Coding Schemes for Relay-Assisted Information Embedding

Abdellatif Zaidi and Luc Vandendorpe, *Fellow, IEEE*

***Abstract*—Cooperative information embedding deals with the problem of embedding unperceived information into some cover signal by different users or partners, cooperatively. It models applications in which embedded signals, or watermarks, transmitted over wireless networks need to be reinforced in order to withstand channel impairments. In cooperative information embedding, each embedder that can reinforce the embedded signal may or may not know the original cover signal. In this paper, we concentrate on the two user cases: 1) an initial embedder and 2) an assisting embedder or helper collaborate to embed some watermark into given digital media content which is transmitted over a wireless network. One important application is that of infrastructure-aided information embedding, a case in which the network provider plays the role of a helper and contributes to securing the distribution of the media, not only by blocking unauthorized signals but also by reinforcing the watermarks in legitimate signals. We investigate the two scenarios in which the helper does or does not know the cover signal. For each scenario, we derive lower and upper bounds on channel capacity. Furthermore, we also design implementable coding schemes and derive the embedding rates practically allowed by these schemes, for both scenarios. Among others, the performance characterization shows that for cooperative information embedding to be effective, careful code design is required at the initial embedder and the helper. The careful design concerns the joint conception of the embedded codes and the exploitation of the knowledge of the cover signal, if any.**

***Index Terms*—Dirty paper coding (DPC), information embedding (IE), noncausal channel state information, relay channel.**

I. Introduction

INFORMATION EMBEDDING (IE) deals with the problem of enhancing the process of embedding extraneous data, referred to as “embedded signal” or “watermark,” into another signal, referred to as “cover-” or “host-signal” [1], [2]. While the array of applications and advances of the theory and interconnections to other problems continue to grow (e.g., in broadcast channel [3], [4]; multiple-access channel [3], [4]; and multicast channel [5] applications), IE has gained much interest in the mid 1990s mainly as a potential solution for multimedia security and digital rights management using watermarking and data-hiding techniques (see, for example, [6]–[10] and references therein). In these applications, rights owners embed in an audio, image, or video signal additional data in an imperceptible manner to establish ownership, usage rules, or track media redistribution. Information embedding has also potential usage in some nonsecurity-oriented applications, such as backwards-compatible upgrades to an existing communication infrastructure [11], interactive advertising, and added services to broadcast audio [12].

The key aspects of IE that make it attractive are the imperceptibility of the embedded data, its statistical covertness, and its ability to potentially withstand several channel degradations, including incidental and intentional attacks. For example, for digital media communication over wireless networks, information embedding techniques and, for instance, watermarking and data-hiding tools can be a relevant supplement to traditional security tools based on encryption and Firewall systems [13], [14]. This is because systems that are based only on encryption are prone to large threats once the digital content is decrypted, whereas watermarks are designed to travel with the host signals wherever these go into the network.

In this paper, we show that in a large wireless network, embedding some secret unperceived information for the sake of securing the communication of digital media content can be performed in a distributed manner. Also, we show the embedding benefits from being handled by more than one embedder or partner. In the setup that we consider, a primary user embeds a watermark at some point of the network and at some other point of the network, the watermark is reinforced by means of a secondary embedder so that it can survive stronger channel degradations on its way to the destination. Thus, in a sense, the secondary embedder acts as a relay for the message or watermark embedded by the primary user. We refer to this situation as “cooperative information embedding,” by reference to the collaborative embedding of the watermark by the cooperating users or partners. Similar techniques that attempt to overcome the inherent challenges of security in a distributed manner include distributed authentication, distributed secure delivery, and distributed intrusion detection (see, for example, [15] and references therein).

One particularly interesting application of collaborative embedding over a wireless network is that of “infrastructure-aided” IE, a situation in which, in agreement with the content owners, the network provider exploits the centralized control that it has over the network to not only monitor the traffic and possibly block the redistribution of unauthorized signals but also to reinforce the watermarks in the legitimate signals that are transmitted over the network. It can also track down illicit users. The

Manuscript received June 30, 2007; revised October 14, 2008. Current version published February 11, 2009. This work was supported in part by the Walloon Region DGTRE Nanotic-COSMOS program, in part by the EU framework program COOPCOM, and in part by the FP7 Network of Excellence NEWCOM++. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fernando Perez-Gonzalez.

The authors are with the École Polytechnique de Louvain, Université Catholique de Louvain, Louvain-la-Neuve 1348, Belgium (e-mail: abdellatif.zaidi@uclouvain.be; luc.vandendorpe@uclouvain.be).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2008.2009588

watermark reinforcement can be implemented at some base stations or other specific access points of the network. Reinforcing the watermark at a base station, for example, makes it easier for the watermark to survive subsequent channel impairments, and, of course, harder for an attacker to remove it. Through this cooperative embedding, the content providers benefit since the network operator contributes to securing the distribution of their media contents, and the latter benefits, in turn, since it gains more credit in the eyes of content providers and retailers, thus encouraging the creation of new business opportunities on the network. This example is discussed in more detail in Section III. For instance, it is argued more on that it is of mutual interest for the content providers and the network provider to collaborate; and it is shown that this can be done without the need to share all of the secret information (messages to embed and keys). The same technique of watermark reinforcement can be used to enhance watermarking solutions applied to emerging audio/video networked applications, such as Pay-TV and video-on-demand (VoD) services on wireless networks.

The question of cooperative IE is very relevant in practice as an assisting-functionality for media security over networks. However, it poses a number of challenging problems among which the question of whether all, only some or none of the cooperating embedders know the original media content. This paper addresses this question in the case of two embedders, a primary user, or initial embedder and a helper. The helper reinforces the message that is embedded by the primary user against subsequent channel degradations on the network. This setup has a connection with state-dependent relay channels [16] and it divides into two different scenarios, according to whether the helper knows the original host signal or not.

For each of the two scenarios, we derive lower and upper bounds on channel capacity. For the scenario in which the helper knows the host signal, the lower bound for the Gaussian case is obtained by a coding scheme in which the initial embedder and the helper employ standard dirty paper coding (DPC) [17] schemes. For the scenario in which the helper does not know the host signal, the lower bound for the Gaussian case is obtained by a coding scheme in which the initial embedder employs a generalized DPC (GDPC) scheme. This GDPC consists of a combination of partial cancellation of the known host signal and DPC.

Furthermore, we also design implementable cooperative embedding schemes and investigate the allowed embedding rates for the two scenarios. For the scenario in which the helper knows the host signal, the coding scheme is based on lattice codes and minimum mean-squared error (MMSE) scaling. For the scenario in which the helper does not know the host signal, the coding scheme is based on a combination of lattice codes, MMSE scaling, and standard (capacity-achieving) Gaussian codes.

It should be mentioned that investigating the problem of cooperative IE from an information-theoretic point-of-view in this paper not only provides a yardstick by which the efficiency of cooperative embedding or watermarking can be measured, but in fact also provides the right guidance to the appropriate design in practice. In particular, one of the insights developed in this paper is that for a networked multimedia communication, securing the distribution of a digital content by means of IE can take advantage from being handled cooperatively (i.e., by more than one user or partner), rather than individually. The enabled improvement is illustrated through the allowed increase in the embedding rate, relative to the situation in which the watermark is not reinforced.

The rest of this paper is organized as follows. After introducing the notation, we recall some known results about DPC in Section II. Also, we give a brief review of the formal statement of the additive Gaussian IE problem as DPC, together with the state of the art of some implementable coding schemes. Then, in Section III, we present our practical motivation. In Section IV, we provide two mathematical models for cooperative IE. In Section V, we analyze the scenario in which the helper knows the host signal; and in Section VI, we analyze the scenario in which the helper does not know the host signal. Finally, in Section VII, we provide some conclusions.

A. Notation

Throughout this paper, boldface fonts denote vectors. We use uppercase letters to denote random variables or vectors (e.g., $\mathbf{X} = (X_1, X_2, \ldots, X_n)$), lowercase letters for their realizations (e.g., $\mathbf{x} = (x_1, x_2, \ldots, x_n)$), and calligraphic fonts for sets (e.g. $\mathcal{X}$). Unless otherwise specified, vectors are assumed to be in the n-dimensional Euclidean space $(\mathbb{R}^n, \|\cdot\|)$ where $\|\cdot\|$ denotes the Euclidean norm of vectors. For a random vector $\mathbf{X}$, we use $\mathbb{E}_{\mathbf{X}}[\cdot]$ to denote the expectation taken with respect to $\mathbf{X}$ and $P_{\mathbf{X}}(\cdot)$ to denote the probability distribution of $\mathbf{X}$. A random vector $\mathbf{X}$ with conditional probability distribution given $\mathbf{S}$ is denoted by $\mathbf{X}|\mathbf{S}$. The Gaussian distribution with mean μ and square deviation σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$. The $n \times n$ identity matrix is denoted by I. For random vectors $\mathbf{X}$, $\mathbf{U}$, and $\mathbf{S}$, the notation

$$\begin{aligned}\mathbf{U}|\mathbf{S} &\sim \mathcal{N}(\alpha\mathbf{S}, PI)\\ \mathbf{X} &= \mathbf{U} - \alpha\mathbf{S}\end{aligned}$$

is used to mean that $\mathbf{X}$ is i.i.d. Gaussian with power P (i.e., $\mathbf{X} \sim \mathcal{N}(0, PI)$), independent of $\mathbf{S}$, and $\mathbf{U}$ is generated as $\mathbf{U} = \mathbf{X} + \alpha\mathbf{S}$ for some scalars α and P. Throughout this paper, the logarithm function is to base 2; and for a scalar $u \in [0, 1]$, its complement to unity is denoted by $\bar{u}$ (i.e., $\bar{u} = 1 - u$).

II. Information Embedding as Dirty Paper Coding

In this section, first we give a brief review of the additive Gaussian IE problem viewed as DPC. Next, we review some coding realizations of DPC based on lattices [18].

A. Additive Gaussian IE as DPC

Fig. 1 depicts a block diagram of the problem of IE. A message m is to be embedded into the host signal $\mathbf{S}$, transmitted through some channels and then received at some receiver. The receiver does not know the host signal and has to decode the embedded message from the received signal. The message m can be represented by a sequence $\{W\}$ of M-ary symbols, $W \in \mathcal{W} = \{1, \ldots, M\}$ so that the embedding of message m amounts to that of the sequence of symbols $\{W\}$. Thus, in the rest of this

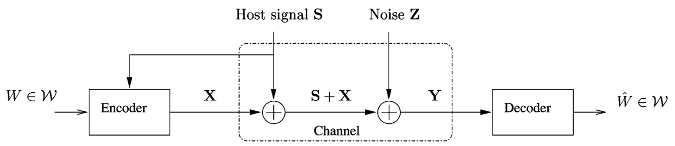


Fig. 1. IE viewed as communication with the noncausal side information at the transmitter.

paper, we will loosely refer to each symbol $W \in \mathcal{W}$ as being a "message."

The embedding process consists in encoding the message W into a signal $\mathbf{X}$, called the "embedded signal" or "watermark," which the embedder then adds to the host signal. The embedding rate R, expressed in number of bits per host sample that the encoder can embed reliably, is such that $M \approx 2^{nR}$. For imperceptibility reasons, embedding should not introduce any perceptible distortion to the host signal; and this imposes an embedding power constraint of the form $\mathbb{E}[||\mathbf{X}||^2] \leq nP$. Also, the embedded signal must survive certain channel degradations, including some common incidental and intentional attacks.

The IE problem shown in Fig. 1 can be viewed as a communication problem with side information (SI) known noncausally at the transmitter but not at the receiver [6]; the SI being the cover signal, the transmitter being the embedder, and the transmission rate being given by the embedding rate. In this model, the SI acts as an interference for the transmission of the message. For the i.i.d. Gaussian case, the relevant work is Costa's "writing on dirty paper" [17], adeptly known as DPC. More specifically, if the SI $\mathbf{S}$ and the noise $\mathbf{Z}$ are independent and i.i.d. Gaussian, with $\mathbf{S} \sim \mathcal{N}(0, QI)$ and $\mathbf{Z} \sim \mathcal{N}(0, NI)$, Costa was the first to show the remarkable result that the additive Gaussian interference $\mathbf{S}$, which is known noncausally only to the transmitter, incurs no loss in capacity relative to the standard interference-free additive white Gaussian noise (AWGN) channel, i.e.,

$$\mathcal{C} = \frac{1}{2}\log\left(1 + \frac{P}{N}\right). \tag{1}$$

The achievability proof is based on a random binning argument for general channels with noncausal state information [19]. It uses a random construction of a Gaussian codebook and a random partition of the codewords of this codebook into "bins." Costa showed that with the choice of the input distribution as

$$\begin{aligned} \mathbf{U}|\mathbf{S} &\sim \mathcal{N}(\alpha\mathbf{S}, PI) \\ \mathbf{X} &= \mathbf{U} - \alpha\mathbf{S}, \end{aligned} \tag{2}$$

where $\mathbf{U}$ is an auxiliary random variable and $\alpha = P/(P+N)$, one achieves the interference-free capacity (1) regardless of the power of the interference $\mathbf{S}$. This theoretical DPC, however, is not feasible in practice due to the huge random codebook which is needed to perform binning. Earlier DPC-based implementations for IE suboptimally set the signal $\mathbf{X}$ to be an appropriate scaled version of the quantization error of the host signal $\mathbf{S}$. Quantization can be scalar-valued [2], [20] or vector-valued (e.g., based on lattices [21]–[24]).

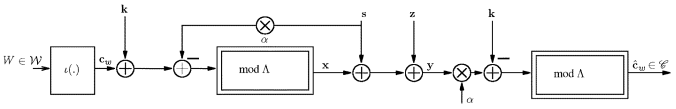


Fig. 2. Lattice coding for DPC based on modulo reduction.

B. Suboptimal Coding

Consider the IE scheme depicted in Fig. 2, where Λ is an n-dimensional lattice with the fundamental Voronoi region $\mathcal{V}(\Lambda)$. We denote $V(\Lambda)$ and $G(\Lambda)$ as the volume and the normalized second moment of lattice Λ, defined as

$$\begin{aligned} V(\Lambda) &\triangleq \mathrm{Vol}\,(\mathcal{V}(\Lambda)) = \int\limits_{\mathcal{V}(\Lambda)} \mathbf{dt} \\ G(\Lambda) &\triangleq \frac{1}{n} V(\Lambda)^{-\frac{n+2}{n}} \int\limits_{\mathcal{V}(\Lambda)} ||\mathbf{t}||^2 \mathbf{dt}. \end{aligned}$$

Assume that the encoder and the decoder share some common randomness in the form of a key $\mathbf{K}$ which is uniformly distributed over $\mathcal{V}(\Lambda)$. The key is used for security purposes [7], [25], and as a capacity-achieving tool [26]. Also, consider an indexing function $\iota(\cdot)$, which maps each message W to be embedded to a unique vector $\mathbf{c}_w$, taken from an appropriately chosen codebook $\mathcal{C} = \{\mathbf{c}_w : w = 1, \ldots, M\}$. The codebook $\mathcal{C}$ may be regarded as a lattice codebook. For each $W \in \mathcal{W}$, the vector $\iota(w) = \mathbf{c}_w$ is the coset leader of the coset $\Lambda_w = \mathbf{c}_w + \Lambda$ of lattice Λ.

Let $\bmod \Lambda$ denote the modulo reduction operation with respect to the fundamental Voronoi region $\mathcal{V}(\Lambda)$ of lattice Λ, defined as $\mathbf{t} \bmod \Lambda \triangleq \mathbf{t} - \mathcal{Q}_\Lambda(\mathbf{t}) \in \mathcal{V}(\Lambda)$. The n-dimensional quantization operator $\mathcal{Q}_\Lambda(\cdot)$ is such that quantization of $\mathbf{t} \in \mathbb{R}^n$ results in the closest lattice point $\boldsymbol{\lambda} \in \Lambda$ to $\mathbf{t}$. The watermarked signal received at the destination is given by the sum of the watermark $\mathbf{x}$, the host signal $\mathbf{s}$, and the unknown noise $\mathbf{z}$, i.e.,

$$\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{z}. \tag{4}$$

The encoding and decoding procedures are given by

$$\mathbf{x}(W; \mathbf{s}, \Lambda) = [\mathbf{c}_w + \mathbf{k} - \alpha\mathbf{s}] \bmod \Lambda \tag{5a}$$

$$\hat{W} = \arg\min_{W \in \mathcal{W}} \min_{\boldsymbol{\lambda} \in \Lambda_w} ||\alpha\mathbf{y} - \mathbf{k} - \boldsymbol{\lambda}|| \tag{5b}$$

where the scale parameter α can be optimized according to different criteria and the input is subject to the embedding power constraint $\sum_{i=1}^{n} x_i^2 \leq nP$. The choice $\alpha = 1$ corresponds to no scaling and is often referred to as zero-forcing DPC (ZF–DPC). Other optimization criteria for the parameter α can be minimum mean-squared error (MMSE-DPC) and minimum error entropy (MEE–DPC) [27]. We note that the sample-by-sample scalar Costa scheme (SCS) [20] corresponds to the special case $\Lambda = \mathbb{Z}$, and quantization index modulation (QIM) [2] corresponds to $\Lambda = \mathbb{Z}^n$. Also, the scheme (5) is closely linked to Tomlinson–Harashima precoding [28], [29].

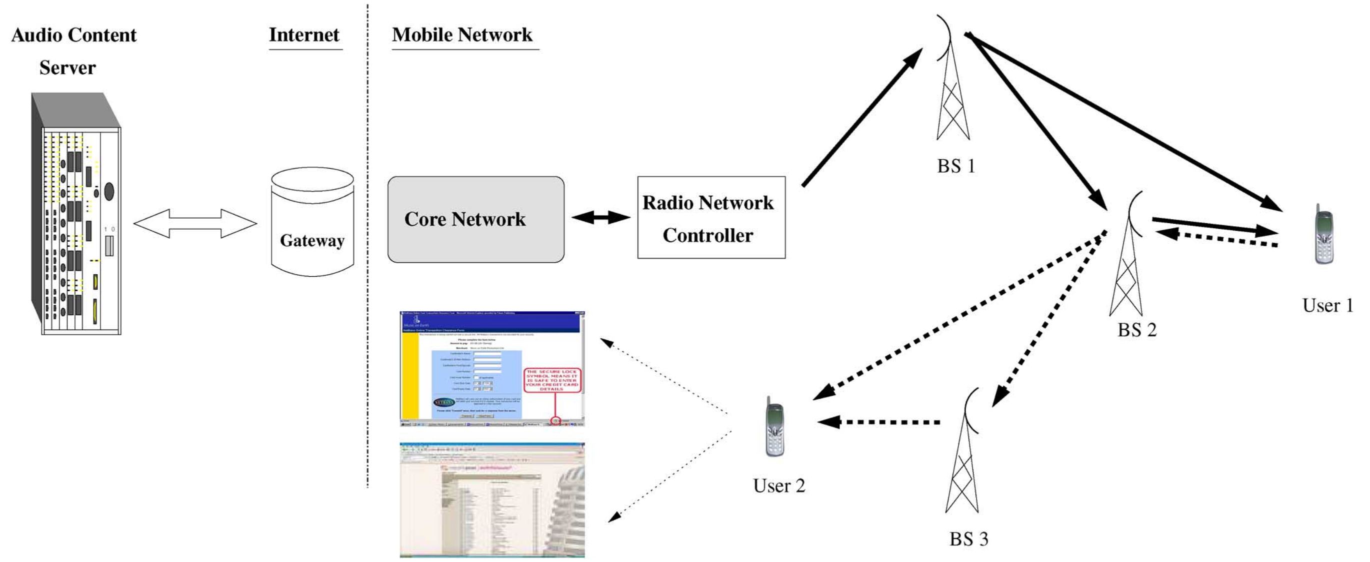


Fig. 3. Architecture for acquisition (solid) and redistribution (dashed) of protected proprietary audio contents on a mobile network.

III. Cooperative Information Embedding: Practical Motivation

For a wireless-networked multimedia communication in which IE is used as an assisting functionality to secure the distribution of digital media content, the embedded signals can be reinforced at some specific access points of the network. This section illustrates one use case scenario.

A. Infrastructure-Aided Information Embedding

The new generation of mobile networks and devices has opened up a new set of business opportunities for multimedia content distribution services on mobile phones. However, one major thwart against the development of such services is the fear of piracy, unauthorized copying, and illegal redistribution. While it seems difficult to build a provably secure end-to-end architecture that circumvents this problem based only on IE techniques, IE (and, for instance, watermarking) has the potential of deterring users from widely redistributing digital contents over networks illegally. The insertion of a watermark into an audio or video file allows content providers and/or the network provider to trace the content back to the phone or device on which it was downloaded. In a mobile network, this can be facilitated by the centralized control that the network provider has on all of the signals that can be exchanged over the network. For instance, the network provider can implement a watermark control at some base stations (and/or other access points of the network) to block the redistribution of illicit digital content [14]. It can also track down the corresponding users. For a mobile operator, the guarantee that no unauthorized content would pass through the network can dispel the concerns of those reluctant content providers, thus stimulating the expanding and development of new services on the network. This shows that it is of reciprocal interest for content and network providers to work together to make the forensic watermarking technology on mobile networks and devices effective. And this can be accomplished without necessarily sharing all of the secret information as will be explained in the example given below.

Moreover, the control that the network provider has on the traffic actually offers even more possibilities for the mobile operator to take part in securing the distribution of proprietary-watermarked digital content. For instance, the mobile operator does not only block unauthorized signals but also strengthens the watermarks in the legitimate signals. Reinforcing the watermark at a base station, for example, makes it easier for the watermark to survive channel degradations on its way to the destination and, of course, harder for an attacker to remove it. Also, it increases the accuracy of any subsequent watermark control handled at the terminal side or another base station by increasing the power of the measured watermark.

B. Use-Case Scenario

The example in this section builds partially upon a secure end-to-end architecture proposed in [14] for an audio content distribution service over a mobile wireless network. In [14], the authors tightly combine the traditional security tools based on cryptography and smart cards, together with watermarking and data-hiding tools. The watermark control is performed at the mobile terminal and the mobile network. Here, the mobile network also possibly reinforces the embedded watermark. Details related to digital rights-management (DRM) architecture and keys management can be found in [14]. For instance, it is interesting to observe that it is possible to adapt the classical open mobile alliance (OMA) DRM architecture to mobile devices' smart cards, through a once-in-a-life registration protocol so that each smart card can communicate with the server through a secure link. These details are omitted here for brevity.

Consider the audio content acquisition and redistribution scenario over a mobile network shown in Fig. 3. A user (user 1) wants to purchase audio content (e.g., music file) from a proprietary audio content server (music catalog). The request is processed by the server and, if the user has obtained the appropriate usage rights, an authentication phase to set up a security context is established and then the content is delivered. At the server side, the contents to be delivered are compressed and then encrypted and wrapped with a proprietary DRM header. Also, prior to the compression, a watermark is embedded into the content by the owner.

The usage rights and restrictions (expressed, for example, in the form of a separate license file which the server uploads into the smart card of the user's device upon payment of the service) are such that a legitimate user can transfer purchased audio content to another user, but the mobile device of the new user should not be able to play the audio content unless the new user has obtained the corresponding usage rights. This control is implemented at the smart card of the mobile device of the user receiving the content [14]. The license contains the watermark and also the keys that are necessary to the decryption of the content. At the mobile device, after decryption, the rendering module checks the watermark on the fly by comparing the watermark extracted from the played content with the one that is in the smart card. Rendition is stopped in case of a discrepancy between the two. As shown in Fig. 3, the service can also provide basic information about the audio content as well as a web link to the server so that new users willing to acquire the usage rights of the received content (or another content) can do so. This information can be inserted at the form of a separate watermark at the very beginning of the audio file (e.g., a prelistening phase of limited duration that can be played without control).

The watermark is used to enable tracking any eventual illegal redistribution over the network. We note that it plays a central role here as a supplement to encryption-based security. And, in fact, the system security relies mostly on the embedded watermark once the content is decrypted. This is because cryptographic protection is of no use if an unauthorized user ever recovers the content after its decryption (e.g., at the mobile device of a legitimate user or by an attacker). Worse, here, cryptographic protection alone does not even prevent a legitimate user from redistributing the content at will once it is deciphered at the used mobile device. Thus, a decrypted content from which the watermark is removed (and the DRM header is stripped off) could be redistributed at will over the network, since it would appear as free uncontrolled (i.e., nonproprietary) content at the mobile device of any receiver.

The aforementioned discussion shows that the embedding of the watermark should be such that its removal (e.g., by an attacker) is difficult. This can be achieved by having the network provider filter the signals sent to and from each terminal, blocking unauthorized signals at some specific base stations and possibly reinforcing the watermarks in the legitimate signals at some other base stations. We will comment on this in Remark 1 and Remark 2.

Suppose, for example, that user 1 sends the audio content to user 2 (after DRM header liftoff and decryption) through the mobile network. On its way to user 2, the watermarked content travels through base station 2 (BS 2). Hence, watermark control at BS 2 is really effective. More specifically, if a watermark is detected at BS 2, but it does not match the one that should be present, this means that the content is pirated (by user 1 or an attacker), and the content is blocked at BS 2. If the correct watermark is detected at BS 2, the content is authorized to go through the network and, in this case, the mobile operator implements a watermark reinforcement at base station 3 (BS 3). For example, BS 3 decodes the embedded message from the signal received from BS 2 on the fly and sends a watermark that carries the same message to user 2. This watermark reinforcement at BS 3 is really effective against the effects of channel impairments on the watermark control operated at the user side. More specifically, without this reinforcement, the watermark that is embedded in the signal received at the mobile device of user 2 directly from BS 2 could be damaged or even removed, with one of the following two effects: 1) in case the watermark is removed, the content could be played freely at the mobile device of user 2 without payment of the license; and 2) in case the watermark is damaged, the rendering module at the mobile device of user 2 would not play the content sent by user 1 even if user 2 acquires the appropriate usage rights from the server (i.e., becomes a legitimate user), because of the discrepancy with the watermark that is in the smart card that would be detected in this case.

Remark 1: As we explained previously, the content owner or provider and the network provider cooperate to secure the audio content redistribution by means of collaborative watermark embedding; this is beneficial for both partners. Furthermore, it should be noted that this can be accomplished without necessarily sharing all of the secret information. For example, let m_s be the message to be hidden by the owner. The owner can obtain another message $m = \Theta_{\mathbf{k}_s}(m_s)$ from the message m_s, where $\Theta_{\mathbf{k}_s}(\cdot)$ is a function which depends on a secret key $\mathbf{k}_s$ that is not revealed to the network provider (i.e., the helper). Then, the owner embeds the message m into the host signal by using another key $\mathbf{k}$. Thus, for the network operator to strengthen the embedded message at the base station, it needs to only know the key $\mathbf{k}$, not the key $\mathbf{k}_s$; and so, it can know the message m but not the original secret message m_s. The destination knows both keys. It decodes the message m by using the key $\mathbf{k}$, and then it obtains the secret message m_s using the key $\mathbf{k}_s$.

Remark 2: For the watermark control, the network provider needs to know the watermark that was originally embedded by the owner only at those base stations which handle all of the traffic (e.g., BS 2). At these base stations, the watermark can be obtained from the server through a secure link similar to that with legitimate users at the delivery of the license. Also, for security purposes, in this example, the network provider reinforces only the watermarks in the legitimate signals (i.e., after a prerequisite watermark control). In nonsecurity-oriented applications of IE, however, the watermark reinforcement could, in principle, be implemented irrespective to this control.

IV. Mathematical Models for Relay-Assisted Information Embedding

In this section, we formalize the notion of cooperative embedding. First, we give an informal description of the basic setup, and then we describe two mathematical models through connection to the relay channel.

A. Basic Setup

We illustrate the main concept as follows. A primary user, or initial embedder, embeds some watermark into the host signal and sends the watermarked signal over the wireless network.

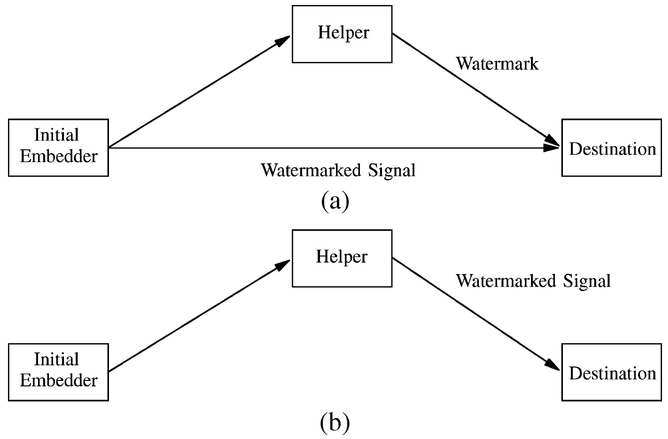


Fig. 4. Possible strategies at the helper: (a) sending (only) a watermark or (b) sending a watermarked signal in which the watermark is reinforced.

This watermarked signal carries some hidden information which is intended to some receiver. The role of the receiver is to decode the embedded message and to convey the hidden information. The transmission is possibly corrupted by a channel attack, the effect of which is assumed to be equivalent to additive noise. If at some point of the network another user or partner (e.g., the network provider at the base station, as in Section III) can determine the embedded message from the watermarked signal sent by the initial embedder, then the former can reencode the hidden information into another watermark that carries the same message and sends it to the destination over the network, using all or part of the available power. The destination receives the sum of the signal watermarked by the initial embedder and the watermark sent by the secondary embedder or helper, possibly corrupted by some channel noise. The embedded information is reinforced since it also benefits from the power at the helper. This strategy is shown in Fig. 4(a).

We note that an alternative strategy, not considered in this paper, consists in sending the watermarked signal to the destination in two phases: from the initial embedder to the helper first, and then from the helper to the destination, as shown in Fig. 4(b). In this case, the helper sends a watermarked signal which carries a reinforced watermark, not only a watermark. However, we mention that, in general, this alternative strategy is suboptimal [w.r.t. the one shown in Fig. 4(a)], for the two scenarios that we will consider in this paper.

B. Channel Models

Let $W \in \mathcal{W}$ be the message to be embedded cooperatively into the host signal $\mathbf{S}$. We denote by P the maximum tolerable embedding distortion (i.e., the distortion above which a watermark could be perceived) (the value of P is usually determined by using some psycoacoustic or psycovisual models). The cooperative embedding and transmission scheme is as follows. The initial embedder encodes the message W into a watermark $\mathbf{X}_1$ and sends the watermarked signal $\mathbf{X}_1 + \mathbf{S}$ over the wireless network. Let P_1 denote the distortion (per-sample) caused on the host signal by embedding at the initial embedder (i.e., $(1/n)\mathbb{E}[||\mathbf{X}_1||^2] = P_1 \leq P$). The sent watermarked signal is received as $\mathbf{Y}_1$ at the helper. The helper decodes the embedded message and then reencodes it into a watermark $\mathbf{X}_2$ that it sends over the network[1] using some power P_2. The power P_2 should be such that $(\sqrt{P_1} + \sqrt{P_2})^2 \leq P$. The destination receives the sum of the watermarked signal sent by the initial embedder and the watermark sent by the helper, possibly corrupted by some channel noise. Thus, the watermarked signal received at the destination contains the watermark $\mathbf{X}_1$ embedded by the initial embedder and the watermark $\mathbf{X}_2$ sent by the helper. The two watermarks carry the same message; and they satisfy a joint embedding distortion constraint $(\sqrt{P_1} + \sqrt{P_2})^2 \leq P$ for imperceptibility reasons at the destination.

Let $\mathbf{Z}_1$ denote the noise corrupting the watermarked signal sent by the initial embedder on its way to the helper, and $\mathbf{Z}$ the noise corrupting it on the way to the destination. We assume that $\mathbf{Z}_1$ and $\mathbf{Z}$ are mutually independent and independent of the host signal $\mathbf{S}$, with $\mathbf{Z}_1 \sim \mathcal{N}(0, N_1 I)$ and $\mathbf{Z} \sim \mathcal{N}(0, NI)$. The host signal also is assumed to be i.i.d Gaussian, with $\mathbf{S} \sim \mathcal{N}(0, QI)$. Then, the watermarked signals received at the helper and the destination are given by

$$\mathbf{Y}_1 = \mathbf{X}_1 + \mathbf{S} + \mathbf{Z}_1 \tag{6a}$$
$$\mathbf{Y} = \mathbf{X}_1 + \mathbf{X}_2 + \mathbf{S} + \mathbf{Z} \tag{6b}$$

respectively.

1) Embedding Distortion: In general, as is shown for the two scenarios considered in Sections V and VI, it is advantageous that the two watermarks by the initial embedder and the helper be correlated. Such correlation makes the cooperative embedding more effective. The induced distortion (per-sample) is given by

$$D_E = \frac{1}{n}\mathbb{E}\left[||\mathbf{X}_1 + \mathbf{X}_2||^2\right]. \tag{7}$$

The maximum value of this distortion is obtained if the two watermarks are proportional (i.e., $\mathbf{X}_2 = \sqrt{P_2/P_1}\mathbf{X}_1$). In this case, the distortion is given by

$$D_{E_{\max}} = (\sqrt{P_1} + \sqrt{P_2})^2. \tag{8}$$

2) Connection to the Relay Channel: The model (6) has close connection with transmission over a state-dependent relay channel (RC) [16]. More precisely, if the helper knows the original host signal, (6) can be viewed as the input–output relation of transmission over a Gaussian RC with SI noncausally known at the source and the relay—the initial embedder playing the role of the source and the helper playing the role of the relay. Similarly, if the helper does not know the original host signal, (6) can be viewed as the input–output relation of transmission over a Gaussian RC with SI noncausally known at only the source.

Moreover, if the noise at the destination can be decomposed as $\mathbf{Z} = \mathbf{Z}_1 + \mathbf{Z}_2$, with $\mathbf{Z}_2 \sim \mathcal{N}(0, (N - N_1)I)$, $N > N_1$, and $\mathbf{Z}_1$, and $\mathbf{Z}_2$ are independent of each other and independent

[1] We assume that the helper receives and transmits at the same time. The results in this paper can be easily specialized to the case in which the helper receives only during a first period, and then transmits only during a second period.

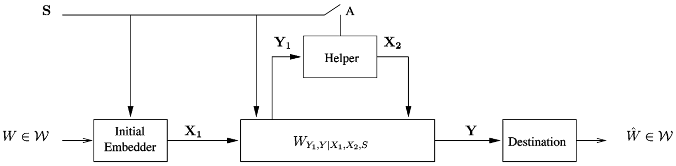


Fig. 5. Relay-assisted IE. The helper may (switch A on) or may not (switch A off) know the original host signal $\mathbf{S}$.

of $\mathbf{S}$, the watermarked signal received at the destination can be written as

$$\begin{aligned}\mathbf{Y} &= (\mathbf{X}_1 + \mathbf{S} + \mathbf{Z}_1) + \mathbf{X}_2 + \mathbf{Z}_2\\ &= \mathbf{Y}_1 + \mathbf{X}_2 + \mathbf{Z}_2. \qquad (9)\end{aligned}$$

In this case, we will say that the cooperative IE channel is degraded by reference to the degradedness of the corresponding Gaussian RC. At a high level, degradedness in cooperative IE means that the destination observes a watermarked content which is more noisy than the one observed by the helper. A block diagram of the general, not necessarily degraded, cooperative IE channel is shown in Fig. 5.

V. Host Known to the Helper

In this section, we study the scenario in which the helper knows the original host signal. First, we derive lower and upper bounds on the ultimate cooperative embedding rate, and then we design a feasible cooperative embedding scheme and analyze the rate allowed by this scheme.

A. Bounds on the Ultimate Embedding Rate

For the Gaussian cooperative IE channel (6) for which the helper knows the original host signal, the following embedding rate in Theorem 1 is achievable (subscript "11" refers to the fact that the initial embedder and the helper know the original host signal).

Theorem 1: The capacity of the Gaussian cooperative information embedding channel (6) for which the helper knows the original host signal is lower-bounded by

$$\mathcal{C}_{11}^{\text{in}} = \max_{0\le\beta\le1} \min\left\{\frac{1}{2}\log\left(1+\frac{\beta P_1}{N_1}\right)\right.$$
$$\left.\frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}P_1P_2}}{N}\right)\right\}. \qquad (10)$$

Proof: The proof consists in computing the lower bound for the discrete memoryless (DM) case in Lemma 1 below using an appropriate jointly Gaussian input distribution that will be specified in the sequel. The result for the DM case can be readily extended to memoryless channels with discrete time and continuous alphabets using standard techniques [30].

Lemma 1: The capacity of the discrete memoryless cooperative embedding channel shown in Fig. 5 in the case in which the helper knows the original host signal is lower-bounded by

$$R = \max_{p(u_1,u_2,x_1,x_2|s)} \min\{I(U_1;Y_1|S,U_2)$$
$$I(U_1,U_2;Y) - I(U_1,U_2;S)\} \qquad (11)$$

where U_1 and U_2 are auxiliary random variables with bounded cardinality.

Proof: The proof of Lemma 1 follows straightforwardly from that for the corresponding RC with side information known noncausally at the source and the relay [31], [32]. Also, this proof is based on a random code construction which is similar to that used for the proof of Lemma 2 in Appendix A and, hence, we omit it for brevity here.

The rest of the proof of Theorem 1 follows through straightforward algebraic manipulations similar to those in [17] to show that the evaluation of the rate (11) with a jointly Gaussian $\mathbf{S}, \mathbf{U}_1, \mathbf{U}_2, \mathbf{X}_1, \mathbf{X}_2$ such that

$$\mathbf{U}_1|\mathbf{S} \sim \mathcal{N}\left(\alpha_1\mathbf{S}, P^{(1)}I\right) \qquad (12a)$$

$$\mathbf{U}_2|\mathbf{S} \sim \mathcal{N}\left(\alpha_2\mathbf{S}, P^{(2)}I\right) \qquad (12b)$$

$$\mathbf{X}_1 = (\mathbf{U}_1 - \alpha_1\mathbf{S}) + \theta(\mathbf{U}_2 - \alpha_2\mathbf{S}) \qquad (12c)$$

$$\mathbf{X}_2 = (1-\theta)(\mathbf{U}_2 - \alpha_2\mathbf{S}) \qquad (12d)$$

with

$$P^{(1)} = \beta P_1, \quad P^{(2)} = (\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2, \quad 0 \le \beta \le 1$$
$$\theta = \frac{\sqrt{\bar{\beta}P_1}}{\sqrt{P^{(2)}}}, \quad \alpha_k = \frac{P^{(k)}}{P^{(1)}+P^{(2)}+N} \quad k=1,2$$

gives the rate (10) in Theorem 1. ■

From (12), we see that the initial embedder employs a combination of two standard DPCs; one DPC to generate $\mathbf{U}_1$ and another DPC to generate $\mathbf{U}_2$. The helper employs only the DPC that is used to generate $\mathbf{U}_2$. More precisely, the watermark embedded by the initial embedder can be written as $\mathbf{X}_1 = \mathbf{X}_1^{(1)} + \mathbf{X}_1^{(2)}$, where $\mathbf{X}_1^{(1)} = \mathbf{U}_1 - \alpha_1\mathbf{S}$ is i.i.d. Gaussian with power $P^{(1)} = \beta P_1$ for some $0 \le \beta \le 1$ and $\mathbf{X}_1^{(2)} = \theta(\mathbf{U}_2 - \alpha_2\mathbf{S})$ is i.i.d. Gaussian with power $\bar{\beta}P_1$. Also, the watermark by the initial embedder and the one by the helper are correlated through $\mathbf{X}_1^{(2)}$, with correlation coefficient $\sqrt{\bar{\beta}} = \sqrt{1-\beta}$. The cooperative part $(\mathbf{X}_1^{(2)} + \mathbf{X}_2) = \mathbf{U}_2 - \alpha_2\mathbf{S}$ of the watermark has power $(\sqrt{P_2} + \sqrt{\bar{\beta}P_1})^2$.

We now comment on the minimization in (10) and on the correlation between the two watermarks $\mathbf{X}_1$ and $\mathbf{X}_2$. First, note that the first term of the minimization in (10) represents the information rate that the helper can decode reliably from the watermarked signal received from the initial embedder. Also, the second term of the minimization in (10) represents the information rate that the destination can decode reliably from the watermarked signal received from the initial embedder and the watermark sent by the helper. Then, the cooperative embedding rate over the channel (6) is given by the minimum of the two information rates and basically expresses that in our setup, the signal sent by the helper actually reinforces the watermark by the initial embedder only if the helper has correctly decoded the message embedded by the initial embedder; otherwise, the signal sent by the helper would look like additional noise at the destination.

Furthermore, from (10), it is easy to see that highly correlated watermarks (i.e., small β) result in a small information rate at the helper and a large information rate at the destination; and weakly correlated watermarks result in the reverse situation. For a given configuration (P_1, P_2, N_1, N), the optimum value of the correlation coefficient $\sqrt{\bar{\beta}}$ can be obtained by equaling the two terms of the minimization in (10). Also, for fixed β, the embedding distortion (per-sample) induced to the host signal at the destination is such that

$$D_E = P_1 + P_2 + 2\sqrt{\bar{\beta}P_1P_2} \leq D_{E_{\max}} = (\sqrt{P_2} + \sqrt{P_1})^2. \tag{13}$$

We now provide an upper bound on the embedding rate by specializing the cutset bound to the Gaussian case.

Proposition 1: The capacity of the Gaussian cooperative information embedding channel (6) for which the helper knows the original host signal is upper-bounded by

$$\begin{aligned} \mathcal{C}_{11}^{\text{out}} &= \max\min\{I(X_1;Y_1,Y|S,X_2), I(X_1,X_2;Y|S)\} \\ &= \max_{0\leq\beta\leq1}\min\left\{\frac{1}{2}\log\left(1+\frac{\beta P_1}{N_1}+\frac{\beta P_1}{N}\right),\right. \\ &\qquad\left.\frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}P_1P_2}}{N}\right)\right\}. \end{aligned} \tag{14}$$

In addition, if the cooperative embedding channel is degraded, the first term of the minimization in (14) can be replaced by $(1/2)\log(1+\beta P_1/N_1)$.

From proposition 1, it can be easily seen that the lower bound in Theorem 1 meets with the cut-set bound and, thus, gives channel capacity, in the special case in which the cooperative IE channel is degraded Gaussian.

Remark 3: We note that an alternative proof of the achievability of rate (10), established for the Gaussian RC with noncausal SI at the source and the relay, is given in [33]. In [33], the authors also give (without proof) a lower bound on channel capacity for the DM case. For the DM case, the achievable rate in [33] differs from (11) in that the first term of the minimization in (11) (i.e., the conditional mutual information $I(U_1;Y_1|S,U_2)$) is replaced by $I(U_1;Y_1|S,X_2)$ in [33]. For the Gaussian case, an optimal choice of the input X_2 is given by a linear combination of the auxiliary random variable U_2 and the state S, as given by (12), and the two achievable rates are equal.

B. *Coding Realization and Feasible Rate*

We now describe a coding realization of the cooperative embedding scenario studied in this section. This coding scheme is based on lattices and MMSE scaling.

Let $W \in \mathcal{W}$ be the message to be embedded. Consider a lattice Λ with volume $V(\Lambda)$ and normalized second moment $G(\Lambda)$ such that

$$G(\Lambda)V(\Lambda)^{\frac{2}{n}} = P_1. \tag{15}$$

We will use the lattices $\Lambda_1 = \sqrt{\beta}\Lambda$ and $\Lambda_2 = ((\sqrt{P_2} + \sqrt{\bar{\beta}P_1})/\sqrt{P_1})\Lambda$, which are appropriate scaled versions of lattice Λ chosen such that modulo-Λ_1 and modulo-Λ_2 reductions provide quantization error signals with powers βP_1 and $(\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2$, respectively. Their normalized second moments and volumes are given by

$$G(\Lambda_1) = G(\Lambda),\ V(\Lambda_1)^{\frac{2}{n}} = \beta V(\Lambda)^{\frac{2}{n}} \tag{16a}$$

$$G(\Lambda_2) = G(\Lambda),\ V(\Lambda_2)^{\frac{2}{n}} = \frac{(\sqrt{\bar{\beta}P_1}+\sqrt{P_2})^2}{P_1}V(\Lambda)^{\frac{2}{n}}. \tag{16b}$$

Also, we use a cryptography key $\mathbf{K}$, which stands for a randomized codebook and is known to the initial embedder, the helper, and the destination; and a mapping function $\iota(\cdot)$ which one-to-one associates messages $\{W\}$ to coset leaders $\{\mathbf{c}_w\}$, all defined as in Section II-B. The key $\mathbf{K}$ is used for security purposes [7], [25] and, in addition, for reasons that will become clear in the sequel, we choose this key to be uniformly distributed over $\mathcal{V}(\Lambda)$ [26].

1) Cooperative Embedding Scheme: The cooperative embedding scheme is as follows. The initial embedder embeds the watermark

$$\mathbf{x}_1(W;\mathbf{s},\Lambda) = \mathbf{x}_1^{(1)} + \mathbf{x}_1^{(2)} \tag{17}$$

with

$$\mathbf{x}_1^{(1)} = [\mathbf{c}_w + \mathbf{k} - \alpha_1\mathbf{s}] \bmod \Lambda_1 \tag{18}$$

$$\mathbf{x}_1^{(2)} = \theta\left([\mathbf{c}_w + \mathbf{k} - \alpha_2(1-\alpha_1)\mathbf{s}] \bmod \Lambda_2\right) \tag{19}$$

and $\theta = \sqrt{\bar{\beta}P_1}/(\sqrt{\bar{\beta}P_1} + \sqrt{P_2})$, for some $\beta \in [0,1], \alpha_1 \in [0,1]$ and $\alpha_2 \in [0,1]$.

Upon reception of the watermarked signal $\mathbf{y}_1 = \mathbf{x}_1 + \mathbf{s} + \mathbf{z}_1$, the helper first decodes the embedded message and then sends the watermark

$$\mathbf{x}_2(W;\mathbf{s},\Lambda) = (1-\theta)\left([\mathbf{c}_w + \mathbf{k} - \alpha_2(1-\alpha_1)\mathbf{s}] \bmod \Lambda_2\right). \tag{20}$$

The decoding procedure at the helper basically consists in computing the signal

$$\tilde{\mathbf{y}}_1 = \mathbf{y}_1 - \left(\mathbf{x}_1^{(2)} + \mathbf{s}\right) = \mathbf{x}_1^{(1)} + \mathbf{z}_1 \tag{21}$$

an operation which transforms the channel from the initial embedder to the helper into an interference-free channel from which the helper can decode the embedded message using standard techniques (e.g., MMSE decoding).

The destination receives the watermarked signal $\mathbf{y} = \mathbf{x}_1^{(1)} + (\mathbf{x}_1^{(2)} + \mathbf{x}_2) + \mathbf{s} + \mathbf{z}$ and, without knowing the host signal $\mathbf{s}$, has to determine the embedded message. The destination first decodes the information carried by the part $\mathbf{x}_1^{(1)}$ of the watermark, subtracts it off, and then decodes the information carried by the cooperative part $(\mathbf{x}_1^{(2)} + \mathbf{x}_2)$ of the watermark. To this end, the decoder at the destination computes the error signal $\mathbf{y}' = [\alpha_1\mathbf{y} - \mathbf{k}] \bmod \Lambda_1$ and then the error signal $\tilde{\mathbf{y}}' = [\alpha_2\tilde{\mathbf{y}} - \mathbf{k}] \bmod \Lambda_2$, with $\tilde{\mathbf{y}} = \mathbf{y} - (\mathbf{x}_1^{(1)} + \alpha_1\mathbf{s})$.[2]

2) Performance Analysis: We focus on the embedding rate allowed by this scheme. For decoding at the helper, which utilizes standard MMSE decoding, it is easy to see that the helper can decode reliably as long as the message at the initial embedder is embedded at a rate less than the mutual information of the interference-free channel (21), i.e., $0.5\log(1+(\beta P_1/N_1))$.

For decoding at the destination, which utilizes modulo-reduction operations, the establishment of the results below relies

[2] We note that although the destination does not know the original host $\mathbf{s}$ and the signal $\mathbf{x}_1^{(1)}$ is a continuum, dirty paper decoding permits knowing the codeword $\mathbf{u}_1 = \mathbf{x}_1^{(1)} + \alpha_1\mathbf{s}$.

principally on the properties of a modulo lattice additive noise (MLAN) channel [34] and on standard properties of the mod-Λ operation [18]. More specifically, it can be easily shown that the error signals computed at the destination can be written as

$$\mathbf{y}' = \left[\mathbf{c}_w + \alpha_1\left(\mathbf{x}_1^{(2)} + \mathbf{x}_2 + \mathbf{z}\right) - (1-\alpha_1)\mathbf{x}_1^{(1)}\right] \bmod \Lambda_1 \quad (22a)$$

$$\tilde{\mathbf{y}}' = \left[\mathbf{c}_w + \alpha_2\mathbf{z} - (1-\alpha_2)\left(\mathbf{x}_1^{(2)} + \mathbf{x}_2\right)\right] \bmod \Lambda_2. \quad (22b)$$

Thus, in the decoding procedure, the destination sees an active channel noise given by the Λ_1-aliased noise $\mathbf{Z}' = [\alpha_1(\mathbf{Z} + \mathbf{X}_1^{(2)} + \mathbf{X}_2) - (1-\alpha_1)\mathbf{X}_1^{(1)}] \bmod \Lambda_1$ in decoding the information carried by the part $\mathbf{X}_1^{(1)}$ of the watermark, and by the Λ_2-aliased noise $\tilde{\mathbf{Z}}' = [\alpha_2\mathbf{Z} - (1-\alpha_2)(\mathbf{X}_1^{(2)} + \mathbf{X}_2)] \bmod \Lambda_2$ in decoding the cooperative part $\mathbf{X}_1^{(2)} + \mathbf{X}_2$ of the watermark. Since the key $\mathbf{K}$ is uniformly distributed over $\mathcal{V}(\Lambda)$, its use as a dither ensures that the inputs $\mathbf{X}_1^{(1)}$, $\mathbf{X}_1^{(2)}$, and $\mathbf{X}_2$ are uniformly distributed regardless of the power Q of the host signal. Thus, the noises $\mathbf{Z}'$ and $\tilde{\mathbf{Z}}'$ are statistically independent of $\mathbf{C}_w$. Then, using the inflated Lattice lemma reported in [26], this gives rise to two MLAN channels: the channel from $\mathbf{C}_w$ to $\mathbf{Y}'$ and the channel from $\mathbf{C}_w$ to $\tilde{\mathbf{Y}}'$. The mutual information of these channels is maximized by a uniform input [35], [36], giving

$$\begin{aligned}\frac{1}{n}I(W;\mathbf{Y}) &= \frac{1}{n}\left[h(\mathbf{Y}') - h(\mathbf{Z}')\right]\\ &= \frac{1}{n}\left[\log V(\Lambda_1) - h(\mathbf{Z}')\right]\\ \frac{1}{n}I(W;\tilde{\mathbf{Y}}) &= \frac{1}{n}\left[h(\tilde{\mathbf{Y}}') - h(\tilde{\mathbf{Z}}')\right]\\ &= \frac{1}{n}\left[\log V(\Lambda_2) - h(\tilde{\mathbf{Z}}')\right] \quad (23)\end{aligned}$$

where $h(\cdot)$ denotes differential entropy.

Finally, using (15) and (16) to substitute the volumes $V(\Lambda_1)$ and $V(\Lambda_2)$ in (23), the rate allowed by the proposed coding realization can be obtained by taking the minimum of the information that the helper can decode reliably from the signal received from the initial embedder (i.e., $0.5\log(1+\beta P_1/N_1)$) and the information that the destination can decode reliably from the signals received from the initial embedder and the helper (i.e., the mutual information sum $(1/n)[I(W;\mathbf{Y}) + I(W;\tilde{\mathbf{Y}})]$). This gives the embedding rate

$$\begin{aligned}R_{11}(\Lambda) = \max\min\Bigg\{&\frac{1}{2}\log\left(1+\frac{\beta P_1}{N_1}\right), \frac{1}{2}\log\left(\frac{\beta P_1}{G(\Lambda)}\right)\\ &+\frac{1}{2}\log\left(\frac{(\sqrt{\bar{\beta}P_1}+\sqrt{P_2})^2}{G(\Lambda)}\right) - \frac{1}{n}h(\mathbf{Z}') - \frac{1}{n}h(\tilde{\mathbf{Z}}')\Bigg\} \quad (24)\end{aligned}$$

where the maximization is over parameters $\beta \in [0,1]$, $\alpha_1 \in [0,1]$, and $\alpha_2 \in [0,1]$.

In general, no closed-form expression can be derived for (24) and the computation of differential entropy and maximization over all possible choices of the tuple $(\beta, \alpha_1, \alpha_2)$ have to be performed numerically.[3] The active channel noises $\mathbf{Z}'$ and $\tilde{\mathbf{Z}}'$ are not Gaussian. An approximation of the rate $R_{11}(\Lambda)$ in (24) (and, in fact, a lower bound) can be obtained by replacing these noise terms by the restrictions to the Voronoi regions of the associated lattices of Gaussian noises with equal first and second moments. More precisely, let $\bar{\mathbf{Z}}$ be the restriction to $\mathcal{V}(\Lambda_1)$ of the noise distributed as $\mathcal{N}(0, \sigma^2 I)$ and $\bar{\mathbf{Z}}'$ the restriction to $\mathcal{V}(\Lambda_2)$ of the noise distributed as $\mathcal{N}(0, \sigma'^2 I)$, with

$$\sigma^2 = \alpha_1^2\left[N + (\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2\right] + \beta(1-\alpha_1)^2 P_1 \quad (25a)$$

$$\sigma'^2 = \alpha_2^2 N + (1-\alpha_2)^2(\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2. \quad (25b)$$

Replacing $\mathbf{Z}'$ and $\tilde{\mathbf{Z}}'$ on the right-hand side of (24) by $\bar{\mathbf{Z}}$ and $\bar{\mathbf{Z}}'$, respectively, one obtains an approximation of the rate $R_{11}(\Lambda)$.[4] However, we note that the computation of the differential entropy terms in this approximation still has to be performed numerically since the noises $\bar{\mathbf{Z}}$ and $\bar{\mathbf{Z}}'$ also are non-Gaussian; they are (only) the restrictions of Gaussian distributions (for similar approaches for computing numerically the rates obtained with lattice codes, the reader may refer to [3] and [27]).

Asymptotic Case: Taking $\alpha_1 = \beta P_1/[\beta P_1 + (\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2 + N]$ and $\alpha_2 = (\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2/[(\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2 + N]$ on the right-hand side of (25), we see that the noise terms $\mathbf{Z}'$ and $\tilde{\mathbf{Z}}'$ (without the effect of the modulo front end) have variances given by $\alpha_1(N + (\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2)$ and $\alpha_2 N$, respectively (per dimension). Then, since for a given second moment a Gaussian random vector has the largest entropy, it follows that:

$$\begin{aligned}&\frac{1}{n}h(\mathbf{Z}') \le \log\left(2\pi e\alpha_1\left[N + (\sqrt{\bar{\beta}P_1} + \sqrt{P_2})^2\right]\right)\\ &\frac{1}{n}h(\tilde{\mathbf{Z}}') \le \log(2\pi e\alpha_2 N). \quad (26)\end{aligned}$$

Substituting (26) in (24), we obtain

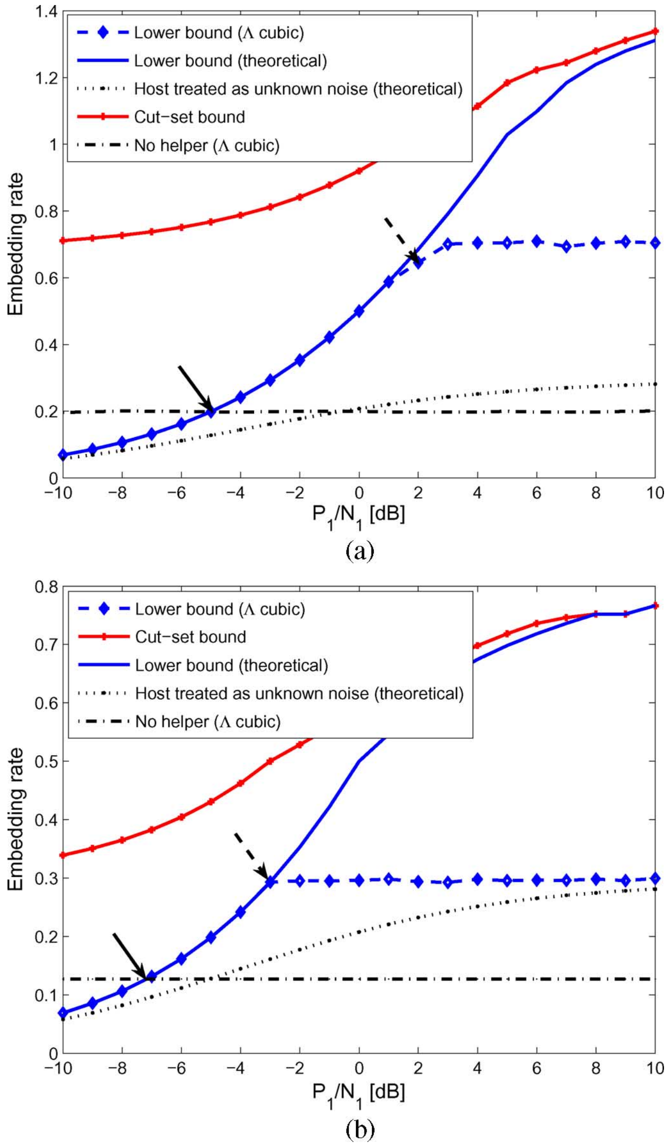
$$\begin{aligned}&R_{11}(\Lambda) \ge \max_{0\le\beta\le 1}\min\left\{\frac{1}{2}\log\left(1+\frac{\beta P_1}{N_1}\right),\right.\\ &\left.\frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{\bar{\beta}P_1P_2}}{N}\right) - \log 2\pi e G(\Lambda)\right\}. \quad (27)\end{aligned}$$

From (27), we see that the gap to the theoretical embedding rate (10) may be made smaller than $\log 2\pi e G(\Lambda)$. For optimal lattices for quantization, we have $G(\Lambda) \to 1/2\pi e$, and this gap goes to zero.

3) Numerical Examples and Discussion: We compute numerically the rate (24) for the cubic lattice ($\Lambda = \mathbb{Z}^n$). We use Monte-Carlo integration for the computation of the differential entropy terms. The results are depicted in Fig. 6 against the watermark-to-noise ratio $\mathrm{WNR}_1 = 10\log_{10}(P_1/N_1)$ [dB] at the helper, for two examples of $\mathrm{WNR}_{\max} = 10\log_{10}(D_{E_{\max}}/N) = 10\log_{10}((\sqrt{P_1} + \sqrt{P_2})^2/N)$ [dB] at the destination: weak channel noise ($\mathrm{WNR}_{\max} = 8$ dB) and relatively strong channel noise ($\mathrm{WNR}_{\max} = 3$ dB). For the second example, the noise may model a Gaussian attack with a variance equal to half the maximum embedding distortion, i.e., $D_{E_{\max}}/2 = (\sqrt{P_1} + \sqrt{P_2})^2/2$. For comparison reasons, Fig. 6 also shows the theoretical embedding rate (10), the cut-set bound (14), and the rate obtained by treating the host signal as unknown additive Gaussian noise. Fig. 6 also shows

[3] Note that in the maximization in (24), the parameter α_1 appears only in $h(\mathbf{Z}')$ and the parameter α_2 appears only in $h(\tilde{\mathbf{Z}}')$.

[4] The accuracy of this approximation is related to the divergence function $D(\cdot\|\cdot)$ [16]. For example, one has $h(\bar{\mathbf{Z}}) - h(\mathbf{Z}') = D(\mathbf{Z}'\|\bar{\mathbf{Z}})$.



(a)

(b)

Fig. 6. Cooperative embedding rate as a function of WNR_1 at the helper, for two examples of $\mathrm{WNR}_{\max}$ at the destination (a) $\mathrm{WNR}_{\max} = 8$ dB and (b) $\mathrm{WNR}_{\max} = 3$ dB. Illustration of the rate (10) (solid, no marker) and the rate (24) obtained with a cubic lattice. Numerical values are $P_2 = P_1$ and $Q = 2P_1$.

the rate obtained in the standard case in which there is no helper (this rate is readily given by the second term of the minimization in (24) in which we put $P_2 = 0$).

The results shown in Fig. 6 motivate the following discussion.

1) As a general comment, we observe that the collaborative embedding by the helper improves the embedding rate (w.r.t., to the standard case in which there is no helper), for a large range of WNR_1. Furthermore, this improvement increases with WNR_1; this is due to the fact that the less noisy the received watermarked signal at the helper is, the more reliable the decoding is at the helper, and so, the more efficient the collaborative embedding is. For large values of WNR_1, the collaborative embedding rate increases only slightly with WNR_1 and is due to the fact that at this range, the embedding rate is determined by how much information the initial embedder and the helper together can embed into the host signal (i.e., the second term of the minimization), and so, it depends essentially on the value of $\mathrm{WNR}_{\max}$ at the destination.

2) Though relatively important, the gap of the feasible rate (24) to the theoretical rate (10) should not be considered as such since, obviously, this gap depends on the value of the operating $\mathrm{WNR}_{\max}$ (observe that this gap is more important if the watermarked signal is more noisy; see the case $\mathrm{WNR}_{\max} = 3$ dB) and also because lattice Λ is set to its simplest (cubic) form for the curves in Fig. 6. Larger feasible embedding rates can be obtained by using lattices with better quantizing properties as we mentioned previously, at the expense of some increase in the computational complexity however.

3) The assumption for the helper to decode the message embedded by the initial embedder prior to reencoding makes collaborative embedding efficient only if WNR_1 is larger than a certain threshold (indicated by the solid arrow in Fig. 6). Below this threshold, the watermark sent by the helper is rather noise-like and can even introduce a decoding ambiguity at the destination. Yet, it is not that cooperative IE is not worth it for such a range of WNR_1, but rather, that the strategy at the helper that requires decoding the embedded message as a prerequisite for the collaborative embedding is a severe constraint which is clearly suboptimal in this case. For very small values of WNR_1, not reinforcing the watermark by the helper may be a better solution.

4) It is interesting to observe that with the described coding realization, reinforcing the watermark by the helper may be advantageous (over not reinforcing the watermark) even for certain values of WNR_1 for which the helper receives a watermarked signal which is more noisy than the one received at the destination (the value of WNR_1 for which $N_1 = N$ is indicated by the dashed arrow). We mention that in this case, the helper can still decode the embedded message more reliably than the destination; this is enabled by the structure of the decoder which benefits from the knowledge of the host signal at the helper to apply MMSE, whereas it is based on modulo-$\mathbb{Z}^n$ reduction operation at the destination.

VI. Host Unknown to the Helper

In this section, we study the important scenario in which the helper does not know the original host signal. First, we derive a lower bound on the ultimate cooperative embedding rate; and then, we design a feasible cooperative embedding scheme and analyze the rate allowed by this scheme.

For convenience, we define the following two functions $r_1(\cdot)$ and $r_2(\cdot)$ which we will use throughout this section.

Definition 1: Let

$$\begin{aligned}
Q'(\rho) &:= (\sqrt{Q} - \sqrt{\rho P_1})^2 \\
\Delta_1(\rho,\beta,\alpha) &:= \beta\bar{\rho}P_1\left(\beta\bar{\rho}P_1 + Q'(\rho) + N_1\right) \\
\Delta_2(\rho,\beta,\alpha) &:= (1-\alpha)^2\beta\bar{\rho}P_1Q'(\rho) + N_1\left(\beta\bar{\rho}P_1 + \alpha^2Q'(\rho)\right) \\
\Delta_3(\rho,\beta,\alpha) &:= \beta\bar{\rho}P_1\left(\bar{\rho}P_1 + P_2 + Q'(\rho) + 2\sqrt{\bar{\beta}\bar{\rho}P_1P_2} + N\right) \\
\Delta_4(\rho,\beta,\alpha) &:= (1-\alpha)^2\beta\bar{\rho}P_1Q'(\rho) + N\left(\beta\bar{\rho}P_1 + \alpha^2Q'(\rho)\right)
\end{aligned}$$

$$r_1(\rho,\beta,\alpha) := \frac{1}{2}\log\left(\frac{\Delta_1(\rho,\beta,\alpha)}{\Delta_2(\rho,\beta,\alpha)}\right)$$
$$r_2(\rho,\beta,\alpha) := \frac{1}{2}\log\left(\frac{\Delta_3(\rho,\beta,\alpha)}{\Delta_4(\rho,\beta,\alpha)}\right)$$

for given $0 \leq \rho \leq \min(1, Q/P_1)$, $0 \leq \beta \leq 1$, $\alpha \in \mathcal{A}(\rho,\beta) := \{x \in \mathbb{R} : r_1(\rho,\beta,x) \geq 0, r_2(\rho,\beta,x) \geq 0\}$.

A. Lower Bound on the Ultimate Embedding Rate

For the Gaussian cooperative IE channel (6) for which the helper does not know the original host signal, the following embedding rate in Theorem 2 is achievable (the subscript "10" refers to the fact that only the initial embedder knows the host signal).

Theorem 2: The capacity of the Gaussian cooperative information embedding channel (6) for which the helper does not know the original host signal is lower-bounded by

$$\mathcal{C}_{10}^{\text{in}} = \max\min\{r_1(\rho,\beta,\alpha_1), r_2(\rho,\beta,\alpha_1)\} \qquad (28)$$

where the maximization is over parameters $0 \leq \rho \leq \min(1, Q/P_1)$, $0 \leq \beta \leq 1$, $\alpha_1 \in \mathcal{A}(\rho,\beta) := \{x \in \mathbb{R} : r_1(\rho,\beta,x) \geq 0, r_2(\rho,\beta,x) \geq 0\}$ and the functions $r_1(\rho,\beta,\alpha_1)$ and $r_2(\rho,\beta,\alpha_1)$ are defined in Definition 1.

Proof: The proof consists in computing the lower bound for the DM case in Lemma 2 below using an appropriate jointly Gaussian input distribution that will be specified in the sequel. The result for the DM case can readily be extended to memoryless channels with discrete time and continuous alphabets using standard techniques [30].

Lemma 2: The capacity of the discrete memoryless cooperative embedding channel shown in Fig. 5 in the case in which the helper does not know the original host signal is lower-bounded by

$$\begin{aligned} R = \max\min\{&I(U_1;Y_1|X_2) - I(U_1;S|X_2) \\ &I(U_1,X_2;Y) - I(U_1;S|X_2)\} \end{aligned} \qquad (29)$$

where the maximization is over all joint measures $P_{S,U_1,X_1,X_2,Y_1,Y}$ of the form

$$P_{S,U_1,X_1,X_2,Y_1,Y} = P_S P_{X_2} P_{U_1,X_1|S,X_2} W_{Y_1,Y|X_1,X_2,S} \qquad (30)$$

and U_1 is an auxiliary random variable with bounded cardinality.

Proof: The proof of Lemma 2 follows straightforwardly from that for the corresponding RC with noncausal SI at only the source [32]. For reasons of completeness, this proof is reproduced in Appendix A.

Fix $0 \leq \beta \leq 1$, $0 \leq \rho \leq \min(1, Q/P_1)$. Through straightforward algebra which we omit for brevity, it can be shown that the evaluation of the first and second terms in the minimization in (29) with a jointly Gaussian $\mathbf{S}, \mathbf{U}_1, \mathbf{X}_1, \mathbf{X}_2$ such that

$$\mathbf{U}_1 \sim \mathcal{N}(\alpha_1\mathbf{S}', \bar{\rho}P_1 I) \qquad (31a)$$
$$\mathbf{X}_1 = \mathbf{U}_1 - \left(\alpha_1 + \frac{\sqrt{\rho P_1}}{\sqrt{Q} - \sqrt{\rho P_1}}\right)\mathbf{S}' \qquad (31b)$$

with $\mathbf{S}' = (1 - \sqrt{\rho P_1}/\sqrt{Q})\mathbf{S}$ and $\mathbf{X}_2$ is i.i.d. Gaussian with power P_2 independent of the host signal $\mathbf{S}$ and correlated with the input $\mathbf{X}_1$ with $\mathbb{E}[X_{1i}X_{2i}] = \sqrt{\bar{\beta}\bar{\rho}P_1P_2}$, gives $r_1(\rho,\beta,\alpha_1)$ and $r_2(\rho,\beta,\alpha_1)$, respectively. Finally, the embedding rate (28) in Theorem 2 is obtained by maximization over all possible values of the tuple (β,ρ,α_1). For fixed (β,ρ), the allowable values of the parameter α_1 are those such that the terms $r_1(\rho,\beta,\alpha_1)$ and $r_2(\rho,\beta,\alpha_1)$ are nonnegative real. ■.

The following lines provide some insights about the chosen input distribution (31). Since the helper does not know it, the original host signal $\mathbf{S}$ acts as unknown interference for the transmission of the watermark $\mathbf{X}_2$. The initial embedder allocates a fraction ρ of its power P_1 to partially cancel the effect of the interference created by the host $\mathbf{S}$ so that the helper can benefit from this cancellation. Then, the initial embedder uses the remaining power (i.e., $\bar{\rho}P_1$) for pure information embedding. Let $\mathbf{U}_w$ be the watermark embedded by the initial embedder; $\mathbf{U}_w$ carries the embedded message W. Thus, the watermark embedded by the initial embedder is given by

$$\mathbf{X}_1 = -\sqrt{\frac{\rho P_1}{Q}}\mathbf{S} + \mathbf{U}_w \qquad (32)$$

where $\mathbf{U}_w$ is i.i.d. Gaussian with power $\bar{\rho}P_1$ and is independent of the host $\mathbf{S}$. The watermark $\mathbf{X}_2$ sent by the helper is an i.i.d. Gaussian signal with power P_2, and is independent of the host $\mathbf{S}$ and correlated with the watermark $\mathbf{U}_w$, with $\mathbb{E}[U_{wi}X_{2i}] = \sqrt{\bar{\beta}\bar{\rho}P_1P_2}$.

With the aforementioned choice of the embedded signals, an alternative representation of the watermarked signals in (6) is

$$\mathbf{Y}_1 = \mathbf{U}_w + \mathbf{S}' + \mathbf{Z}_1 \qquad (33)$$
$$\mathbf{Y} = \mathbf{U}_w + \mathbf{X}_2 + \mathbf{S}' + \mathbf{Z} \qquad (34)$$

where the effective interference $\mathbf{S}' = (1 - \sqrt{\rho P_1}/\sqrt{Q})\mathbf{S}$ is unknown to the helper but it has less power than the interference $\mathbf{S}$ (i.e., $Q'(\rho) := \mathbb{E}[S_i'^2] = (\sqrt{Q} - \sqrt{\rho P_1})^2 \leq Q$). Then, for the cooperative IE channel (34), the initial embedder applies a standard DPC to produce $\mathbf{U}_w$, by taking $\mathbf{S}'$ as SI as

$$\mathbf{U}_1 \sim \mathcal{N}(\alpha_1\mathbf{S}', \bar{\rho}P_1 I) \qquad (35a)$$
$$\mathbf{U}_w = \mathbf{U}_1 - \alpha_1\mathbf{S}'. \qquad (35b)$$

Finally, combining (32) and (35), we obtain (31).

Remark 4: For the scenario at hand, in essence, we have used a DPC which allows arbitrary negative correlation between the watermark at the initial embedder and the SI (i.e., $\mathbb{E}[X_{1i}S_i] = -\sqrt{\rho P_1 Q}$). This negative correlation can be viewed as a partial cancellation of the known interference. As a consequence, the signal sent by the helper faces less interference on its way to the destination [cf. (34)]. Thus, in a sense, the initial embedder helps the helper so that the latter can assist in embedding in turn. We refer to this strategy as a "generalized" DPC (GDPC). It specializes to standard DPC by putting $\rho = 0$. Both "GDPC" and "standard DPC" can be implemented using lattice codes. We defer the discussion of the implementation issues until we present a brief comparison of the rates allowed theoretically by these two coding schemes.

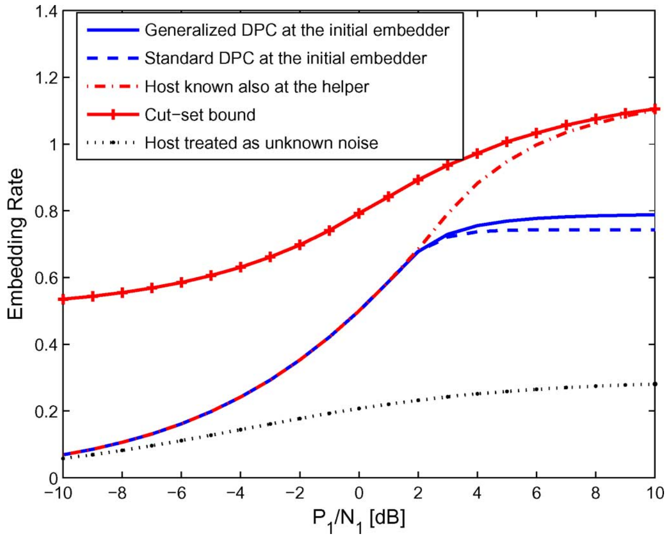


Fig. 7. Illustration of the improvement brought by GDPC over the standard DPC at the initial embedder, as a function of WNR_1 at the helper. Numerical values are $P_1 = P_2 = 1, Q = 2$, and $N = 1$.

1) "Generalized DPC" versus "Standard DPC": This section illustrates the embedding rates obtained with GDPC and standard DPC, with the help of an example. We illustrate the effect of applying GDPC at the initial embedder to improve the cooperative embedding in the case in which the helper does not know the original host signal.

Fig. 7 depicts the evolution of the embedding rate (28) obtained with GDPC at the initial embedder against $\mathrm{WNR}_1 = 10\log_{10}(P_1/N_1)$ [dB]. Also shown for comparison are the rate obtained with standard DPC at the initial embedder, the rate (10), a lower bound obtained by treating the host signal as unknown noise at both the helper and the destination, and the cut-set bound (14).

We see that even though only the initial embedder knows the host signal, both embedders benefit from this knowledge: the helper benefits since its watermark faces less interference and so, the initial embedder benefits in turn. However, the improvement brought by GDPC over standard DPC is mainly visible at large WNR_1 (i.e., when the helper decodes the embedded message correctly). Similar to the scenario discussed in Section V, at such a range of WNR_1, the allowed embedding rate is determined by how much information the two embedders together can embed into the host signal [i.e., the term $r_2(\rho, \beta, \alpha_1)$ in rate (28)]. At small WNR_1, however, the collaborative embedding rate is constrained by the amount of information that the helper can decode reliably [i.e., the term $r_1(\rho, \beta, \alpha_1)$ in rate (28)]. Hence, at small WNR_1, there is no need for the initial embedder to partially cancel the effect of the interference for the helper, since this would be accomplished at the cost of some power that could be used to embed a more powerful watermark that the latter can decode more reliably instead.

Remark 5: An alternative interpretation of the improvement allowed by GDPC over standard DPC is as follows. At high WNR_1, the initial embedder and the helper form two fictitious users sending information to the same destination over a state-dependent Gaussian multiaccess channel (MAC), with only one of the two users knowing (noncausally) the states of the channel. In [37] and [38], it is shown that GDPC at the informed encoder is relevant, and it gives a larger rate region than standard DPC.

B. Coding Realization and Feasible Rate

We now describe a coding realization of the cooperative embedding scheme studied in this section. This scheme is based on a combination of lattices, MMSE scaling, and standard (capacity-achieving) codes.

Let $W \in \mathcal{W}$ be the message to be embedded. Consider a lattice Λ with volume $V(\Lambda)$ and normalized second moment $G(\Lambda)$ such that

$$G(\Lambda)V(\Lambda)^{\frac{2}{n}} = P_1. \tag{36}$$

We will use the scaled lattice $\Lambda_1 = \sqrt{\beta\bar{\rho}}\Lambda$ chosen here such that the modulo-Λ_1 reduction gives a quantization error signal that has power $\beta\bar{\rho}P_1$, for some $0 \le \beta \le 1$ and $0 \le \rho \le \min(1, Q/P_1)$. This lattice has a normalized second moment and volume given by

$$G(\Lambda_1) = G(\Lambda), \quad V(\Lambda_1)^{\frac{2}{n}} = \beta\bar{\rho}V(\Lambda)^{\frac{2}{n}}. \tag{37}$$

Also, proceeding similarly as in Section V, we use a cryptography key $\mathbf{K}$ which is uniformly distributed over $\mathcal{V}(\Lambda)$ and stands for a randomized codebook which is known to the initial embedder, the helper and the destination, and a mapping function $\iota(\cdot)$ which one-to-one associates messages $\{W\}$ to coset leaders $\{\mathbf{c}_w\}$. Similar to Section V, the key $\mathbf{K}$ is used for security purposes and its uniform distribution is chosen for capacity-achieving purposes.

1) Cooperative Embedding Scheme: The cooperative embedding scheme is as follows. Let $\mathbf{s}' = (1 - \sqrt{\rho P_1}/\sqrt{Q})\mathbf{s}$. The initial embedder embeds the watermark

$$\mathbf{x}_1(W; \mathbf{s}, \Lambda) = -\sqrt{\frac{\rho P_1}{Q}}\mathbf{s} + \mathbf{u}_w^{(1)} + \mathbf{u}_w^{(2)} \tag{38}$$

with

$$\mathbf{u}_w^{(1)} = \left[\mathbf{c}_w + \mathbf{k} - \alpha_1\left(1 - \sqrt{\frac{\rho P_1}{Q}}\right)\mathbf{s}\right] \bmod \Lambda_1 \tag{39}$$

for some $0 \le \alpha_1 \le 1$, and $\mathbf{u}_w^{(2)}$ is i.i.d. Gaussian with power $\bar{\beta}\bar{\rho}P_1$ and is independent of $\mathbf{u}_w^{(1)}$ and the host signal $\mathbf{s}$. The watermark $\mathbf{u}_w^{(2)}$ can be obtained by encoding the message W using any provably good (i.e., capacity-achieving) code such as the well-known turbo and low-density parity-check codes.

The helper receives the watermarked signal

$$\mathbf{y}_1 = \mathbf{x}_1 + \mathbf{s} + \mathbf{z}_1 = \mathbf{u}_w^{(1)} + \mathbf{u}_w^{(2)} + \mathbf{s}' + \mathbf{z}_1 \tag{40}$$

from which it decodes the embedded message and then it sends $\mathbf{x}_2 = \sqrt{P_2/\bar{\beta}\bar{\rho}P_1}\mathbf{u}_w^{(2)}$.

The decoding procedure at the helper is based on a modulo-reduction operation. More specifically, the helper first computes the signal[5] $\tilde{\mathbf{y}}_1 = \mathbf{y}_1 - \mathbf{u}_w^{(2)}$ and then computes the error signal $\tilde{\mathbf{y}}_1' = [\alpha_1 \tilde{\mathbf{y}}_1 - \mathbf{k}] \bmod \Lambda_1$. It decodes the embedded message as

$$\hat{W} = \underset{W \in \mathcal{W}}{\arg\min} \quad \min_{\boldsymbol{\lambda} \in \Lambda_1} \|\alpha_1 \tilde{\mathbf{y}}_1 - \mathbf{k} - \boldsymbol{\lambda} - \mathbf{c}_w\|. \tag{41}$$

The destination receives the watermarked signal

$$\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{s} + \mathbf{z} \tag{42a}$$

$$= \mathbf{u}_w^{(1)} + \left(\mathbf{u}_w^{(2)} + \mathbf{x}_2\right) + \mathbf{s}' + \mathbf{z} \tag{42b}$$

and without knowing the original host signal $\mathbf{s}$, it has to determine the embedded message. To this end, the destination first decodes the information carried by the part $(\mathbf{u}_w^{(2)} + \mathbf{x}_2)$ of the cooperative watermark, subtracts it off, and then decodes the information carried by the part $\mathbf{u}_w^{(1)}$ of the watermark. The decoding of $(\mathbf{u}_w^{(2)} + \mathbf{x}_2)$ can be performed in a standard manner, considering $\mathbf{u}_w^{(1)} + \mathbf{s}'$ as unknown noise, such as for a regular Gaussian channel. For the decoding of $\mathbf{u}_w^{(1)}$, the decoder at the destination computes the error signal $\tilde{\mathbf{y}}' = [\alpha_1 \tilde{\mathbf{y}} - \mathbf{k}] \bmod \Lambda_1$, with $\tilde{\mathbf{y}} = \mathbf{y} - (\mathbf{u}_w^{(2)} + \mathbf{x}_2) = \mathbf{u}_w^{(1)} + \mathbf{s}' + \mathbf{z}$, since the channel from $\mathbf{C}_w$ to $\tilde{\mathbf{Y}}$ is a DPC with SI $\mathbf{S}'$ noncausally known to the transmitter but not to the receiver.

2) Performance Analysis: We focus on the embedding rate allowed by this scheme. Proceeding similarly to Section V, we obtain

$$\tilde{\mathbf{y}}_1' = \left[\mathbf{c}_w + \alpha_1 \mathbf{z}_1 - (1-\alpha_1)\mathbf{u}_w^{(1)}\right] \bmod \Lambda_1 \tag{43a}$$

$$\tilde{\mathbf{y}}' = \left[\mathbf{c}_w + \alpha_1 \mathbf{z} - (1-\alpha_1)\mathbf{u}_w^{(1)}\right] \bmod \Lambda_1. \tag{43b}$$

Thus, for the decoding procedure, the helper sees an active channel noise given by the Λ_1-aliased noise $\tilde{\mathbf{Z}}_1' = [\alpha_1 \mathbf{Z}_1 - (1-\alpha_1)\mathbf{U}_w^{(1)}] \bmod \Lambda_1$. The destination sees the noise $\mathbf{Z}' = \mathbf{Z} + \mathbf{U}_w^{(1)} + \mathbf{S}'$ in decoding the cooperative part $(\mathbf{U}_w^{(2)} + \mathbf{X}_2)$ of the watermark and the Λ_1-aliased noise $\tilde{\mathbf{Z}}' = [\alpha_1 \mathbf{Z} - (1-\alpha_1)\mathbf{U}_w^{(1)}] \bmod \Lambda_1$ in decoding the part $\mathbf{U}_w^{(1)}$ of the watermark. Since the key $\mathbf{K}$ is uniformly distributed over $\mathcal{V}(\Lambda)$, its use as a dither ensures that $\mathbf{U}_w^{(1)}$ is uniformly distributed over $\mathcal{V}(\Lambda_1)$ regardless of the power Q of the host signal. Thus, the noises $\tilde{\mathbf{Z}}_1'$ and $\tilde{\mathbf{Z}}'$ are statistically independent of $\mathbf{C}_w$. Then, using the inflated lattice lemma [26], this gives rise to two MLAN channels: the channel from $\mathbf{C}_w$ to $\tilde{\mathbf{Y}}_1$ and the channel from $\mathbf{C}_w$ to $\tilde{\mathbf{Y}}$. The mutual information of these channels is maximized by a uniform input, giving

$$\begin{aligned} \frac{1}{n} I(W; \tilde{\mathbf{Y}}_1) &= \frac{1}{n}\left[h\left(\tilde{\mathbf{Y}}_1'\right) - h\left(\tilde{\mathbf{Z}}_1'\right)\right] \\ &= \frac{1}{n}\left[\log V(\Lambda_1) - h\left(\tilde{\mathbf{Z}}_1'\right)\right] \\ \frac{1}{n} I(W; \tilde{\mathbf{Y}}) &= \frac{1}{n}\left[h(\tilde{\mathbf{Y}}') - h(\tilde{\mathbf{Z}}')\right] \\ &= \frac{1}{n}\left[\log V(\Lambda_1) - h(\tilde{\mathbf{Z}}')\right]. \end{aligned} \tag{44}$$

The cooperative part $(\mathbf{U}_w^{(2)} + \mathbf{X}_2)$ of the watermark carries information at a rate given by

$$\frac{1}{n} I(\mathbf{X}_2; \mathbf{Y}) = \frac{1}{2}\log\left(1 + \frac{(\sqrt{\bar{\beta}\bar{\rho}P_1} + \sqrt{P_2})^2}{N + \beta\bar{\rho}P_1 + Q'(\rho)}\right). \tag{45}$$

The embedding rate allowed by the considered coding realization can be obtained by taking the minimum of the information that the helper can decode reliably from the signal received from the initial embedder (i.e., $(1/n)I(W; \tilde{\mathbf{Y}}_1)$) and the information that the destination can decode reliably from the signals received from the initial embedder and the helper (i.e., the sum $(1/n)[I(\mathbf{X}_2; \mathbf{Y}) + I(W; \tilde{\mathbf{Y}})]$). Substituting (37) in (44) and using (44) and (45), this gives the embedding rate

$$\begin{aligned} R_{10}(\Lambda) = \max\min\Bigg\{ &\frac{1}{2}\log\left(\frac{\beta\bar{\rho}P_1}{G(\Lambda)}\right) - \frac{1}{n}h\left(\tilde{\mathbf{Z}}_1'\right), \\ &\frac{1}{2}\log\left(1 + \frac{(\sqrt{\bar{\beta}\bar{\rho}P_1} + \sqrt{P_2})^2}{N + \beta\bar{\rho}P_1 + Q'(\rho)}\right) \\ &+ \frac{1}{2}\log\left(\frac{\beta\bar{\rho}P_1}{G(\Lambda)}\right) - \frac{1}{n}h(\tilde{\mathbf{Z}}')\Bigg\} \end{aligned} \tag{46}$$

where the maximization is over parameters $0 \le \beta \le 1$, $0 \le \rho \le \min(1, Q/P_1)$, and $0 \le \alpha_1 \le 1$ such that the terms in the minimization in (46) are nonnegative real.

Similar to the scenario in Section V, no closed-form expression can be derived for (46) and the computation of differential entropy and the maximization over all possible choices of the tuple (ρ, β, α_1) have to be performed numerically. In particular, note that one cannot derive a closed-form expression for the optimal α_1 since this parameter appears in both terms of the minimization, through the differential entropy terms $h(\tilde{\mathbf{Z}}_1')$ and $h(\tilde{\mathbf{Z}}')$.

3) Numerical Examples and Discussion: We numerically compute the rate (46) for the cubic lattice $\mathbb{Z}^n$, using Monte-Carlo integration for the computation of the differential entropy terms. The results are depicted in Fig. 8 against $\mathrm{WNR}_1 = 10\log_{10}(P_1/N_1)$ [dB] at the helper, for two examples of $\mathrm{WNR}_{\max} = 10\log_{10}((\sqrt{P_1} + \sqrt{P_2})^2/N)$ [dB] at the destination $\mathrm{WNR}_{\max} = 8$ dB and $\mathrm{WNR}_{\max} = 3$ dB. For comparison reasons, Fig. 8 also shows the theoretical embedding rate (28) obtained with GDPC, the rate (10), the embedding rate obtained in the standard case in which there is no helper (computed as in Section V), the cutset bound (14), and the lower bound obtained by treating the host signal as unknown noise. Fig. 8(a) also shows the rates given by standard DPC theoretically and using $\Lambda = \mathbb{Z}^n$ (obtained by putting $\rho = 0$ in (28) and (46), respectively).

The results shown in Fig. 8 motivate the following discussion, which we use to close this paper.

1) For this scenario also, we observe that the collaborative embedding by the helper can improve the embedding rate (w.r.t., to the standard case in which there is no helper), especially for large values of WNR_1. However, unlike the scenario in which the helper knows the host signal, here, the collaborative embedding is advantageous only if the helper receives a watermarked signal which is less noisy than the one received at the destination (the value of

[5] Note that this is possible since the helper knows $\mathbf{x}_2$, and the input $\mathbf{u}_w^{(2)}$ is proportional to $\mathbf{x}_2$.

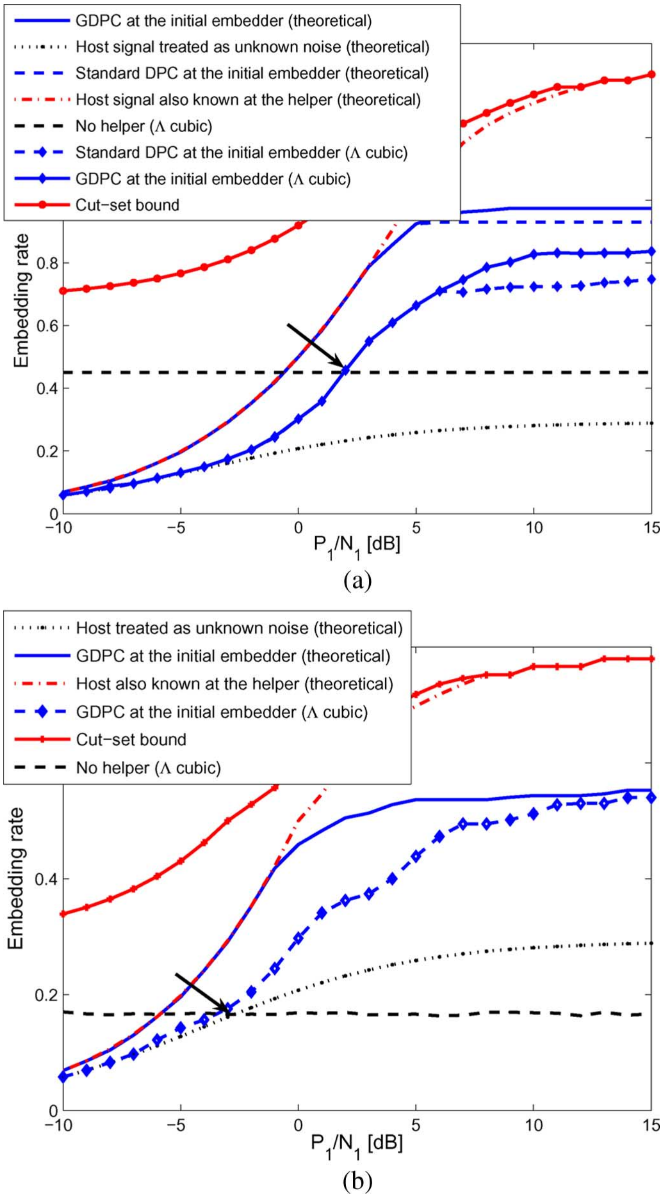


Fig. 8. Cooperative embedding rate as a function of WNR_1 at the helper, for two examples of $\mathrm{WNR}_{\max}$ at the destination (a) $\mathrm{WNR}_{\max} = 8\ \mathrm{dB}$ and (b) $\mathrm{WNR}_{\max} = 3\ \mathrm{dB}$. Illustration of the rates allowed in practice by GDPC (46) and standard DPC (given by (46) with $\rho = 0$) using a cubic lattice. Numerical values are $P_2 = P_1$, $Q = 2P_1$.

WNR_1 for which $N_1 = N$ is indicated by the arrow in Fig. 8).

2) Similar to the theoretical setup in Section VI-A, the practical implementation of GDPC with a cubic lattice also improves upon the one of standard DPC with the same lattice. Also, the improvement is especially visible at large values of WNR_1. Furthermore, with both coding schemes, larger embedding rates can be obtained by using lattices with better quantizing properties.

3) Following Costa's result, the so called "standard DPC" scheme is the one that the initial embedder would naturally apply in order to exploit the knowledge of the host signal (recall that Costa showed that an additive Gaussian interference which is known noncausally to the transmitter should not be combated by this transmitter neither totally nor partially). Thus, that partial interference cancellation (through GDPC) turns out to be relevant in our setup means that if embedding is to be performed by different embedders jointly, then a certain degree of coordination between these embedders is needed since the appropriate embedding strategy of one embedder may depend on those of the other embedders and also on whether they know the host signal.

VII. Conclusion

In this paper, we investigate the problem of embedding information into digital media content transmitted over wireless networks, cooperatively by two users or partners—an initial embedder and an assisting embedder or helper. The helper reinforces the watermark embedded by the initial embedder so that it can survive subsequent channel degradations on the network. The considered setup has application in reinforcing digital watermarks transmitted over wireless multimedia networks.

We analyze the cooperative embedding for the two conceptually different scenarios in which the helper does or does not know the original host signal. For each of the two scenarios, we derive lower and upper bounds on the cooperative embedding rate. For the scenario in which the helper knows the original host signal, the lower bound is obtained with a coding scheme in which the initial embedder and the helper employ standard DPC schemes. For the scenario in which the helper does not know the original host signal, the lower bound is obtained with a coding scheme in which the initial embedder employs a generalized DPC (GDPC) scheme. GDPC consists of a combination of partial-state cancellation and standard DPC. Furthermore, we also design implementable coding schemes for the considered two scenarios and determine the embedding rates allowed by these schemes.

Appendix

A. *Proof of Lemma 2*

We describe a random coding scheme that we use to prove the achievability of rate (29). This coding scheme is based on a combination of regular-encoding sliding-window decoding [39] and Gel'fand–Pinsker binning [19].

Codebook Generation: Fix a measure $P_{S,U_1,X_1,X_2,Y_1,Y}$ satisfying (30). Fix $\epsilon > 0$ and denote

$$J = 2^{n[I(U_1;S|X_2)+2\epsilon]} \tag{A1a}$$

$$M = 2^{n[R-4\epsilon]}. \tag{A1b}$$

We generate two statistically independent codebooks (codebooks 1 and 2) by following the steps that are outlined below twice. These codebooks will be used for blocks with odd and even indices, respectively (see the Encoding step).

1) We generate M i.i.d. codewords $\{\mathbf{x}_2(w')\}$ indexed by $w' = 1, 2, \ldots, M$, each of length n and with i.i.d. components drawn according to P_{X_2}.

2) For each codeword $\mathbf{x}_2(w')$, we generate a collection of $M \times J$ i.i.d. auxiliary codewords $\{\mathbf{u}_1(w', w, j)\}$ indexed by $w = 1, 2, \ldots, M$, $j = 1, \cdots, J$. These codewords

have length n each and their components are drawn i.i.d. according to $P_{U_1|X_2}$.

To embed a message W by the initial embedder, this message is divided into B blocks $w_1, w_2, \cdots, w_B$ of nR bits each; and embedding is performed over $B+1$ blocks by selecting the appropriate codewords for each block, as will be specified. For fixed n, the rate $R(B/(B+1))$ approaches R as $B \longrightarrow +\infty$.

Encoding: We encode messages using codebooks 1 and 2, respectively, for blocks with odd and even indices. Using independent codebooks for blocks with odd and even indices makes the error events correspond to these blocks independent and, hence, the corresponding probabilities are easier to evaluate.

Continuing with the strategy, assume that at the beginning of block i, w_i is the new message to be embedded by the initial embedder and w_{i-1} is the message embedded in the previous block $i-1$. Also, assume that at the beginning of block i, the helper has decoded w_{i-1} correctly. Then, the helper sends $\mathbf{x}_2(w_{i-1})$. In order to embed w_i, the initial embedder searches for the smallest $j \in \{1, \cdots, J\}$ such that $\mathbf{u}_1(w_{i-1}, w_i, j)$ is jointly typical with $\mathbf{s}(i)$ given $\mathbf{x}_2(w_{i-1})$, where $\mathbf{s}(i)$ denotes the state in block i. Denote this j by $j^\star = j(\mathbf{s}(i), w_{i-1})$. If such $j^\star$ is not found, or if the observed state is not typical, an error is declared and $j^\star$ is set to J. Then, the initial embedder embeds a vector $\mathbf{x}_1$ which is drawn i.i.d. conditionally given $(\mathbf{u}_1(w_{i-1}, w_i, j^\star), \mathbf{x}_2(w_{i-1}), \mathbf{s}(i))$ [using the appropriate marginal induced by the distribution (30)].

Decoding: Decoding is based on a combination of joint typicality and sliding window. The decoding procedures at the end of block i are as follows.

1) The helper, having known w_{i-1}, declares that $\hat{w}_i$ is embedded if there is a unique $\hat{w}_i$ such that $\mathbf{u}_1(w_{i-1}, \hat{w}_i, j)$ is jointly typical with $\mathbf{y}_1(i)$ given $\mathbf{x}_2(w_{i-1})$. Random coding arguments guarantee that the decoding error in this step is small for sufficiently large n if
$$R < I(U_1; Y_1|X_2) - I(U_1; S|X_2). \qquad \text{(A2)}$$

2) The destination knows w_{i-2} and decodes w_{i-1} based on the information received in block $i-1$ and block i. It declares that the message $\hat{w}_{i-1}$ is embedded if there is a unique $\hat{w}_{i-1}$ such that $\mathbf{x}_2(\hat{w}_{i-1})$ is jointly typical with $\mathbf{y}(i)$ and $\mathbf{u}_1(w_{i-2}, \hat{w}_{i-1}, j)$ is jointly typical with $\mathbf{y}(i-1)$ given $\mathbf{x}_2(w_{i-2})$. Random coding arguments guarantee that the decoding error in this step is small for sufficiently large n if
$$R < I(U_1, X_2; Y) - I(U_1; S|X_2). \qquad \text{(A3)}$$

Combining (A2) and (A3), we obtain (29), and this completes the proof of Lemma 2.

References

[1] B. Chen, S. C. Draper, and G. Wornell, "Information embedding and related problems: Recent results and applications," presented at the Allerton Conf. Comm., Control, Computing, Monticello, IL, Oct. 2001.

[2] B. Chen and G. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[3] A. Zaidi, P. Piantanida, and P. Duhamel, "Broadcast- and MAC-aware coding strategies for multiple user information embedding," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2974–2992, Jun. 2007.

[4] S. Kotagiri and J. N. Laneman, "Variations on information embedding in multiple access and broacast channels," *IEEE Trans. Inf. Theory*, 2008, submitted for publication.

[5] A. Khisti, U. Erez, A. Lapidoth, and G. Wornell, "Carbon copying onto dirty paper," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1814–1827, May 2007.

[6] I. Cox, M. Miller, and A. McKellips, "Watermarking as communication with side information," in *Proc. Int. Conf. Multimedia Computing Systems*, Jul. 1999, pp. 1127–1141.

[7] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 563–593, Mar. 2003.

[8] F. Pérez-González, C. Mosquera, M. Barni, and A. Abrardo, "Rational dither modulation: A high-rate data-hiding method invariant to gain attacks," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pt. 2, pp. 3960–3975, Oct. 2005.

[9] M. Barni and F. Bartolini, "Data hiding for fighting piracy," *IEEE Signal Process. Mag.*, vol. 21, no. 2, pp. 28–39, Mar. 2004.

[10] H. S. Malvar and D. A. F. Florêncio, "Improved spread spectrum: A new modulation technique for robust watermarking," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 898–905, Apr. 2003.

[11] R. Puri, K. Ramchandran, and S. S. Pradhan, "On seamless digital upgrade of analog transmission systems using coding with side information," presented at the Allerton Conf. Communication, Control, Computing, Urbana-Champaign, IL, Oct. 2002.

[12] M. Löytynoja, A. Keskinarkaus, N. Cvejic, and T. Seppänen, "Watermark-enabled value added services to broadcast audio," in *Proc. IEEE Int. Conf. Digital Ecosystem Technologies*, Feb. 2008, pp. 388–396.

[13] S. Voloshynovskiy, F. Deguillaume, O. Koval, and T. Pun, "Information-theoretic data-hiding for public network security, services control and secure communications," in *Proc. IEEE Int. Conf. Telecomm. Modern Satellite, Cable and Broadcasting Service*, Oct. 2003, vol. 1, pp. 3–17.

[14] A. Benjelloun-Touimi, J.-B. Fischer, C. Fontaine, C. Giraud, and M. Milhau, "Enhanced security architecture for music distribution on mobile," in *Proc. Eur. Symp. Research in Computer Security*, Sep. 2006, pp. 97–109.

[15] Y. Zhang, W. Lee, and Y. Huang, "Intrusion detection techniques for mobile wireless networks," *J. Wireless Netw.*, vol. 9, pp. 545–556, Nov. 2003.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[17] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[18] J. H. Conway and N. J. A. Sloane, *Sphere Packing, Lattices and Groups*, 3rd ed. New York: Wiley, 1988.

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

[20] J. J. Eggers, R. Bäuml, R. Tzschoppe, and B. Girod, "Scalar costa scheme for information embedding," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 1003–1019, Apr. 2003.

[21] R. F. H. Fischer, R. Tzschoppe, and R. Bäeuml, "Lattice costa schemes using subspace projection for digital watermarking," in *Proc. ITG Conf. Source Channel Coding*, Erlangen, Germany, Jan. 2004, pp. 127–134.

[22] P. Moulin and R. Koetter, "Data-hiding codes," *Proc. IEEE*, vol. 93, no. 12, pp. 2083–2126, Dec. 2005.

[23] A. Zaidi and P. Duhamel, "Modulo lattice additive noise channel for QIM watermarking," in *Proc Int. Conf. Image Processing*, Genova, Italy, Sep. 2005, pp. 993–996.

[24] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.

[25] L. Pérez-Freire, P. Comesaña, J.-R. Troncoso-Pastoriza, and F. Pérez-González, "Watermarking security: A survey," *Trans. Data Hiding and Multimedia Security I*, vol. 21, pp. 41–72, Oct. 2006.

[26] U. Erez, S. Shamai (Shitz), and R. Zamir, "Capacity and lattice strategies for cancelling known interference," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.

[27] R. F. H. Fischer, "The modulo-lattice channel: The key feature in precoding schemes," *Int. J. Electron. Commun.*, pp. 244–253, Oct. 2005.

[28] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *IEEE Electon. Lett.*, vol. 7, no. 5, pp. 138–139, Mar. 1971.

[29] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. COM-20, no. 4, pp. 774–780, Aug. 1972.

[30] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[31] A. Zaidi, L. Vandendorpe, and P. Duhamel, "Lower bounds on the capacity regions of the multi-relay channel and the multi-relay broadcast channel with non-causal side-information," in *Proc. IEEE Int. Conf. Communications*, Glasgow, U.K., Jun. 2007, pp. 6005–6011.

[32] A. Zaidi and L. Vandendorpe, "Rate regions for the partially-cooperative relay-broadcast channel with non-causal side information," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1246–1250.

[33] Y.-H. Kim, A. Sutivong, and S. Sigurjonsson, "Multiple user writing on dirty paper," in *Proc. IEEE Int. Symp. Info. Theory*, Chicago, IL, June 2004, p. 534.

[34] G. D. Forney, M. D. Trott, and S. Y. Chung, "Sphere-bound-achieving cosets codes and multilevel coset codes," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 820–850, May 2000.

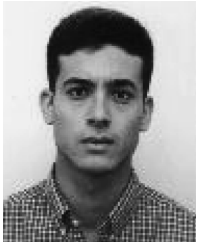
[35] G. D. Forney and L. F. Wei, "Multidimensional constellations—Part I: Introductions figures of merit, and generalized cross constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 877–892, Aug. 1989.

[36] J. G. D. Forney, "Multidimensional constellations—Part II: Voronoi constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 941–958, Aug. 1989.

[37] S. Kotagiri and J. Laneman, "Multiaccess channels with state information known to one encoders: A case of degraded message sets," presented at the IEEE Int. Symp. Information Theory, Nice, France, 2007.

[38] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdù, "Cooperative multiple access encoding with states available at one transmitter," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4448–4469, Oct. 2008.

[39] A. B. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 841–850, Nov. 1982.



Abdellatif Zaidi was born in Tunisia in 1978. He received the Eng. degree in electrical engineering from École Nationale Supérieure de Techniques Avancés, ENSTA ParisTech, Paris, France, in 2002 and the M.Sc. and Ph.D. degrees from École Nationale Supérieure des Télécommunications, TELECOM ParisTech, Paris, in 2002 and 2005, respectively.

From 2002 to 2005, he was with the Communications and Electronics Department at TELECOM ParisTech, and the Signals and Systems Lab., CNRS/Supélec, Supélec. Currently, he is a Research Assistant with École Polytechnique de Louvain, Université Catholique de Louvain, Louvain-la-Neuve, Belgium. He was Research Visitor at the University of Notre Dame, Notre Dame, IN, in 2007 and 2008. His research interests cover a broad range of topics from signal processing for communication and multiuser information theory. Of particular interest are the problems of coding for side-informed channels, secure communication, coding and interference mitigation in multiuser channels, and relaying problems and cooperative communication with applications to sensor networking and ad-hoc wireless networks.

Luc Vandendorpe (M'93–SM'99–F'06) was born in Mouscron, Belgium, in 1962. He received the electrical engineering degree (Hons.) and the Ph.D. degree from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium in 1985 and 1991, respectively.

Since 1985, he has been with the Communications and Remote Sensing Laboratory of UCL where he first worked in the field of bit-rate reduction techniques for video coding. In 1992, he was a Visiting Scientist and Research Fellow at the Telecommunications and Traffic Control Systems Group of the Delft Technical University, Delft, The Netherlands, where he worked on spread-spectrum techniques for personal communications systems. From 1992 to 1997, he was a Senior Research Associate of the Belgian National Science Foundation at UCL, and Invited Assistant Professor. Currently, he is Professor and Head of the Electrical Engineering Department of UCL. He is mainly interested in digital communication systems: equalization, joint detection/synchronization for code-division multiple access (CDMA), orthogonal frequency-division multiple access (OFDM) (multicarrier), multiinput multioutput (MIMO), distributed MIMO, and turbo-based communications systems (UMTS, xDSL, WIMAX) as well as localization. He was an editor of the IEEE TRANSACTIONS ON COMMUNICATIONS FOR SYNCHRONIZATION AND EQUALIZATION between 2000 and 2002, associate editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS between 2003 and 2005, and associate editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2005 to 2007. He is Editor-in-Chief for the *Eurasip Journal on Wireless Communications and Networks*.

Dr. Vandendorpe was Co-Recipient of the Biennal Alcatel-Bell Award from the Belgian National Science Foundation (NSF) for a contribution in the field of image coding in 1990. In 2000, he was Co-Recipient (with J. Louveaux and F. Deryck) of the Biennal Siemens Award from the Belgian NSF for a contribution about filter-bank-based multicarrier transmission. In 2004, he was Co-Winner (with J. Czyz) of the Face Authentication Competition, FAC 2004. He is or has been a TPC member of IEEE VTC since 1999, IEEE Globecom 2003 Communications Theory Symposium, the 2003 Turbo Symposium, IEEE VTC Fall 2003, IEEE SPAWC 2005, IEEE SPAWC 2006, IEEE SAM 2008, Eusipco 2008, ISSSTA 2008, PIMRC 2008, WPMC 2008, and TURBO 2008. He was Co-Technical Chair (with P. Duhamel) for IEEE ICASSP 2006. He was Chair of the IEEE Benelux joint chapter on Communications and Vehicular Technology between 1999 and 2003. He was an elected member of the Signal Processing for Communications committee between 2000 and 2005. Currently, he is an elected member of the Sensor Array and Multichannel Signal Processing committee of the Signal Processing Society.