is given by: $x_n = \tilde{\alpha}\mathcal{Q}_\Delta\{s_n - \frac{w_n}{\mathcal{M}}\Delta\} - \tilde{\alpha}(s_n - \frac{w_n}{\mathcal{M}}\Delta)$

with $\Delta = \frac{\sqrt{12P}}{\tilde{\alpha}}$ and $\tilde{\alpha} = \sqrt{\frac{P}{P+2.71N}}$, where $\mathcal{Q}_\Delta$ is a scalar uniform quantizer with constant step size $\Delta$. Here, the quantized values of $S$ form the codebook. To decode a transmitted message, the decoder quantizes the received signal $y = x + s + z$ and then looks for the closest codebook entry. To briefly summarize, simple hard decision decoding of $w_n$ is achieved by scalar quantization of $y_n$ followed by a thresholding procedure, i.e. let $r_n = \mathcal{Q}_\Delta\{y_n\} - y_n$, the decoded $\hat{w_n}$ is the closet integer to $\frac{r_n}{\Delta/\mathcal{M}}$. The optimum parameter $\alpha$ is $\tilde{\alpha}$, obtained by maximixing the mutual information $I(r; W)$. The maximal archievable rate for this practical coding is given by

$$\mathcal{C}_{\text{SCS}} = \max_{\alpha} I(r; W) \leq \mathcal{C}_{\text{DPC}}. \quad (2)$$

In a single user setting, as depicted above, the performances of SCS have been shown to be close to theorical DPC [6]. This is clearly a motivation to adapt it to the multiple user case. In a data hiding context, "Multiple user" means encoding several messages into the same host signal, with different robustness requirements. In the next sections two different approaches are first proposed and then compared.

## 3. MULTIPLE WATERMARKING: "DOUBLE DPC"

Consider a watermarking system with two watermarks that jointly satisfy a power constraint $P$. The main purpose of this system is to encode two messages $m_1$ and $m_2$ into the same host signal $S$. To this end, $m_1$ and $m_2$ are first mapped to two sequences of $\mathcal{M}_1$-ary and $\mathcal{M}_2$-ary symbols, repectively. The resulting watermarking problem is equivalent to that of transmitting a pair of indexes $(W_1, W_2) \in \{1, \ldots, \mathcal{M}_1\} \times \{1, \ldots, \mathcal{M}_2\}$ over the watermark channel. Let $X_1$ (carrying $W_1$) and $X_2$ (carrying $W_2$) be, respectively, the *fragile* and *robust* watermark. These watermarks are supposed to survive channel distorsions $N_1$ and $N_2$, with $N_1 \leq N_2$. Thus, the watermak signal is $X = X_1 + X_2$. The input power constraint can be satisfied using a power-sharing technique, i.e. $E[{X_1}^2] = \gamma P$ and $E[{X_2}^2] = (1-\gamma)P$, with $\gamma \in [0, 1]$[1]. Decoder 1 decodes $\hat{W}_1$ from $Y_1$ at rate $R_1$, and Decoder 2 decodes $\hat{W}_2$ from $Y_2$ at rate $R_2$. The decoding is successful if $(\hat{W}_1, \hat{W}_2) = (W_1, W_2)$. A simplified corresponding diagram is shown in Fig. 2. Designing a watermark system for coding two watermarks consists in finding encoder/decoders strategies (schemes and optimal parameters) that allow to simultaneously encode and independently decode $W_1$ and $W_2$ at the highest possible rates $R_1$ and $R_2$. A simple, rather intuitive, approach is based on using two independent single-user DPC (or SCSs for a suboptimal practical implementation). In essence, this coding corresponds to the following steps: (i) to form $X_2$ [2], use a first DPC (DPC1) taking into account the known state $S$ and the unknown noise $Z_2$, (ii) to form $X_1$, use an other DPC (DPC2) taking into account the sum state $S + X_2$ and unknown noise $Z_1$, (iii) finally, transmit the sum $X = X_1 + X_2$ over the channel. The resulting coding scheme and optimal parameters (separately optimized) are given by:

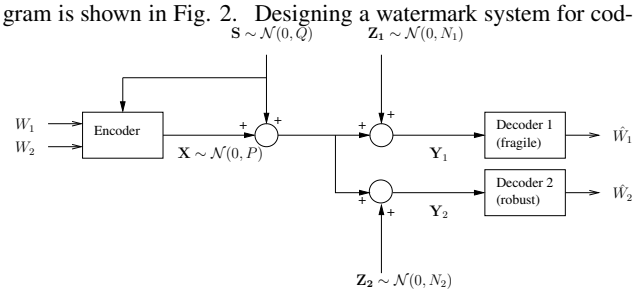


**Fig. 2**. *Two users watermarking viewed as a double DPC.*

(a) Channel $Y_2$ (DPC1): $X_2 = U_2 - \alpha_2 S$ where

$$U_2 \sim \mathcal{N}\left(\alpha_2 S, (1-\gamma)P\right) \text{ with } \alpha_2 = \frac{(1-\gamma)P}{(1-\gamma)P + N_2} \quad (3)$$

(b) Channel $Y_1$ (DPC2): $X_1 = U_1 - \alpha_1(S + X_2)$ where

$$U_1 \sim \mathcal{N}\left(\alpha_1(S + X_2), \gamma P\right) \text{ with } \alpha_1 = \frac{\gamma P}{\gamma P + N_1}, \quad (4)$$

The theoretically achievable rates $R_1$ and $R_2$ corresponding to DPC1 and DPC2 are given by

$$R_1 = \frac{1}{2}\log_2\left(1 + \frac{\gamma P}{N_1}\right) \text{ and} R_2 = R(\alpha_2, (1-\gamma)P, Q, \gamma P + N_2), \quad (5)$$

with $R(\alpha, P, Q, N) = \frac{1}{2}\log_2\left(\frac{P(P+Q+N)}{PQ(1-\alpha)^2 + N(P+\alpha^2 Q)}\right)$.

For a practical implementation of this coding, we remplace these

[1] This means $E[{X_1}^2] + E[{X_2}^2] = P$.

[2] It could be argued that the robust watermark must be encoded first.

single-user DPC by two single-user SCS, with parameters:

$$\tilde{\alpha_1} = \sqrt{\frac{\gamma P}{\gamma P + 2.71 N_1}}, \quad \tilde{\alpha_2} = \sqrt{\frac{(1-\gamma)P}{(1-\gamma)P + 2.71 N_2}} \tag{6}$$

Maximal achievable rate pairs for this coding scheme (denoted

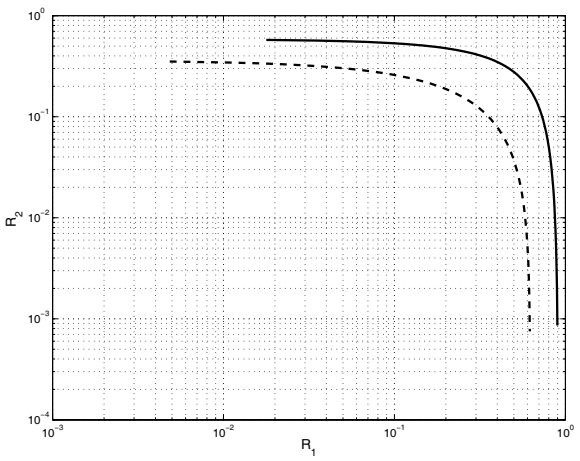


**Fig. 3**. *Theoretical and practical achievable capacity regions. Solid line correspond to the double ideal DPC given by equations (3), (4) and (5). Dashed lines correspond to the double SCS coding given by equations (6) and (2).*

hereafter as $\tilde{R_1}$ and $\tilde{R_2}$) are computed numerically using Eqs. (2) and (6). Fig. 3 shows the practical rate pair ($\tilde{R_1}$, $\tilde{R_2}$) with respect to the theoretical rate pair ($R_1$, $R_2$). Eq. (5), shows that DPC2, as given by (4) is optimal, because the achievable rate $R_1$ is equal to that of a channel with no interfering signals $S$ and $X_2$. DPC1, however, is not. The reason is as follows: if we take this multi-user watermarking system as a Gaussian Broadcast Channel, we see that $R_2 \leq R_2^{(\max)}$ where $R_2^{(\max)} = \frac{1}{2}\log_2\left(1 + \frac{(1-\gamma)P}{\gamma P + N_2}\right)$ [7]. Thus, in the following section we show that the encoding strategy of $W_2$ can be improved to bring the rate $R_2$ close to $R_2^{(\max)}$. The corresponding scheme is denoted as "Joint DPC".

## 4. MULTIPLE WATERMARKING: "JOINT DPC"

In this section, we first present a parallel with the GBC with state information noncausally known to the transmitter. Few recent findings in network information theory are briefly reviewed and then a more efficient coding scheme is derived.

### 4.1. Connection to the Gaussian BC

The communication scenario depicted in Fig. 2 is basically that of a GBC with state information noncausally known to the transmitter but not to the receivers. In [8], it has been shown that the capacity region of this channel is given by

$$R_1 \leq \frac{1}{2}\log_2\left(1 + \frac{\gamma P}{N_1}\right), \quad R_2 \leq \frac{1}{2}\log_2\left(1 + \frac{(1-\gamma)P}{\gamma P + N_2}\right) \tag{7}$$

which is the capacity region of a GBC with no interfering signal $S$. This region (7) can be attained by an appropriate succesive encoding scheme that uses two well designed DPCs, as proved in [8]. The encoding of $W_1$ (DPC2) is still given by equation (4). For the other message, the most important point is to consider $X_1$ as unknown Gaussian noise and then combine it with the channel noise $Z_2$. We use a DPC (DPC2) to form an optimal $X_2$ with respect to that "noise":

$$U_2 \sim \mathcal{N}(\alpha_2 S, (1-\gamma)P) \text{ with } \alpha_2 = \frac{(1-\gamma)P}{P + N_2} \tag{8a}$$

$$X_2 = U_2 - \alpha_2 S. \tag{8b}$$

Note that, even if the inteference due to $X_1$ is not cancelled, the upper bound for $R_2$ in (7) is attained. We now design a practical Jointly DPC coding based on SCS.

### 4.2. Capacity region of the Jointly Scalar DPC

Consider a SCS implementation of the Joint DPC presented above. The optimal parameters defining this coding scheme are $(\tilde{\alpha_1}, \tilde{\alpha_2})$ and $(\Delta_1, \Delta_2)$. The maximal archievable rate pair for this practical coding is given by the convex hull of all rate pairs $(\tilde{R_1}, \tilde{R_2})$ satisfying

$$\tilde{R_1} \leq I(r_1, W_1|W_2), \quad \tilde{R_2} \leq I(r_2, W_2) \tag{9}$$

where $r_i = \mathcal{Q}_{\Delta_i}\{y_i\} - y_i$ and $i \in \{1, 2\}$. Let two SCS silar to SCS1 and SCS2 described in section 3. The optimal value for $\tilde{\alpha_2}$ is obtained by maximixing $\tilde{R_2}$ in (9). For the other channel, we simply use the value $\tilde{\alpha_1}$ in (6),

$$\tilde{\alpha_2} = \sqrt{\frac{(1-\gamma)P}{(1-\gamma)P + 2.71(\gamma P + N_2)}}, \quad \tilde{\alpha_1} = \sqrt{\frac{\gamma P}{\gamma P + 2.71 N_1}}. \tag{10}$$

To evaluate practical rates in (9), we need computing $p_{r_2}(r_2|W_2)$ and $p_{r_1}(r_1|W_1; W_2)$ probabilities. This can be done using the high resolution quantization assumption ($Q \gg P$), which is reasonable in most watermarking applications (see [6]). In Fig. 4 we observe the resulting rates compared to the ideal DPC for BC (given by equations (8) and (4)). Note that the depicted curves naturally depend on the ratios $P/N_1$ and $P/N_2$, which have been set to $P/N_1 = 10P/N_2 = 10$. Improvement over the "Double DPC" is made possible through increasing the achievable rate $R_2$. We can see that, for asymptotically large alphabet size, the Joint Scalar DPC performs close to the ideal DPC derived from the broadcast solution.

### 4.3. BER analysis and discussion

Another performance evaluation is based on measured BERs for hard decision of binary Jointly Scalar DPC as shown in Fig. 5. Curves are obtained with a Monte Carlo simulation. The signal-to-noise power ratios are given by $\text{SNR}_1 = 10\log_{10}\left(\frac{\gamma P}{N_1}\right) \in [-10, 9.6]$ dB and $\text{SNR}_2 = 10\log_{10}\left(\frac{(1-\gamma)P}{\gamma P + N_2}\right) \in [-11, 7]$ dB. At such SNR-range, it has been shown that repetition coding is almost optimal. In Fig. 5, curves are obtained with $(\rho_1, \rho_2) = (4, 4)$, meaning that $W_1$ and $W_2$ are being repeated 4 times each. In practical situations, $\rho_1$ and $\rho_2$ should be chosen in light of the desired bit-error-rates. The choice $(\rho_1, \rho_2) = (4, 4)$ should be taken just as a toy example, but it is already sufficient to point out the additional degree of freedom provided by $\rho_1$ and $\rho_2$ to attain targeted decoding reliabilities. Also, It can be easily checked that, when plotted separately versus the SNR, the curves are identical to those of a SCS with a signal-to-noise power ratio equal to $\text{SNR}_1$

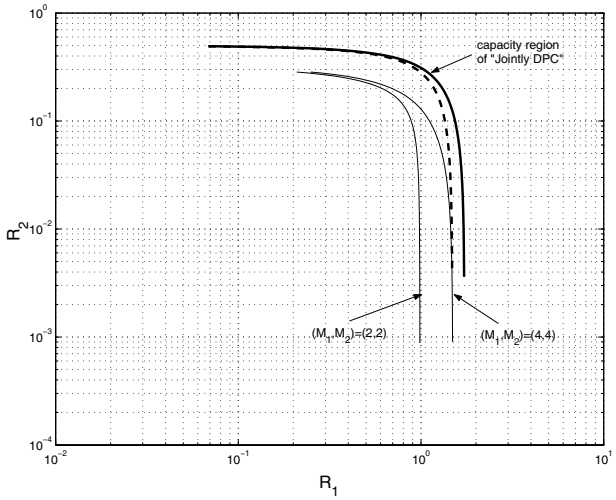


**Fig. 4**. *Achievable rates of the Joint Scalar DPC, for both binary and quaternary alphabets (solid line). The upper bound is the capacity region curve of the ideal DPC for BC. Dashed line correspond to maximum achievable rate pairs for the Double DPC.*

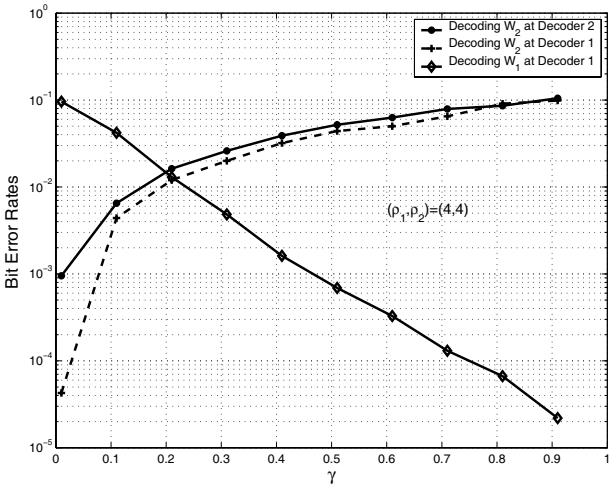


**Fig. 5**. *Bit error probabilities for binary transmission with repetition coding. Each message is being repeated* 4 *times. Plotted curves correpond to* $P = 5$*;* $N_2 \approx 1$ *and* $N_1 = 0.5$*. Signal-to-Noise power ratios are* $SNR_1 \in [-10, 9.6]$ *dB and* $SNR_2 \in [-11, 7]$ *dB.*

and $\mathrm{SNR}_2$, respectively. Channel coding results as applied previously to the Eggers SCS can also be applied to both channels $Y_1$ and $Y_2$ taken separately. As $\gamma \in [0, 1]$ increases, the power part of the signal $X$ carrying $W_1$ becomes stronger and that for $W_2$ becomes weaker. Corresponding BERs monotoneously decreases and increases, respectively.

Fig. 5 shows that the *worst* channel $Y_2$ (more noisy) has naturally a much larger error-probability than that of $Y_1$. For a joint coding, however, it is interesting to note that decoding message $W_2$ at decoder 1 (fragile) should, in principle, be more reliable than that at the decoder 2 since the noise $Z_1$ is less-powerfull than $Z_2$. Therefore, decoder 1 should be able to jointly decode $W_1$ and $W_2$. In a context of data hiding, this means that the robust watermark (supposed to survive channel degradations up to $N_2$) should be reliably detected by the other decoder. This is a dividend for Broadcast Channels where successive coding requires that Decoder 1 first decodes $W_2$ and then $W_1$. In our scheme, it performs only slightly better than Decoder 2 when decoding $W_2$. The difference around $\gamma = 0$ could be larger, so as to correspond to about 3 dB of SNR improvement at the decoder input. This is due to the fact that, parameter $\tilde{\alpha_1}$ is not optimally tuned for decoding $W_2$ from channel 1 output. Further investigations are needed to closely fit with Broadcast requirements.

## 5. CONCLUSION

In this paper, we investigated two practical (low-complexity) coding and decoding schemes in order to implement a two-users Dirty Paper Coding. The first method consists in a "Double DPC", which are independently superimposed. The second one corresponds to a "Joint DPC" scheme derived from the ideal DPC for Broadcast Channels with noncausally side information known to the encoder. We have shown that the second method provides better Bit Error Rates and information rates. On the other hand, we observed that the parameter $\alpha$ for the SCS is not optimized to allow the better channel (less robust watermark, $Y_1$) to decode both messages ($W_1$ and $W_2$). However, it is sufficient for the sole decoding of $W_1$. We also pointed out the role of repetition coding as an additional degree of freedom to allow different desired decoding reliabilities. The practical coding schemes of this paper have been proposed within the context of data hiding but should be easily exploitable for other applications such as data compression or traditional multi-user communications where the transmitter/source has several messages or data flows to encode. Multiple watermarking has been considered since encoding several watermarks for *several* users or a watermark that is attacked *several* times, readily fits with the proposed schemes: for embedding two messages, for example, the first mark has to be very robust but does not carry much information (receiver 2) and the second mark is asked to be very informative at the cost of reduced robustness (receiver 1).

## 6. REFERENCES


[1] S. I. Gel'fand and M. S. Pinsker, "coding for channel with random parameters," *Problems of Control and IT.*, vol. 9, pp. 19–31, 1980.

[2] M. H. M. COSTA, "Writing on dirty papers," *IEEE Trans. on IT*, vol. IT-29, pp. 439–441, may 1983.

[3] A. S. Cohen and A. Lapidoth, "Generalized writing on dirty paper," in *Proc. ISIT 2002*, Lausanne-Switzerland, July 2002.

[4] W. Yu, A. Sutivong, D. Julian, T. M. Cover, and M. Chiang, "Writing on colored paper," in *Proc. IEEE ISIT*, June 2001.

[5] B. Chen and G. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, pp. 1423–1443, may 2001.

[6] J. J. Eggers, R. Buml, R. Tzschoppe, and B. Girod, "Scalar costa scheme for information embedding," *IEEE Transactions on Signal Processing*, pp. 1–39, 2002.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, J. W. . S. INC., Ed., New York, 1991.

[8] Y.-H. Kim, A. Sutivong, and S. Sigurjonsson, "Multiple user writing on dirty paper," in *Proc. ISIT 2004*, Juin 2004, p. 534.