Hindawi Publishing Corporation
EURASIP Journal on Information Security
Volume 2009, Article ID 914937, 14 pages
doi:10.1155/2009/914937

Research Article

High-Rate Data-Hiding Robust to Linear Filtering for Colored Hosts

Michele Scagliola,[1] Fernando Pérez-González,[2] and Pietro Guccione[1]

[1] *Dipartimento di Elettrotecnica ed Elettronica, Politecnico di Bari, Via E. Orabona 4, 70125 Bari, Italy*
[2] *Signal Theory and Communications Department, University of Vigo, 36200 Vigo, Spain*

Correspondence should be addressed to Michele Scagliola, m.scagliola@poliba.it

Received 30 April 2009; Revised 24 July 2009; Accepted 24 September 2009

Recommended by Alessandro Piva

The discrete Fourier transform-rational dither modulation (DFT-RDM) has been proposed as a way to provide robustness to linear-time-invariant (LTI) filtering for quantization-based watermarking systems. This scheme has been proven to provide high rates for white Gaussian hosts but those rates considerably decrease for nonwhite hosts. In this paper the theoretical analysis of DFT-RDM is generalized to colored Gaussian hosts supplied with an explanation of the performance degradation with respect to white Gaussian hosts. Moreover the characterization of the watermark-to-noise ratio in the frequency domain is shown as an useful tool to give a simple and intuitive measure of performance. Afterwards an extension of DFT-RDM is proposed to improve its performance for colored hosts without assuming any additional knowledge on the attack filter. Our analysis is validated by experiments and the results of several simulations for different attack filters confirm the performance improvement afforded by the whitening operation for both Gaussian colored hosts and audio tracks.

1. Introduction

Quantization index modulation (QIM) [1] is a wide class of watermarking methods which are proven to yield optimum performance in additive white Gaussian channels without downgrading the host signal fidelity. The main drawback of quantization-based schemes is their sensitivity to valumetric distortions; these attacks vary the amplitude of the watermarked signal so that, even if they do not usually reduce the perceived quality of the media, the produced mismatch between encoder and decoder lattice volumes severely increases the bit-error rate (BER). Consequently, a great effort has been spent by researchers in developing quantization-based methods robust to valumetric distortions and the problem can be considered somewhat solved by different approaches, that is, [2–4].

Linear-time-invariant (LTI) filtering attack is in some sense related to valumetric distortions; in spite of the simplicity and wide use of filtering in signal processing, literature about this attack for quantization-based schemes is scarce. This is even more dramatic since basic quantization-based schemes are not able to cope with filtering attacks; in fact it has been proven that by cutting away with a lowpass filter only one percent of the signal spectrum, the resulting BER for binary time-domain Dither Modulation (DM) is already 0.5 [5].

Apart from the work done by Wang et al. [6], where the decoder is assumed to have some information about the attack filter and the maximum-likelihood criterion is used to estimate the frequency gain, the LTI filtering attack has been addressed only in [5]. In that work, it is proposed an extension of the rational dither modulation (RDM) scheme [4], which is robust to LTI filtering without assuming any prior knowledge about the attack filter. The main idea relies on the amplitude scaling invariance of RDM and on the convolution theorem [7], so that an RDM-like channel is constructed on a subset of the frequency channels in the discrete Fourier transform (DFT) domain. Analytical and experimental results in [5] demonstrate that a high-rate can be reached for white Gaussian hosts, but experiments carried out with audio signals have shown a severe loss of performance for nonstationary, non-Gaussian, and colored

hosts. On the other hand, the analysis developed in [5] is focused uniquely on white Gaussian hosts, so that it cannot be straightforwardly used to justify the experimental results obtained for nonwhite hosts.

In this paper the behavior of DFT-RDM for Gaussian colored hosts is investigated. By modeling the colored host with an autoregressive (AR) [7] random process, the analysis of DFT-RDM is generalized, providing an explanation for the loss of performance with respect to white Gaussian hosts. This is essentially due to the combination of two facts: (1) the power of an RDM watermark signal is proportional to the host signal power, and (2) the influence of the nonflat power spectral density (psd) of the host on the self-noise that in turn is due to a block-DFT operation. Moreover, we introduce the per-channel watermark-to-noise ratio (WNR) as a simple measure to evaluate the reliability of each RDM-like channel.

We also propose an extension of DFT-RDM that improves performance in the case of colored hosts under the hypotheses of a blind watermarking scheme and total ignorance about the attack filter both at the embedder and the decoder. In such case, low error probabilities are obtained by performing DFT-RDM embedding and decoding after a whitening operation, without any penalty in terms of embedding distortion and payload.

The paper is organized as follows. In Section 2 some notations are introduced while DFT-RDM is revised in Section 3. The behavior of DFT-RDM with a Gaussian colored host is analyzed in Section 4 and in Section 5 the proposed extension of DFT-RDM is presented. Numerical simulations that validate the developed analysis and show the performance of the proposed approach are given in Section 6; finally in Section 7 some conclusions are drawn.

2. Notation

We assume 1D real-valued hosts arranged in vectors, which are denoted by boldface letters, so that $\mathbf{x}$ is a vector and x_l is its lth element. As customary in data-hiding applications, if the vector $\mathbf{x}$ is the host signal, after the watermark embedding the watermarked signal is denoted by $\mathbf{y}$ and the watermark signal is by definition $\mathbf{w} \triangleq \mathbf{y} - \mathbf{x}$. The vector $\mathbf{z}$ denotes the samples received by the decoder at the channel output.

Uppercase letters will be used for random variables, that is, X_l is a random variable modeling the lth sample of the host signal, and $\{X_l\}$ is the random process related to the whole sequence $\{x_l\}$. Finally, to denote a variable in the DFT domain, the tilde will be used, so that the random variable $\widetilde{X}_{m,k}$ is the kth coefficient of the DFT computed on the mth block of the host signal. Similarly, if h_l is the impulse response of a real-valued LTI filter, $H(e^{j\omega})$ denotes its Fourier transform so that we have $\widetilde{h}_k \triangleq H(e^{j2\pi k/N})$.

Finally, for zero-mean hosts we define the document-to-watermark ratio (DWR) as the ratio between the host signal variance σ_x^2 and the embedding distortion D_w, which is the average power of the watermark signal, as customary.

3. Review of DFT-RDM

The discrete Fourier transform-rational dither modulation (DFT-RDM) method has been proposed in [5] to counteract linear-time-invariant (LTI) filtering. This scheme is based on RDM [4], which is a high-rate quantization-based data-hiding method invariant to amplitude scaling, and on the convolution theorem [7], which allows to represent the filter output as a multiplication in the Fourier domain of the input signal and the filter response.

In a real application DFT-RDM uses the discrete Fourier transform in a block-by-block basis instead of the full-sequence Fourier transform [5], which would be impractical due to its computational complexity and the memory required by RDM. In the adopted framework the exact multiplication in the DFT domain would only be achieved with a circular convolution, whereas the filtered signal is obtained through an ordinary convolution. As a consequence, the effect of filtering on each DFT channel cannot be modeled by a pure scaling, but a host-dependent error has to be considered too.

Assuming nonoverlapping DFT blocks of length N, let $\mathbf{x}_m$ be the mth block of the host signal and $\widetilde{x}_{m,k}$ the kth coefficient of the DFT of such block:

$$\widetilde{x}_{m,k} = \sum_{l=0}^{N-1} x_{m,l} \exp\left(-j\frac{2\pi k}{N}l\right). \tag{1}$$

The information bits are embedded into the absolute value of the DFT coefficients, taking care in preserving the symmetry of the DFT for real signals. Essentially, on each of the first $N/2+1$ discrete frequencies an RDM-like channel is constructed so that the absolute value of the watermarked signal is

$$|\widetilde{y}_{m,k}| = g(\widetilde{\mathbf{y}}_{m-1,k})\, Q_{b_{m,k}}\left(\frac{|\widetilde{x}_{m,k}|}{g(\widetilde{\mathbf{y}}_{m-1,k})}\right), \tag{2}$$

where $\widetilde{\mathbf{y}}_{m-1,k} \triangleq (\widetilde{y}_{m-1,k}, \widetilde{y}_{m-2,k}, \ldots, \widetilde{y}_{m-L,k})^T$ and $0 \le k \le N/2$. The phase of $\widetilde{y}_{m,k}$ is set equal to the phase of $\widetilde{x}_{m,k}$ so that the embedding distortion is minimized; in order to preserve symmetry, the remaining DFT coefficients are updated according to the rule $\widetilde{y}_{m,k} = \widetilde{y}^*_{m,N-k}$ for $N/2 + 1 < k < N$, where the superscript $*$ denotes the complex conjugate. The watermarked signal is then mapped back into the original domain through a nonoverlapping block-by-block inverse DFT of the marked coefficients:

$$y_{m,l} = \frac{1}{N}\sum_{k=0}^{N-1} \widetilde{y}_{m,k} \exp\left(j\frac{2\pi l}{N}k\right). \tag{3}$$

Due to the orthogonality of the DFT, the DWR in the DFT domain is identical to that in the time domain. Hence DFT-RDM inherits from the standard RDM the relations between quantization step-size, power of the watermark signal and DWR. It is worth noting that all the RDM-like channels use the same quantization step-size, which is computed from the knowledge of the target overall DWR.

At the decoder, with $\mathbf{z}_m$ denoting the mth block of the received signal, the relative DFT coefficients are computed

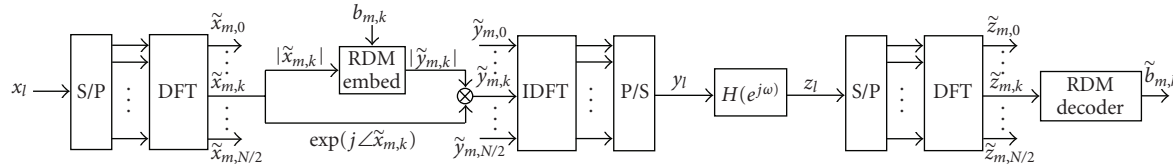

Figure 1: Block scheme of the whole embedding/decoding chain for DFT-RDM.

and the standard RDM decoder is then applied to estimate the embedded information bits. Assuming $z_l = y_l * h_l$, under the hypothesis of N sufficiently large to approximate an ordinary convolution, we have $\widetilde{z}_{m,k} \approx \widetilde{h}_k \widetilde{y}_{m,k}$, from which the RDM decoder is able to recover the correct information bits. The whole embedding/decoding block scheme is shown in Figure 1.

Due to the effects of the circular convolution, the random variable representing the kth received DFT coefficient can be written as $\widetilde{Z}_{m,k} = \widetilde{h}_k(\widetilde{X}_{m,k} + \widetilde{W}_{m,k} + \widetilde{N}_{m,k})$, where $\widetilde{N}_{m,k}$ models the deviation from a pure multiplication (which would correspond to full-length DFTs) and so it will be referred to as *per-channel multiplication error*. Under the hypothesis of large DWR and using the filter-bank interpretation of the DFT [7], this term can be expressed as

$$\widetilde{N}_{m,k} \approx X_l * f_{l,k}\,|_{l=mN+N-1}, \tag{4}$$

where $f_{l,k}$ is given by

$$f_{l,k} \triangleq \left(\frac{h_l}{\widetilde{h}_k} - \delta_l\right) * \phi^*_{l,k}, \quad k = 0, 1, \ldots, N-1, \tag{5}$$

and, by definition, $\phi_{l,k} \triangleq v_l \exp(-j2\pi lk/N)$ for $l, k = 0, \ldots, N-1$ and is zero otherwise, with δ_l denoting the Kronecker's delta. Here $\phi_{l,k}$ represents the impulse response of the kth DFT basis function multiplied by a window $\mathbf{v} = (v_0, v_1, \ldots, v_{N-1})^T$ whose purpose will be made clear shortly. Hence, from (4) and (5) it can be seen that the per-channel multiplication error is strictly dependent on both the filter coefficients h_l and the host signal. Let $\{X_l\}$ be a zero-mean white process with variance σ_x^2, then the process $\{\widetilde{X}_l * f_{l,k}\}$ can be assumed stationary as discussed in [5], and so $\widetilde{N}_{m,k}$ will approximately have zero mean and variance:

$$\sigma^2_{\widetilde{N}}(k) = \frac{\sigma_x^2}{2\pi}\int_{-\pi}^{\pi} \left|\Phi(e^{j\omega})\right|^2 \left|1 - \frac{H\left(e^{j(\omega+2\pi k/N)}\right)}{H\left(e^{j2\pi k/N}\right)}\right|^2 d\omega, \tag{6}$$

where $\Phi(e^{j\omega})$ is the Fourier transform of the window $\mathbf{v}$.

To reduce the error probability, in [5] two improvements have been proposed: windowing and spreading. The former entails multiplying the block $\mathbf{x}_m$ by a properly designed window $\mathbf{v}$ before computing the DFT coefficients at the price of an increased peak-to-average distortion. The latter amounts to adding M length-N blocks $\mathbf{x}_m$ and then applying the DFT-RDM embedding on N samples. By spreading, the robustness against filtering is increased while the payload is reduced by a factor of M.

Full details on DFT-RDM and its performance can be found in [5], where guidelines are provided for the case of white Gaussian hosts to assist the designer in the parameter selection that leads to acceptable BER values. Unfortunately, the results of some experiments with audio signals (which are nonstationary, non-Gaussian and colored hosts) reported in [5] show a considerable increase of the BER with respect to white Gaussian hosts using the same system parameters.

4. Performance Analysis for Colored Gaussian Hosts

In this section the analysis of DFT-RDM is extended to colored hosts using a frequency-domain approach and introducing some new tools. As shown in Section 6, and similarly to the experimental results for audio signals reported in [5], if a watermark is embedded in a colored host using DFT-RDM and then filtered with a conventional audio equalizer, the measured BER is noticeably greater than the BER for a white host using the same system parameters. The rationale for this behavior can be found in the inner working of DFT-RDM, which is essentially an RDM-like scheme for every DFT channel, and in the influence of a nonflat psd on the per-channel multiplication error. In [5] this error was characterized in the time domain; in contrast, we pursue here a frequency-domain approach, which is needed to separate each RDM-like channel and will lead to a somewhat simpler expression. However, the main novelty of our analysis lies in the usage of the per-channel watermark-to-noise ratio (WNR), which is a very convenient and intuitive measure that is directly related to the BER.

To better understand the behavior of DFT-RDM for audio signals, we have focused on colored Gaussian hosts modeled by an Autoregressive (AR) random process [7]. Hence, given a zero-mean white Gaussian host $\mathbf{x_0}$ with psd $\sigma^2_{x_0}$, the colored host $\mathbf{x}$ can be regarded to as the output of an all-pole filter $H_{\mathrm{AR}}(z) = 1/A(z)$ excited by $\mathbf{x_0}$. The host power spectral density can then be written as

$$S_x\left(e^{j\omega}\right) = \frac{\sigma^2_{x_0}}{\left|A(e^{j\omega})\right|^2}. \tag{7}$$

The idea is to work with a colored host whose psd resembles that of a generic audio signal, which typically has most of its power concentrated at lower frequencies. Hereinafter for colored hosts we will assume an AR signal which models the spectral contents of this generic audio signal.

We are interested in evaluating the performance (as measured by the BER) on each DFT channel; to this end, we

will rely on the watermark-to-noise ratio (WNR). It is very important to remark that while the WNR is usually defined as the ratio between the powers of the watermark signal and the attack noise, since in our framework the only impairment is the filtering, we will define the per-channel WNR as the ratio between the power of the watermark signal and that of the multiplication-error for each frequency channel:

$$\mathrm{WNR}(k) \triangleq \frac{E\left\{\left|\widetilde{Y}_{m,k} - \widetilde{X}_{m,k}\right|^2\right\}}{\sigma^2_{\widetilde{N}}(k)}, \tag{8}$$

where $E\{\cdot\}$ denotes the statistical expectation.

As a first step towards obtaining the per-channel WNR, the per-channel host power in the DFT domain has to be derived. To this aim, the filter-bank interpretation of the DFT [7] can be adopted, according to which it is possible to get

$$\widetilde{X}_{m,k} = X_l * \phi^*_{l,k}\Big|_{l=mN+N-1}. \tag{9}$$

The variance of the zero-mean process $\widetilde{X}_{m,k}$ is given by $\sigma^2_{\widetilde{X}}(k) = E\{|X_l * \phi^*_{l,k}|^2\}$ and can be computed by applying Parseval's relation, so that we have

$$\sigma^2_{\widetilde{X}}(k) = \frac{\sigma^2_{x_0}}{2\pi}\int_{-\pi}^{\pi}\left|\Phi(e^{j\omega})\right|^2\left|\frac{1}{A\left(e^{j(\omega+2\pi k/N)}\right)}\right|^2 d\omega. \tag{10}$$

According to the corresponding relation in [4] and for $p = 2$ in the g function, after the RDM embedding, the per-channel watermark signal power is

$$\sigma^2_{\widetilde{W}}(k) = \frac{\Delta^2}{3}\sigma^2_{\widetilde{X}}(k), \tag{11}$$

where the quantization step-size Δ is set to have a watermarked signal with the desired DWR. Since the per-channel watermark signal power is proportional to the per-channel host power because of the properties of RDM, a larger watermark signal originates from those host DFT channels having stronger spectral contents. Hence, in the lower frequencies of an audio-like colored host, the per-channel watermark signal will be much larger than the corresponding to higher-frequencies. This shaping of the per-channel watermark power alters the behavior of DFT-RDM with respect to that of a white Gaussian host, where the per-channel watermark power is uniform, as analyzed in [5].

On the other hand, the spectral shaping of the host influences also the per-channel multiplication error, which for high DWRs can be approximated by $\widetilde{N}_{m,k} = (X_l + W_l) * f_{l,k}|_{l=mN+N-1} \approx X_l * f_{l,k}|_{l=mN+N-1}$, as it has been explained in Section 3.

Recalling (6) and assuming reasonably the stationarity of $\widetilde{N}_{m,k}$, its variance can be written as

$$\begin{aligned}\sigma^2_{\widetilde{N}}(k) = {} & \frac{\sigma^2_{x_0}}{2\pi}\int_{-\pi}^{\pi}\left|\Phi(e^{j\omega})\right|^2\left|\frac{1}{A\left(e^{j(\omega+2\pi k/N)}\right)}\right|^2 \\ & \times\left|1 - \frac{H\left(e^{j(\omega+2\pi k/N)}\right)}{H\left(e^{j2\pi k/N}\right)}\right|^2 d\omega.\end{aligned} \tag{12}$$

The watermark-to-noise ratio can be useful to infer whether the RDM channel is able to correctly convey the information bits, because the probability of the error approaches 0.5 when the power of the watermark signal is approximately equal to that of the additive noise. Thus, the per-channel WNR is computed as the ratio between (11) and (12):

$$\mathrm{WNR}(k) = \frac{(\Delta^2/3)\int_{-\pi}^{\pi}\left|\Phi(e^{j\omega})\right|^2\left|1/A\left(e^{j(\omega+2\pi k/N)}\right)\right|^2 d\omega}{\int_{-\pi}^{\pi}\left|\Phi(e^{j\omega})\right|^2\left|1/A\left(e^{j(\omega+2\pi k/N)}\right)\right|^2\left|1 - H\left(e^{j(\omega+2\pi k/N)}\right)/H\left(e^{j2\pi k/N}\right)\right|^2 d\omega}. \tag{13}$$

To easily understand the influence of the spectral shaping of the host on the WNR, it is useful to approximate the per-channel host power as

$$\sigma^2_{\widetilde{X}}(k) \approx N\frac{\sigma^2_{x_0}}{\left|A(e^{j2\pi k/N})\right|^2}. \tag{14}$$

By this approximation, which is valid only in the case of a rectangular window, the effects of computing the DFT on finite-length blocks are neglected. Consequently, the WNR can be approximated as follows:

$$\mathrm{WNR}(k) \approx \frac{N\Delta^2/3}{\int_{-\pi}^{\pi}\left|\Phi(e^{j\omega})\right|^2\left|A\left(e^{j2\pi k/N}\right)/A\left(e^{j(\omega+2\pi k/N)}\right)\right|^2\left|1 - H\left(e^{j(\omega+2\pi k/N)}\right)/H\left(e^{j2\pi k/N}\right)\right|^2 d\omega}. \tag{15}$$

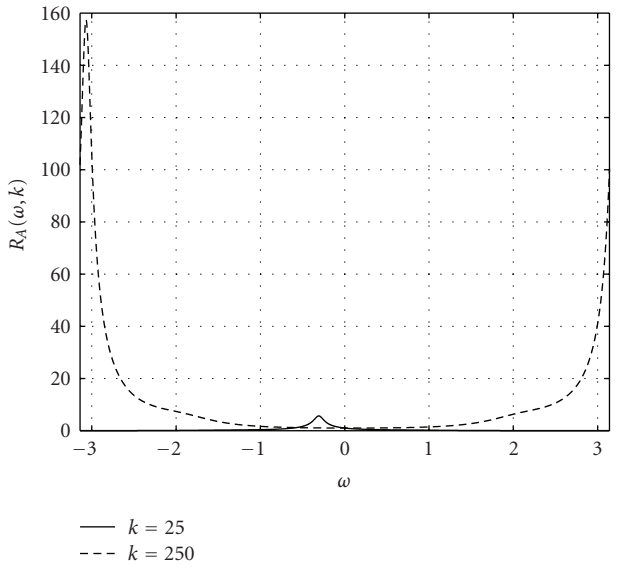

FIGURE 2: $R_A(\omega, k)$ versus discrete frequency for $k = 25$ and $k = 250$ ($N = 512$).

If the host signal is white, then the ratio $R_A(\omega, k) \triangleq |A(e^{j2\pi k/N})/A(e^{j(\omega+2\pi k/N)})|$ is equal to 1 for every k and consequently WNR(k) depends only on the attack filter; if the host is colored, this ratio is a function which has great variations for the different channels k thus affecting heavily WNR(k). As shown in Figure 2, because of the high-pass behavior of $A(z)$, for k corresponding to the high-frequency channels, the function $R_A(\omega, k)$ takes values much larger than those corresponding to low frequencies. Therefore, the spectral shaping of the host yields less robustness in high-frequency channels compared to low-frequency channels; however, strictly speaking, the per-channel WNR also depends on the attack filter, as is evident from (15).

In [5] the per-channel bit-error probability has been derived analytically relying on the results in [4], where the bit-error probability of an RDM channel is derived for i.i.d. host samples and additive noise independent of the host signal. If $P_{e,\mathrm{RDM}}(L, s)$ denotes the bit-error probability of classical RDM, with s the effective signal-to-noise ratio, the bit-error probability of the kth channel of DFT-RDM is

$$P_e(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{e,\mathrm{RDM}}\left(L, \frac{\Delta\ \sigma_{\widetilde{X}}(k)}{\sigma_{\widetilde{N}}(k,\theta)}\right)d\theta, \tag{16}$$

where $\sigma_{\widetilde{N}}(k,\theta)$ is the magnitude of the per-channel multiplication error projected onto $e^{j\theta}$; see [5].

An upper bound for the bit-error probability was also provided in [5]. Since the bound $\sigma_{\widetilde{N}}(k,\theta) \le \sigma_{\widetilde{N}}(k)$ is always verified for every θ, the upper bound can be computed by substituting $\sigma_{\widetilde{N}}(k,\theta)$ in (16) by the standard deviation of the per-channel multiplication error $\sigma_{\widetilde{N}}(k)$. Refer to [5] for more details on the analysis.

The upper bound formula allows to link directly the per-channel WNR and the per-channel bit-error probability. In fact, according to (11), we can substitute $\sigma_{\widetilde{X}}(k) = (\sqrt{3}/\Delta)\sigma_{\widetilde{W}}(k)$ into (16) and using the bound $\sigma_{\widetilde{N}}(k,\theta) \le \sigma_{\widetilde{N}}(k)$ we have

$$\begin{aligned} P_e(k) &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{e,\mathrm{RDM}}\left(L, \frac{\sqrt{3}\ \sigma_{\widetilde{W}}(k)}{\sigma_{\widetilde{N}}(k)}\right)d\theta \\ &= P_{e,\mathrm{RDM}}\left(L, \sqrt{3\ \mathrm{WNR}(k)}\right). \end{aligned} \tag{17}$$

If this analytical model is applied to colored hosts, the predicted error probabilities will be only an approximation of the actual BERs. The inaccuracy of the analytical model is expected to be noticeable for those DFT channels whose $\widetilde{X}_{m,k}$ is more correlated with the neighboring channels; in this case, the per-channel multiplication error will increase due to the leakage from those host samples at adjacent channels. To evaluate the correlation between the kth channel and the tth channel, the correlation coefficient $\rho_{k,t}$ can be employed. Using the approximate expression of the per-channel host power we can write

$$\begin{aligned} \rho_{k,t} &\triangleq \frac{E\left\{\widetilde{X}_k \widetilde{X}_t^*\right\}}{\left(E\left\{\left|\widetilde{X}_k\right|^2\right\}E\left\{\left|\widetilde{X}_t\right|^2\right\}\right)^{1/2}} \\ &\approx \frac{\left|A\left(e^{j2\pi k/N}\right)\right|\,\left|A\left(e^{j2\pi t/N}\right)\right|}{N\sigma_{x_0}^2} E\left\{\widetilde{X}_k \widetilde{X}_t^*\right\}. \end{aligned} \tag{18}$$

The analysis carried out here for DFT-RDM and colored hosts gives a first explanation of the experimental results that were given in [5] for DFT-RDM applied to audio signals.

5. Whitening and DFT-RDM

From the analysis of DFT-RDM for colored hosts developed in Section 4, any colored host will have unavoidably different watermark signal powers for different DFT channels; consequently, there will be some DFT channels more exposed than others to the per-channel multiplication error, as it has been explained above. Assuming that neither the embedder nor the decoder has any prior knowledge about the attack filter, it is reasonable to embed in every DFT channel with the same watermark power. Clearly, this choice does not assure the best BER for every attack filter but it is a trade-off to have a good BER even if the attack filter is unknown. The optimum would be to shape the per-channel watermark power so that it is larger in those DFT channels which are less modified by the attack filter, but this assumes prior knowledge; so we have decided not to follow this path.

On the other hand, according to [5], if the host signal is white, the per-channel multiplication error is approximately independent on both the host and the watermark signal, so the correlation between neighboring channels, which usually leads to higher per-channel error probabilities, becomes small.

These considerations lead to whiten the host signal and use the same embedding power on every DFT channel.

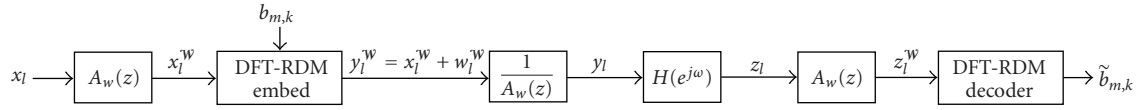

Figure 3: Block scheme of the whole embedding/decoding chain for W-DFT-RDM.

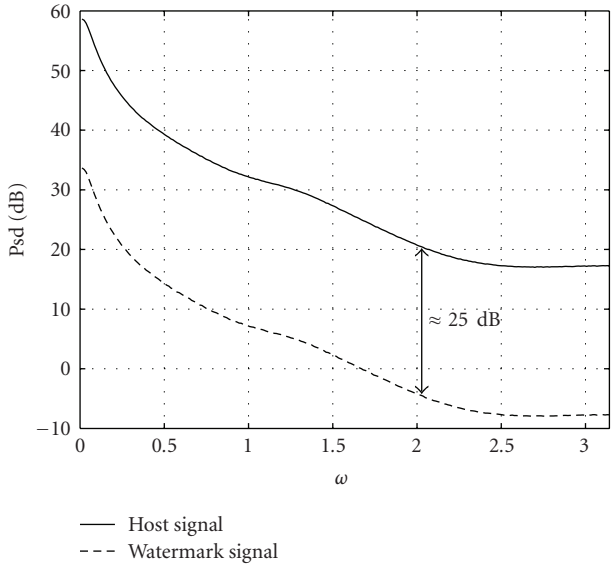

Figure 4: Experimental power spectral densities of host and watermark signal after reconstruction filtering for DWR = 25 dB.

The idea is then to perform the DFT-RDM embedding in the host signal $\mathbf{x}^{\mathcal{W}}$ obtained as the output of a whitening filter $A_w(z)$ excited by the colored host $\mathbf{x}$. Hereinafter the superscript $\mathcal{W}$ is used to denote signals which are obtained by whitening. After the embedding, the watermarked signal $\mathbf{y}^{\mathcal{W}}$ is filtered by the inverse of the whitening filter to reshape the signal, as shown in Figure 3, where the whole block scheme is depicted. At the decoder side the received host signal $\mathbf{z}$ feeds the whitening filter $A_w(z)$ and from the obtained signal $\mathbf{z}^{\mathcal{W}}$ the DFT-RDM decoder recovers the estimated hidden message.

In this section we will assume that the host is an AR random signal which is generated as described in Section 4 by the all-pole filter $1/A(z)$. If the whitening filter $A_w(z)$ is equal to $A(z)$, then we have $\mathbf{x}^{\mathcal{W}} = \mathbf{x_0}$, which is a white Gaussian host with power spectral density $\sigma^2_{x_0}$ by construction of the colored signal. After DFT-RDM embedding, the watermarked signal can be expressed as $\mathbf{y}^{\mathcal{W}} = \mathbf{x}^{\mathcal{W}} + \mathbf{w}^{\mathcal{W}}$. Since DFT-RDM embedding is performed on the white signal $\mathbf{x_0}$, the resulting watermark signal $\mathbf{w}^{\mathcal{W}}$ can be also assumed to be white and uncorrelated with the host signal from the properties of DFT-RDM. Consequently, the reconstruction filter shapes both the host and watermark signal in the same way, so that their power spectral densities have approximately the same trend, as it is shown in Figure 4.

Moreover, given the whiteness of the watermark signal and the superposition principle, the overall DWR is not changed by the reconstruction filter:

$$\mathrm{DWR} = \frac{\sigma_x^2}{\sigma_w^2} = \frac{\int_{-\pi}^{\pi} \sigma^2_{x_0} / \left|A(e^{j\omega})\right|^2 \, d\omega}{\int_{-\pi}^{\pi} \sigma^2_{w^{\mathcal{W}}} / \left|A(e^{j\omega})\right|^2 \, d\omega} = \frac{\sigma^2_{x_0}}{\sigma^2_{w^{\mathcal{W}}}}, \tag{19}$$

and it is approximately equal to the DWR measured on each DFT-RDM channel, as expected according to (11). Thus, even if DFT-RDM is applied to the host signal after whitening, the relation between the overall DWR and Δ is the same as in DFT-RDM, as described in [5]. From this it can be inferred that DFT-RDM with whitening does not incur in any penalty in terms of embedding distortion with respect to DFT-RDM, which is a desirable property of the proposed extension.

At the decoder side, after the whitening filter $A_w(z)$, we have $z_l^{\mathcal{W}} = y_l^{\mathcal{W}} * h_l$; hence the white watermarked signal $\mathbf{y}^{\mathcal{W}}$ goes through an equivalent channel where there is only the attack filter. Consequently, even if the host is colored, using the above proposed scheme we expect the same performance as for DFT-RDM applied to a white host for the same attack filter and the same system parameters.

We have tested the above presented scheme with audio signals. Since audio signals are nonstationary and the whitening filter $A_w(z)$ is the inverse of an AR filter which resembles the spectral contents of a generic audio signal, we can no longer expect $\mathbf{x}^{\mathcal{W}}$ to be really a white signal. However, $\mathbf{x}^{\mathcal{W}}$ will usually have a per-channel host power more evenly distributed than the original host.

6. Experimental Results

Some experiments are here presented to validate the analysis carried out in Section 4 and to verify the effectiveness of DFT-RDM applied to colored hosts after a whitening filtering. In all the experiments the DWR was set to 25 dB, in the g function the memory L was set to 100 and p was set to 2. An AR model with order $Q = 10$ is assumed in all the experiments. Unless otherwise specified, we assume that the DFT length is $N = 512$ and that neither spreading nor windowing is used.

The colored host signal is the output of an all-pole filter $1/A_{\mathrm{av}}(z)$ whose coefficients have been obtained by AR modeling of several audio tracks in order to resemble the power spectral density of a generic (average) audio signal; Figure 5 represents the magnitude of the frequency response of the filter $A_{\mathrm{av}}(e^{j\omega})$ that has been used in the subsequent simulations.

Figure 6 illustrates the per-channel watermark signal power; while the matching between the experimental results and the analytical values obtained substituting (10) in (11) is excellent, a mismatch in the high-frequency channels is

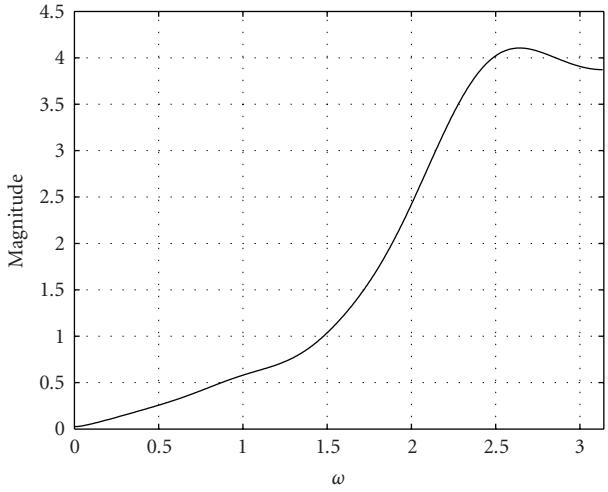

FIGURE 5: Magnitude of the frequency response of the filter $A_{\text{av}}(e^{j\omega})$ with order $Q = 10$.

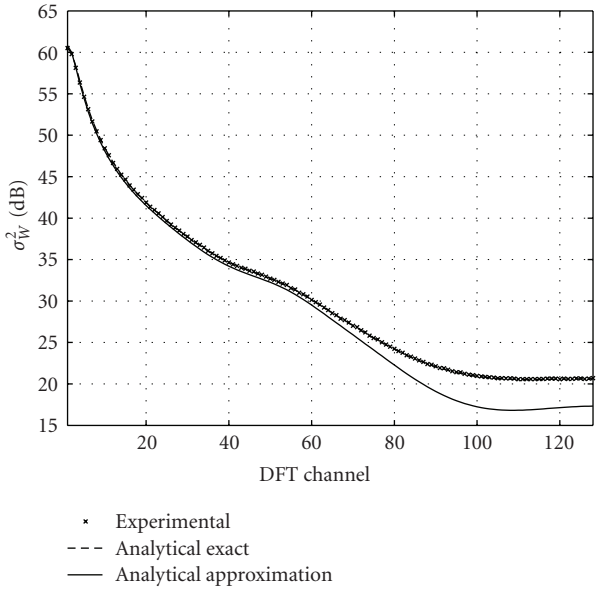

FIGURE 6: Per-channel watermark signal power in dB ($\sigma^2_{x_0} = 1000$ and $N = 256$).

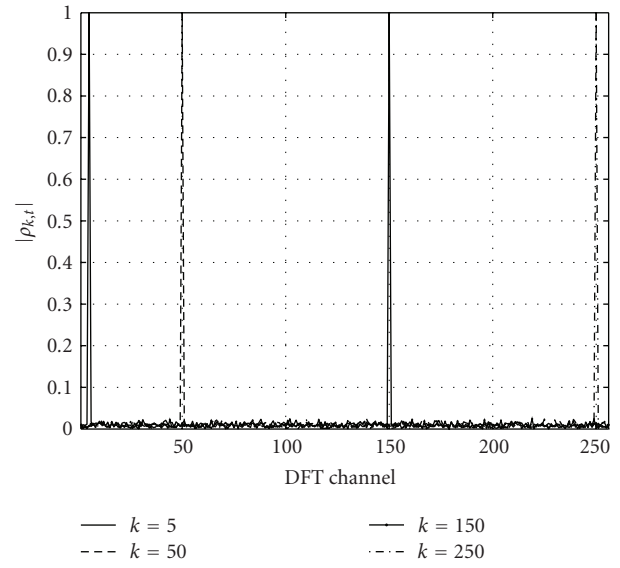

FIGURE 7: Magnitude of the correlation coefficient $|\rho_{k,t}|$ for white watermarked signal evaluated at channels $k = 5$, $k = 50$, $k = 150$, and $k = 250$.

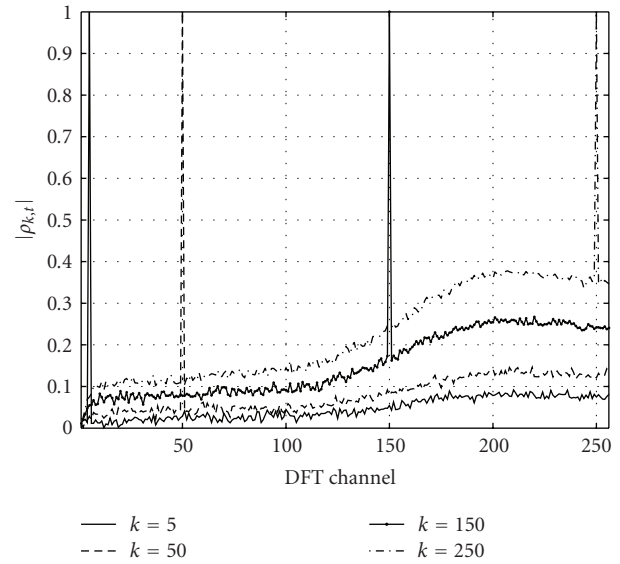

FIGURE 8: Magnitude of the correlation coefficient $|\rho_{k,t}|$ for colored watermarked signal evaluated at channels $k = 5$, $k = 50$, $k = 150$ and $k = 250$.

apparent when using in (11) the approximate formula (14) for the per-channel host power.

In order to verify the existing correlation between channels for colored hosts, the magnitude of the correlation coefficient $|\rho_{k,t}|$ has been evaluated on the watermarked signal $\widetilde{Y}_{m,k}$ according to (18).

First, we plot in Figure 7 the magnitude of the correlation coefficient for several DFT channels when the watermarked signal is white Gaussian. As it can be verified, the correlation between neighboring channels is very small for all k, t, with $k \neq t$. Obviously, for $k = t$ we have $\rho_{k,t} = 1$, since the correlation coefficient corresponds to the normalized autocorrelation.

In contrast, for a colored host the correlation coefficient is strictly dependent on the selected channels, as it is evident in Figure 8. As expected from (18) for a high-pass filter $A(z)$, the correlation between two low-frequency neighboring channels is quite small, while it noticeably increases when neighboring higher-frequency pairs are considered. Since the analytical results are less accurate when DFT channels become more correlated, we should expect worse predictions for high-frequency channels.

Then we have tested the watermarking system with a lowpass filter with cut-off frequency $\omega_c = 0.8\pi$ rad.

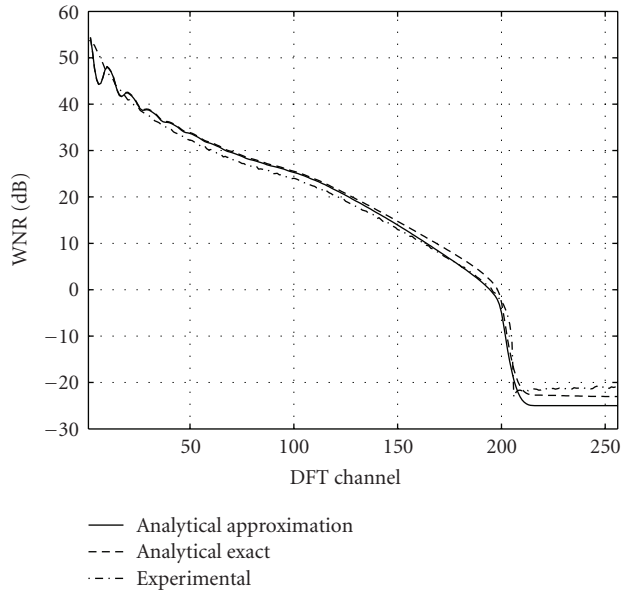

Figure 9: WNR versus DFT channel for colored host and lowpass filter with $\omega_c = 0.8\pi$.

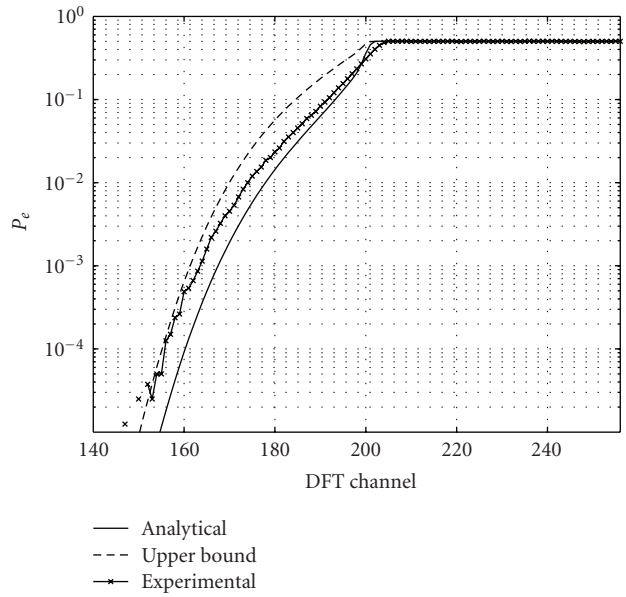

Figure 10: BER versus DFT channel for colored host and lowpass filter with $\omega_c = 0.8\pi$.

Figure 9 compares the experimentally evaluated WNR with the analytical WNR computed according to (13) and the analytical approximation of the WNR obtained from (15). It is worth noting that the WNR is much larger for the low-frequency channels where the host power is also larger and the filter response is flat. In Figure 10 the experimental BER is compared with the analytically derived BER and its upper bound, according to the formulas in [5] (here and in the following, the analytical BER is computed using the exact formula of the per-channel signal power); here we show only the range of channels having an experimental BER larger than 10^{-5}. From the comparison of Figures 9 and 10 it can be verified that the error probability is approximately 0.5 for those DFT channels whose WNR is lower than 0 dB, as we have already discussed.

To understand how different AR models influence the WNR, the analytical WNR for the lowpass filter has been computed using (13) for different orders Q of the AR model. In Figure 11 the WNR for AR(10) is compared with the analytical WNR computed for AR(3), AR(7), AR(50) and AR(100). For $Q = 3$ the WNR is slightly lower than that of $Q = 10$, while for $Q = 7$ approximately the same WNR of $Q = 10$ is obtained. As the order of AR model increases, the WNR has more ripples but it has always the same average trend of that for AR(10), as it is shown in Figures 11(c) and 11(d). We conclude that the order of the AR model has little impact on the final results.

Then we have tested the watermarking system with a lowpass filter having passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad, with a smooth transition in the middle. Figure 12 compares the experimental WNR and the analytical one; there is a noticeable difference in the frequency range where the interchannel correlation is larger. In Figure 13 the experimental BER is compared with the analytically derived BER and its upper bound (again only the range of channels having an experimental BER larger than 10^{-5} is shown). Moreover, it can be seen that the analytical error probability matches the experimental one since all the channels with $P_e < 0.5$ are not in the range of high correlation. It is worth noting that the error probability is approximately 0.5 in the majority of channels belonging to the transition band, which is approximately between channels 102 and 204.

Finally, a ten-band graphic audio equalizer has been used as attack filter. In the following experiments we have set the equalizer subband filters so that they produce the overall frequency response depicted in Figure 14, which is the same that was used in the experiments presented in [5]. Figures 15 and 16 illustrate the analytical WNRs and the comparison of the experimental BERs with the analytical ones, respectively. From the per-channel WNR shown in Figure 15 it can be inferred that the expected error probability will be very high, especially for the high-frequency channels, and this behavior is confirmed by the experimental BERs shown in Figure 16. One can also notice that the analytical error probabilities provide a good prediction only for the low-frequency channels. We conjecture that the observed inaccuracies are due to the correlation among neighboring channels of the colored host, which could be increased even further by the equalizer. Above all, this experiment reveals that DFT-RDM applied to a colored host does not guarantee at all the robustness of the watermark against an equalizer attack, especially as neither windowing nor spreading is here used, since the overall BER is approximately 0.48. It is worth noting that by embedding the watermark with the same system parameters into a white Gaussian host, the overall BER is approximately 0.21. On the other hand, the analysis and the experiments carried out for DFT-RDM with a colored host and an equalizer attack provide a qualitative explanation

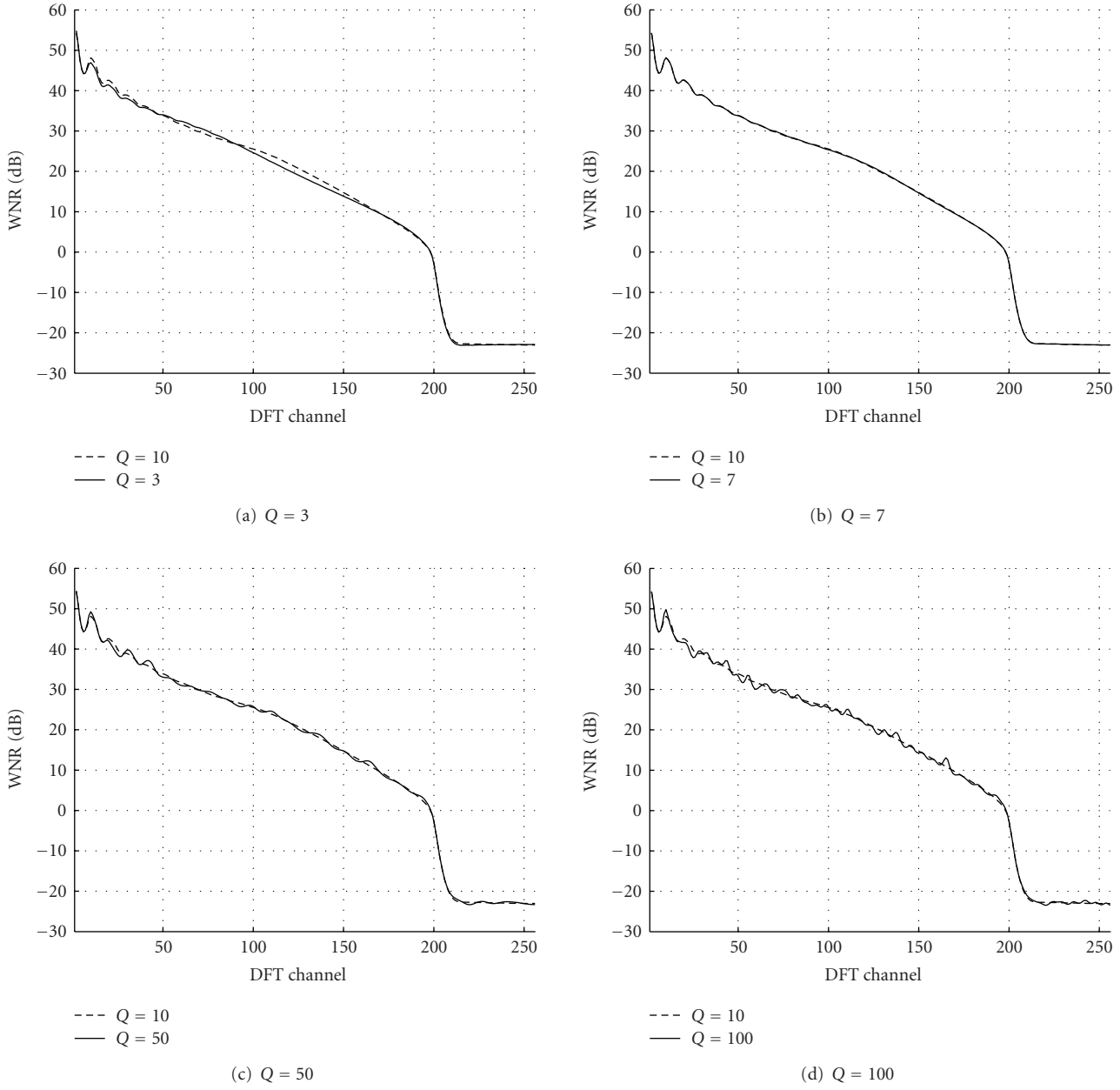

Figure 11: Analytical WNR versus DFT channel for different order AR filters.

of the experimental results reported in [5] for DFT-RDM applied to audio signals.

Some experiments were conducted to verify the effectiveness of the extension of DFT-RDM proposed in Section 5, hereinafter denoted by the subscript W-DFT-RDM; in the following, the host will be assumed to be colored by $1/A_{\text{av}}(z)$, whereas perfect whitening is assumed, that is, $A_w(z) = A_{\text{av}}(z)$.

First of all, we have compared the performance of W-DFT-RDM with that of DFT-RDM applied to both white and colored hosts. The experimental BERs measured for different attack filters are presented in Figures 17, 18 and 19, where it is verified that the BER of DFT-RDM applied to a white host matches always that of W-DFT-RDM applied to a colored host, as it was expected.

In Figure 17 are shown the experimental BERs measured for the lowpass attack filter with cut-off frequency $\omega_c = 0.8\pi$ rad. It can be noticed that for the given attack filter, the overall error probability of DFT-RDM applied directly to the colored host is $P_e \approx 0.12$, which is less than the overall error probability of DFT-RDM for a white host ($P_e \approx 0.134$). This behavior can be easily explained by the fact that the per-channel watermark signal power is larger at low-frequency channels which are not modified at all by the attack filter. This result confirms the conclusion that whitening does not always assure the best BER for every attack filter.

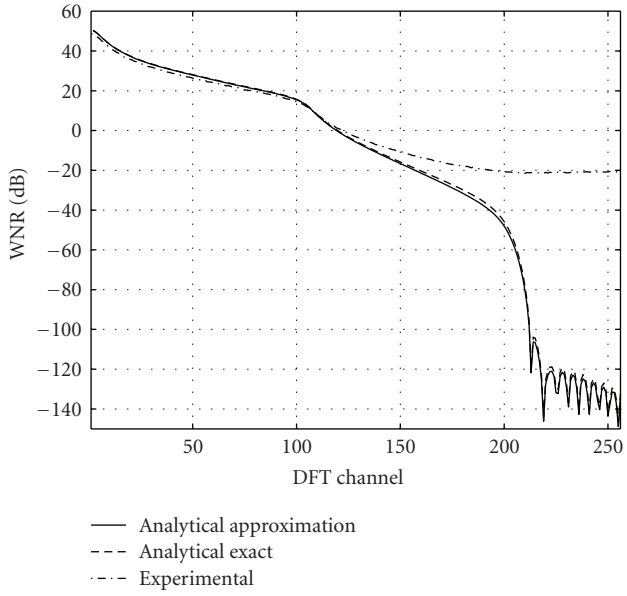

Figure 12: WNR versus DFT channel for colored host and lowpass filter with passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad.

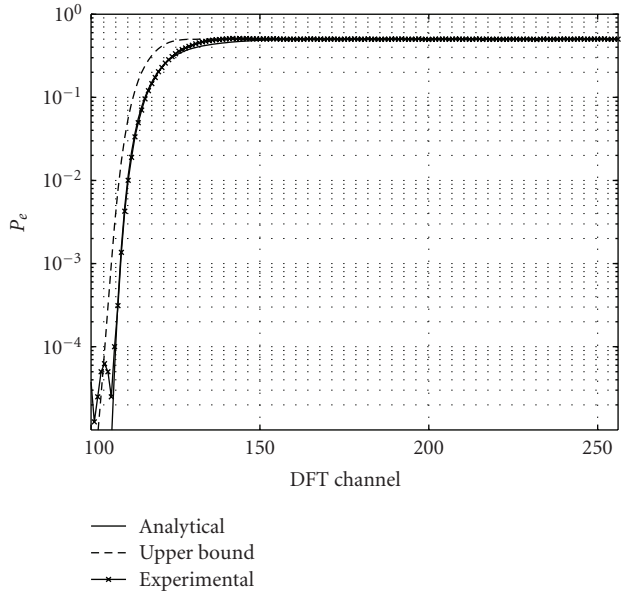

Figure 13: Ber versus DFT channel for colored host and lowpass filter with passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad.

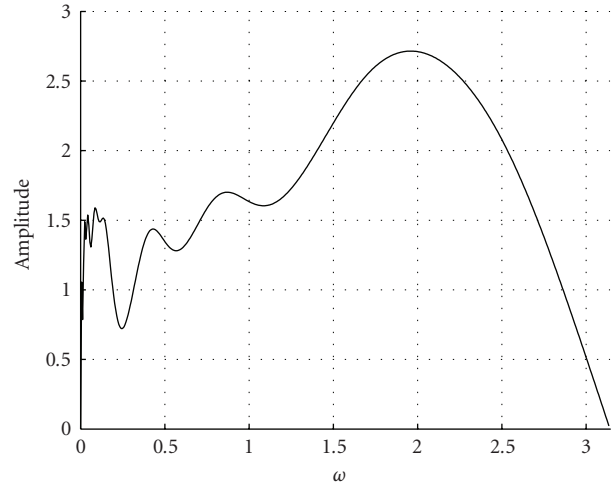

Figure 14: Magnitude of the ten-band audio equalizer used in the experiments.

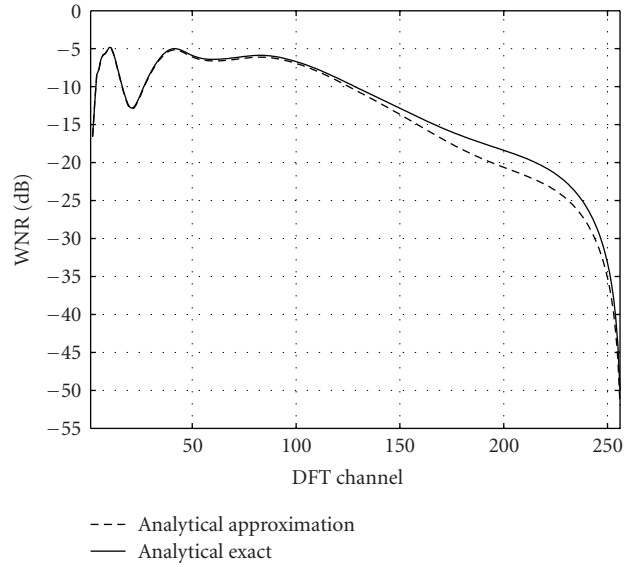

Figure 15: WNR versus DFT channel for colored host and ten-band equalizer attack.

Then we have tested the watermarking methods with the lowpass filter having passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad, with a smooth transition in the middle. The experimental BERs are shown in Figure 18. In this case, the error probability of W-DFT-RDM is approximately 0.5 only in the stopband, while for DFT-RDM applied to a colored host it is 0.5 in the transition band too. This yields the overall error probability of DFT-RDM ($P_e \approx 0.26$), which is larger than that of W-DFT-RDM ($P_e \approx 0.21$).

In Figure 19 are shown the BERs for the ten-band graphic audio equalizer. With this attack filter, since the filtering effect is spread over all frequencies, W-DFT-RDM outperforms DFT-RDM for colored hosts (the overall error probabilities are respectively $P_e \approx 0.21$ and $P_e \approx 0.48$).

We have also compared the behavior of W-DFT-RDM and of DFT-RDM using real audio tracks sampled at 44.1 kHz with 16 bits as host signal. These experiments have been conducted using for all the audio tracks a fixed whitening filter, which is again $A_w(z) = A_{\text{av}}(z)$. We remark here that perfect whitening does not occur with audio tracks since the whitening filter $A_w(z)$ is the inverse of an AR filter which resembles the spectral contents of a generic audio signal. The measured DWRs have been obtained fixing the target DWR at 25 dB; we remark here that with nonstationary, non-Gaussian and nonwhite hosts

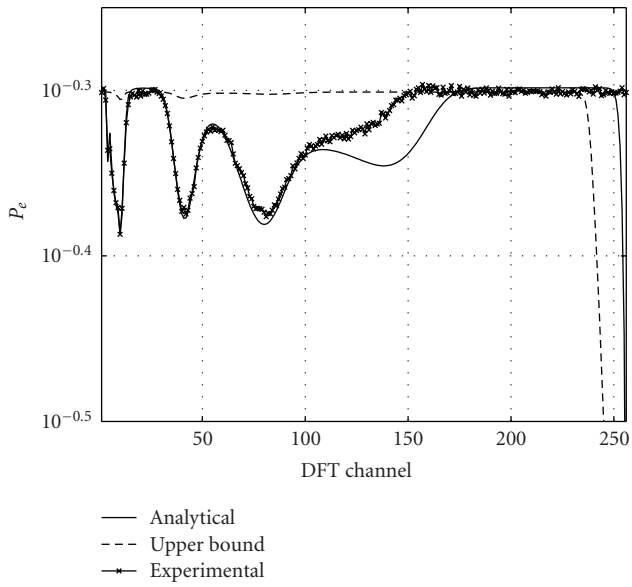

Figure 16: BER versus DFT channel for colored host and ten-band equalizer attack.

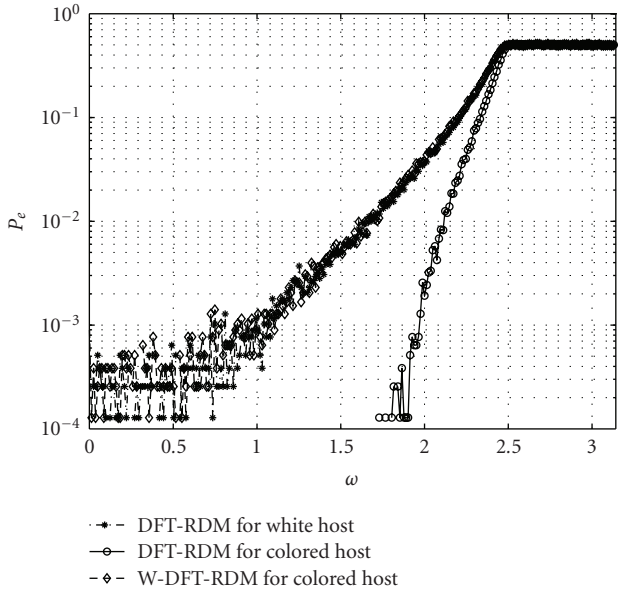

Figure 17: BERs versus discrete frequency for lowpass filter with $\omega_c = 0.8\pi$ rad.

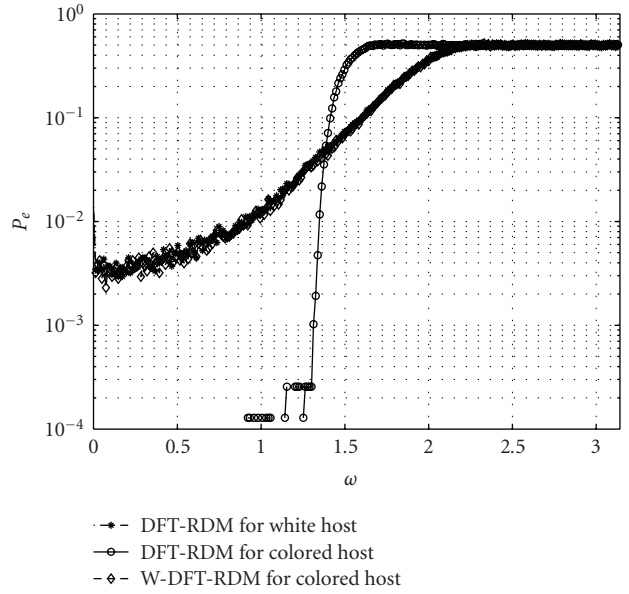

Figure 18: BERs versus discrete frequency for lowpass filter with passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad.

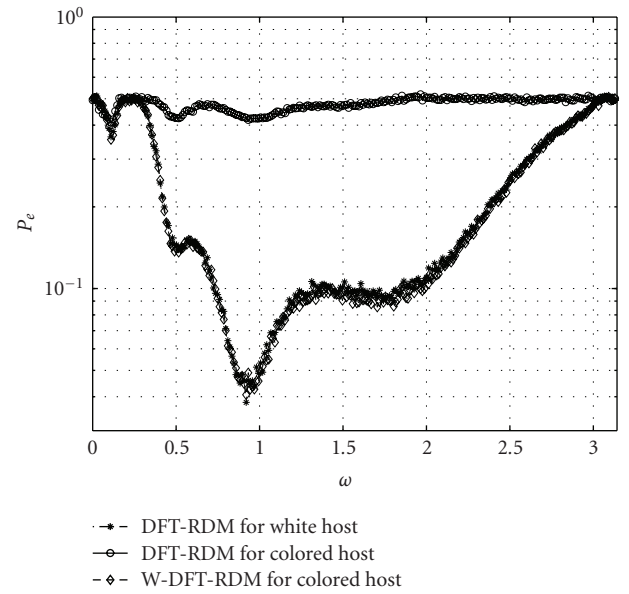

Figure 19: BERs versus discrete frequency for the ten-band equalizer attack.

the analytical derivation of the DWR for DFT-RDM is only an approximation.

In Tables 1, 2, and 3 the overall error probabilities evaluated for a spreading factor $M = 1$ (i.e., no spreading) and a rectangular window are given. Notice that in all the experiments, for the same audio track, the DWRs produced by the two embedding techniques are approximately equal.

Table 1 shows the experimental results for the lowpass filter with cut-off frequency $\omega_c = 0.8\pi$ rad. As it was to be expected from the results presented before for a colored host, for audio signals DFT-RDM has also lower bit error probabilities than W-DFT-RDM. Similar results have been obtained attacking the watermarked host with the lowpass filter having passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad. As it is shown in Table 2, the overall error probabilities for DFT-RDM are mostly lower than the respective ones for W-DFT-RDM; however, the behavior depends on the particular audio track, as it can be noticed from the results obtained for the tracks "Spff" and "Spfg." In contrast, for the ten-band equalizer attack, W-DFT-RDM yields an improved overall BER for all the audio tracks.

Table 1: Overall error probabilities for the lowpass filter with $\omega_c = 0.8\pi$ rad ($M = 1$ and rectangular window).

Track	DFT-RDM		W-DFT-RDM	
	DWR (dB)	BER	DWR (dB)	BER
Bass	24.82	0.110	24.72	0.131
Jarre	25.01	0.125	24.98	0.185
REM	24.96	0.102	24.75	0.129
Sopr	24.97	0.116	24.79	0.131
Spff	24.81	0.108	24.97	0.114
Spfg	24.66	0.106	24.53	0.115
Trpt	25.05	0.100	24.79	0.114
Vioo	25.23	0.105	25.23	0.162

Table 2: Overall error probabilities for the lowpass filter with passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad ($M = 1$ and rectangular window).

Track	DFT-RDM		W-DFT-RDM	
	DWR (dB)	BER	DWR (dB)	BER
Bass	24.82	0.237	24.74	0.249
Jarre	24.97	0.274	24.96	0.299
REM	24.96	0.219	24.73	0.263
Sopr	24.98	0.235	24.79	0.264
Spff	24.79	0.247	24.94	0.179
Spfg	24.64	0.237	24.52	0.172
Trpt	25.03	0.177	24.76	0.264
Vioo	25.24	0.246	25.22	0.283

We must remark that the BERs given above for both DFT-RDM-based schemes would be unacceptable in a watermarking application, thus the experiments have been repeated using a spreading factor $M = 8$ and the optimal window, which has been computed according to [5]. We remind that spreading grants a robustness improvement at the expense of a reduction of the data rate, which becomes 1/16 bits/sample for $M = 8$. From the inspection of the DWRs listed in Tables 4, 5, and 6 it can be noticed that in all the experiments, for the same audio track, the DWRs produced by the two embedding techniques are approximately equal.

Table 4 shows the results for the lowpass filter with cut-off frequency $\omega_c = 0.8\pi$ rad. Here, for every audio track, both DFT-RDM-based schemes reach the minimum error probability, which corresponds to the correct detection of all those watermark bits embedded in DFT channels within the passband and is approximately 0.1.

From the comparison of the results for the lowpass attacking filter with passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad, that are listed in Table 5, we can notice that whitening yields a minimum error probability, that is again approximately 0.1, in almost all the experiments. Moreover, DFT-RDM has always an overall error probability higher than W-DFT-RDM and away from the minimum error probability.

The overall error probabilities presented in Table 6 confirm the better behavior of W-DFT-RDM for the equalizer attack. In fact, for every audio track the BER of W-DFT-RDM is always lower, with an improvement with respect to DFT-RDM that goes from a factor of 1.5 to 7 in terms of error probability, depending on the audio track.

Even though linear filtering does not encompass MPEG Layer-3 (MP3) compression, this can be very roughly seen as a lowpass filtering with cut-off frequency equal to the sampling frequency of the audio track after MP3 compression. Hence, we have conducted several experiments to verify the robustness of DFT-RDM-based techniques to MP3 compression. The real audio tracks, whose sampling frequency is 44.1 kHz, have been marked, compressed using LAME 3.97 [8] to perform MP3 encoding, and, finally, the watermark has been retrieved. In Table 7 are listed the BERs measured for both DFT-RDM-based techniques using a spreading factor $M = 8$ and the optimal window. These results have been obtained for constant bit-rate MP3 encoding of the watermarked audio tracks, but approximately the same error probabilities have been measured for average bit-rate MP3 encoding. It is worth noting that in these experiments the minimum error probability is approximately 0.137, that corresponds to the correct detection of all the watermark samples embedded up to 32 kHz, which is the sampling frequency of the audio tracks compressed by LAME for the considered bit-rates. From the inspection of the results in Table 7, it can be noticed that the error probabilities of W-DFT-RDM are always lower than those of DFT-RDM for the same bit-rate. Even if the measured error probabilities are considerably dependent on the particular audio track, W-DFT-RDM approaches the minimum error probability for almost all audio tracks and an encoding bit-rate equal to

Table 3: Overall error probabilities for the ten-band equalizer attack ($M = 1$ and rectangular window).

Track	DFT-RDM		W-DFT-RDM	
	DWR (dB)	BER	DWR (dB)	BER
Bass	24.81	0.463	24.70	0.359
Jarre	24.99	0.471	24.98	0.308
REM	24.95	0.481	24.76	0.392
Sopr	24.96	0.457	24.78	0.344
Spff	24.79	0.370	24.93	0.166
Spfg	24.69	0.364	24.55	0.149
Trpt	25.01	0.488	24.74	0.481
Vioo	25.21	0.493	25.19	0.360

Table 4: Overall error probabilities for the lowpass filter with $\omega_c = 0.8\pi$ rad ($M = 8$ and optimal window).

Track	DFT-RDM		W-DFT-RDM	
	DWR (dB)	BER	DWR (dB)	BER
Bass	22.96	0.112	22.93	0.100
Jarre	22.99	0.100	23.11	0.101
REM	26.60	0.100	26.33	0.102
Sopr	23.78	0.110	23.74	0.100
Spff	23.33	0.101	23.45	0.099
Spfg	25.37	0.100	25.25	0.100
Trpt	31.01	0.101	30.73	0.101
Vioo	25.29	0.101	25.27	0.099

Table 5: Overall error probabilities for the lowpass filter with passband $[0, 0.4\pi]$ rad and stopband $[0.8\pi, \pi)$ rad ($M = 8$ and optimal window).

Track	DFT-RDM		W-DFT-RDM	
	DWR (dB)	BER	DWR (dB)	BER
Bass	22.95	0.235	22.94	0.113
Jarre	22.98	0.139	23.10	0.101
REM	26.60	0.176	26.30	0.102
Sopr	23.82	0.241	23.76	0.101
Spff	23.36	0.148	23.45	0.100
Spfg	25.36	0.141	25.26	0.100
Trpt	31.05	0.247	30.74	0.158
Vioo	25.25	0.213	25.30	0.109

Table 6: Overall error probabilities for the ten-band equalizer attack ($M = 8$ and optimal window).

Track	DFT-RDM		W-DFT-RDM	
	DWR (dB)	BER	DWR (dB)	BER
Bass	22.96	0.229	22.90	0.0325
Jarre	23.00	0.0380	23.14	0.0180
REM	26.61	0.0688	26.33	0.0130
Sopr	23.79	0.248	23.71	0.0383
Spff	23.34	0.0580	23.44	0.0115
Spfg	25.36	0.0542	25.27	0.0349
Trpt	31.02	0.3514	30.75	0.0512
Vioo	25.33	0.186	25.29	0.0282

Table 7: Overall error probabilities for MP3 compression attacks ($M = 8$ and optimal window).

Track	DFT-RDM				W-DFT-RDM			
	DWR (dB)	80 kbps	160 kbps	320 kbps	DWR (dB)	80 kbps	160 kbps	320 kbps
Bass	22.96	0.389	0.346	0.322	22.90	0.339	0.232	0.145
Jarre	23.00	0.409	0.229	0.155	23.14	0.402	0.213	0.140
REM	26.61	0.405	0.258	0.223	26.33	0.360	0.187	0.143
Sopr	23.79	0.399	0.354	0.337	23.71	0.345	0.220	0.146
Spff	23.34	0.292	0.185	0.148	23.44	0.280	0.175	0.138
Spfg	25.36	0.252	0.172	0.144	25.27	0.246	0.167	0.138
Trpt	31.02	0.402	0.371	0.366	30.75	0.389	0.346	0.328
Vioo	25.33	0.389	0.319	0.290	25.29	0.363	0.257	0.167

320 kbps. On the other hand, the BERs measured for DFT-RDM can be far away from the minimum error probability even if the audio tracks are encoded at the maximum allowed bit-rate.

7. Conclusions

A thorough analysis of the behavior of DFT-RDM for colored Gaussian hosts has been performed. An explanation to the performance loss with respect to white Gaussian hosts has been given. We have also provided an extension of DFT-RDM for colored hosts without any additional knowledge on the attack filter; this extension consists in using a fixed whitening filter that captures the average properties of audio signals. The analysis has been validated by experimental results which confirm the performance improvement afforded by the proposed solution. Moreover W-DFT-RDM has been tested with audio signals providing a BER decrease which encourages us to continue on this research line. W-DFT-RDM for audio tracks is not able to fill the performance gap with respect to DFT-RDM for white hosts since a fixed (and nonperfectly matched) average whitening filter is used at both the embedder and the decoder. A further improvement could be obtained by using a host-adaptive whitening filter at the embedder which, assuming a blind framework, should be retrieved at the decoder side, at least with some approximation. Finally, even though encouraging BER results have been obtained for MP3 compression, an accurate analysis of DFT-RDM-based techniques against compression is needed in order to assess the real bounds.

References

[1] B. Chen and G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423–1443, 2001.

[2] I. D. Shterev and R. L. Lagendijk, "Amplitude scale estimation for quantization-based watermarking," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4146–4155, 2006.

[3] M. L. Miller, G. J. Doërr, and I. J. Cox, "Applying informed coding and embedding to design a robust high-capacity watermark," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 792–807, 2004.

[4] F. Pérez-González, C. Mosquera, M. Barni, and A. Abrardo, "Rational dither modulation: a high-rate data-hiding method invariant to gain attacks," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3960–3975, 2005.

[5] F. Pérez-González and C. Mosquera, "Quantization-based data hiding robust to linear-time-invariant filtering," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, pp. 137–152, 2008.

[6] J. Wang, I. D. Shterev, and R. L. Lagendijk, "Scale estimation in two-band filter attacks on QIM watermarks," in *Security, Steganography, and Watermarking of Multimedia Contents VIII*, E. J. Delp III and P. W. Wong, Eds., vol. 6072 of *Proceedings of SPIE*, pp. 118–127, San Jose, Calif, USA, January 2006.

[7] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing*, Prentice Hall, Upper Saddle River, NJ, USA, 4th edition, 2006.

[8] The Lame Project, http://lame.sourceforge.net/.