EURASIP Journal on Advances in Signal Processing
a SpringerOpen Journal

RESEARCH Open Access

Performance analysis and improvement of dither modulation under the composite attacks

Xinshan Zhu[1,2*] and Jie Ding[1,3]

Abstract

The first goal of this article is to analyze the performance of dither modulation (DM) against the composite attacks including valumetric scaling, additive noise and constant change. The analyzes are developed under the assumptions that the host vector and noise vector are mutually independent and both of them have independently and identically distributed components. We derive the general expressions of the probability density functions of several concerned signals and the decoding error probability. The specific theoretical results are provided for the case of generalized Gaussian host and noise. Based on the analyzes, the performance of DM is predicted for different scenarios with a high degree of accuracy and evaluated for different distribution models of host and noise signals. Numerical simulations confirm the validity of the given theoretical analyzes. Then, we address to improve the robustness of DM against valumetric scaling plus constant change. The normalized dither modulation (NDM) is presented, which works by constructing a gain-invariant vector with zero mean for quantization. Performance analysis shows that NDM is theoretically invariant to valumetric scaling and constant change and achieves similar performance to DM in other aspects. The performance of NDM is further improved by weighting the quantization errors. Experiments on real images also show the advantage of NDM over DM subject to amplitude scaling and constant change.

Keywords: digital watermarking, quantization index modulation, composite attacks, valumetric scaling, constant change

1 Introduction

In the past decade, much attention has been paid to the quantization-based watermarking for canceling the host signal interference. One of the most important methods proposed so far is quantization index modulation (QIM) [1]. The basic QIM algorithm includes a number of variants, i.e., dither modulation (DM), distortion compensated dither modulation (DC-DM) (also known as scalar Costa scheme (SCS) [2]) and spread transform dither modulation (STDM) [1]. The theoretical performance of QIM methods is a key issue and has received considerable attention.

Initially, the Gaussian channel is often used in the analyzes and the performance of QIM methods has been extensively investigated in this case. A relatively crude approximation to the decoding error probability of QIM was given in [1] for the additive white Gaussian noise (AWGN) attacks. The performance of SCS was completely analyzed by Eggers et al. [2] under the AWGN attacks. In [3], the performance of scalar DC-QIM against AWGN was theoretically evaluated from the detection viewpoint. Recently, a new logarithmic QIM (LQIM) was presented in [4] and its performance was analyzed in the presence of AWGN. It has been pointed out in [5] that the performance of QIM methods may be overstated under Gaussian channels. In the second phase, a deeper analysis is done for QIM taking into account a much wider variety of attacks. The careful performance analyzes were presented by Pérez-Gonzàlez et al. [5] for a large class of QIM methods in the cases of uniform and Gaussian noise. Bartolini et al. [6] concentrated on analyzing the performance of STDM in the presence of two important classes of non additive attacks, the gain attack plus noise addition and the quantization attack. In [7], the authors proposed an improved DM scheme to resist linear-time-invariant

* Correspondence: xszhu_hm@hotmail.com
[1]School of Information Engineering, Yangzhou University, Yangzhou 225009, China
Full list of author information is available at the end of the article

filtering and provided a thorough analysis of it. We notice that most of previous analyzes make use of the Gaussian host assumption and even neglect the statistical properties of the host signal.

The conventional QIM has a serious drawback, i.e., the weakness against valumetric scaling. Spherical codes were utilized to cope with this problem in [8]. However, watermark embedding and recovery get very complicated [9]. Oostveen et al. [10] proposed to choose the adaptive quantization step size to be proportional to a local average of the host signal samples. Despite its robustness against valumetric scaling, the method presents a nonzero probability of error even for null distortions and becomes more sensitive to constant change. Rational dithered modulation (RDM) was developed in [9] using a gain-invariant adaptive quantization step size at both embedder and decoder. The RDM achieves invariance to valumetric scaling, but is still sensitive to constant change. Li and Cox [11] applied the modified Watson's perceptual model to provide resistance to valumetric scaling for QIM watermarking. The modification to Watson's model results in the degradation in quality and performance loss with respect to constant change.

The first objective of this article is to analyze the performance of DM against composite attacks, which is lacking in the literature. Obviously, in watermarking applications, it is more often that the watermark undergoes multiple attacks. Specifically, the combination of valumetric scaling, additive noise and constant change will be considered. On the other hand, most of previous analyzes are restricted to the Gaussian noise channel, even sometime regardless of the distribution of the host signal, which we will try to overcome. The generalized Gaussian distribution (GGD) is adopted to model both the host signal and the noise signal in our analysis. Since the GGD is a parametric family of distributions, we will observe how the choice of distribution model affects the performance of DM. Next, the weakness of DM is concerned. DM itself is largely vulnerable to valumetric scaling as well as constant change. Several existing improved DM schemes achieve the robustness against valumetric scaling, but becomes more sensitive to constant change. We will present the normalized DM (NDM) considering both of them. Under the light of the performance analyzes done for DM in this article, we show that NDM approaches the performance of DM, with the great advantage of insensitivity to valumetric scaling and constant change.

The rest of this article is organized as follows. Section 2 reviews the original DM and describes the problems to be solved. Next, Section 3 accurately derives the general PDF models for several relevant signals. In Section 4, the performance of DM under the composite attacks is mathematically analyzed by the derived PDFs. The decoding error probability is given in closed form and the theoretical results are confirmed by numerical simulations. Then, in Section 5, the NDM method is presented and its performance is theoretically evaluated. Section 6 provides a useful strategy to improve the performance of NDM. In Section 7, a series of tests on real data are done to verify the validity of analytical derivations and evaluate the presented approaches. Finally, Section 8 concludes.

Notation: In the remainder of this article, we use boldface lower-case letters to denote column vectors, e.g., $\boldsymbol{x}$, and scalar variables are denoted by italicized lower-case letters, e.g., x. The probability distribution function (PDF) of a random variable (r.v.) x is denoted by $p_X(x)$, whereas if x is discrete its probability mass function (PMF) is designated by $P_X(x)$. We write $x \sim p_X(x)$ to indicate that a r.v. x is distributed as $p_X(x)$. $\mathrm{p}_{X|Y}(x|y)$ means the conditional probability of x given y. And the subscripts of the distribution functions will be dropped wherever it is clear the random variables they refer to. Finally, the mathematical expectation and standard deviation are respectively represented by μ_x and σ_x for a r.v. x.

2 Review of DM and problem

We will concentrate our attention on DM in this study. The uncoded binary DM can be summarized as follows.

Let $\boldsymbol{x} \in \mathbb{R}^N$ be a host signal vector in which we wish to embed the watermark message m. First, the message m is represented by a vector $\boldsymbol{b}$ with NR_m binary antipodal components, i.e., $b_j = \pm 1$, $j = 1, ..., NR_m$, where R_m denotes the bit rate. The host signal $\boldsymbol{x}$ is then decomposed into NR_m subvectors (blocks) of length $L = \lfloor 1/R_m \rfloor$, denoted by $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{NR_m}$. In the binary DM, two L-dimensional uniform quantizers $Q_{-1}(\cdot)$ and $Q_{+1}(\cdot)$ are constructed, whose centroids are given by the lattices and $\Lambda_{-1} = 2\Delta\mathbb{Z}^L + \boldsymbol{d}$ and $\Lambda_{+1} = 2\Delta\mathbb{Z}^L + \boldsymbol{d} + \Delta\mathbf{a}$ with $\boldsymbol{d} \in \mathbb{R}^L$ a key-dependent dithering vector and $\mathbf{a} = (1, ..., 1)^T$. Each message bit b_j is hidden by using $Q_{bj}(\cdot)$ on $\boldsymbol{x}_j$, resulting in the watermarked signal $\boldsymbol{y} \in \mathbb{R}^N$ as

$$\boldsymbol{y}_j = Q_{b_j}(\boldsymbol{x}_j), \quad j = 1, \ldots, NR_m. \tag{1}$$

The watermark detector receives a distorted, watermarked signal, $\boldsymbol{z}$, and decodes a message $\hat{m}$ using the minimal distance decoder

$$\widehat{b}_j = \arg\min_{-1,1} \left\| Q_{b_j}(\boldsymbol{z}_j) - \boldsymbol{z}_j \right\|, \quad j = 1, \ldots, NR_m, \tag{2}$$

where $||\cdot||$ stands for Euclidean (i.e., ℓ_2) norm.

In practical watermarking applications, the watermarked signal might undergo composite attacks. It is

well known that quantization-based watermarking is vulnerable to valumetric scaling attack. While the vector at the input of the decoder is scaled by ρ_j, i.e., $\boldsymbol{z}_j = \rho_j \boldsymbol{y}_j$, the quantization bins at the decoder are not scaled accordingly, thus producing a mismatch between embedder and decoder that dramatically affects performance. Also, the original DM is not robust to constant change distortion, i.e., $\boldsymbol{z}_j = \boldsymbol{y}_j + c_j\mathbf{a}$ with c_j a constant value. No decoding error is made for $|c_j| < \Delta/2$, however, the bit error rate (BER) is equal to 1 for $\Delta/2 < |c_j| < 3\Delta/2$. In this study, the two attacks are considered together with additive noise $\boldsymbol{v}_j$, yielding the attacked signal as

$$\boldsymbol{z}_j = \rho_j \boldsymbol{y}_j + \boldsymbol{v}_j + c_j\mathbf{a}. \tag{3}$$

We will analyze the performance of DM in the case of (3), and present an improved DM resisting both valumetric scaling and constant change. In the subsequent analysis, $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ and $\boldsymbol{v}$ are all regarded as random vectors. And we assume that both $\boldsymbol{x}$ and $\boldsymbol{v}$ have independently and identically distributed (i.i.d.) components and $\boldsymbol{v}$ is independent from $\boldsymbol{y}$. Since the mean value of additive noise $\boldsymbol{v}_j$ can be counted by the third term in the right-hand side of (3), it is reasonable to assume that $\mu_v = 0$.

3 PDF models

Define the extracted vector $\boldsymbol{r}, \boldsymbol{r} \triangleq Q_b(\boldsymbol{z}) - \boldsymbol{z}$. Obviously, a crucial aspect when performing a rigorous analysis lies in computing the PDF of $\boldsymbol{r}$. Let us begin with the issue.

3.1 PDF model of the watermarked signal

We use a lower-case letter to indicate any element of the vector denoted by the boldface one. The previously used index j is dropped for no specific values (or sub-vectors) are concerned. Given $x \sim p_X(x)$, from the relation (1), the PDF of the watermarked signal y conditioned on a transmitted symbol b is written as

$$p_Y(y|b) = \sum_{k=-\infty}^{\infty} \delta(y - y_k) \int_{y_k-\Delta}^{y_k+\Delta} p_X(x)dx, \tag{4}$$

where the variable y_k is defined as $y_k = 2k\Delta + (b + 1)\Delta/2 + d$ and $\delta(\cdot)$ denotes the delta function.

A few observations are in order about the PDF of y. First, for different dither value d, the PDF $p_Y(y|b)$ is different. That means each element of $\boldsymbol{y}$ obeys different distributions by randomly selecting $\boldsymbol{d}$ during embedding. However, due to the fact $P_Y(y_k+2\Delta|b) = P_Y(y_{k+1}|b)$ exists, it is sufficient for us to consider the case $d \in [-\Delta,\Delta)$. Further, if the PDF $p_X(x)$ is symmetric, i.e., $p_X(x) = p_X(-x)$, from (4), the PDF $p_Y(y)$ satisfies $p_Y(y|b = -1) = p_Y(-y|b = 1)$ for the case of $d = -\Delta/2$ and $p_Y(y|b) = p_Y(-y|b)$ for the case of $d = 0$. The former indicates that the PDFs $p_Y(y|b = -1)$ and $p_Y(y|b = 1)$ are mirrors of each other and the latter indicates that the PDF $p_Y(y)$ is even. These two properties of $p_Y(y)$ are exhibited in Figure 1.

3.2 PDF model of the attacked signal

Taking the Equation (3) into account and using the fact that for any $\rho > 0 p_{\rho Y}(y) = \frac{1}{\rho} p_Y\left(\frac{y}{\rho}\right)$ holds, the conditional PDF of $\boldsymbol{z}$ given the transmitted symbol b can be obtained by convolution [12]

$$p_Z(z|b) = \sum_{k=-\infty}^{\infty} P_Y(y_k|b) p_v(z - \rho y_k - c), \tag{5}$$

where the convolution follows from the independence between $\boldsymbol{y}$ and $\boldsymbol{v}$. Observing (5), if the effect of different d on $P_Y(y)$ is ignored (this generally holds when the embedding distortion is acceptable), the PDF $p_Z(z|b)$ with $d \neq 0$ can be approximately viewed as the translate of $p_Z(z|b)$ with $d = 0$, that is, $p_Z(z + \rho d|b, d \neq 0) \approx p_Z(z|b, d = 0)$.

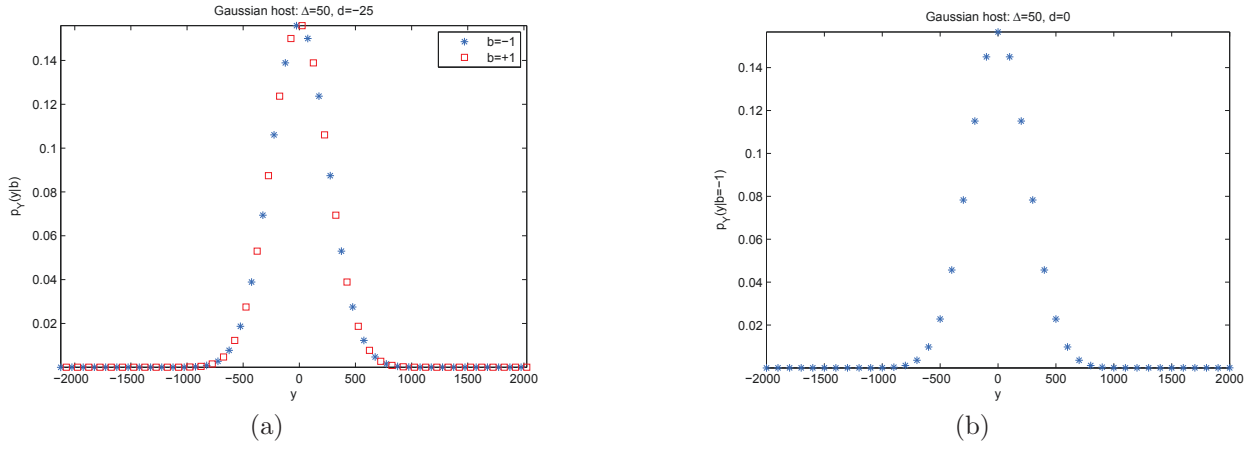

Figure 1 The PDF curves of y for zero-mean Gaussian host data with variance 255^2: (a) $d = -\frac{\Delta}{2}$, (b) $d = 0$.

Moreover, when both x and v are distributed symmetrically around the origin, we have the mirror property $p_Z(z + 2c|b = -1) = p_Z(-z|b = 1)$ for the case $d = -\Delta/2$, and the symmetric property $p_Z(z + 2c|b) = p_Z(-z|b)$ for the case $d = 0$.

Figure 2a depicts qualitatively the PDFs of z for zero-mean Gaussian host data with variance 255^2 and zero-mean Gaussian noise. It can be seen that there is a bell curve present around each discrete value of y due to the existence of Gaussian noise, and the two adjacent ones even overlap for a large noise strength. Meanwhile, the distance between two discrete points of y is scaled by the scaling factor ρ and $p_Z(z)$ is translated by constant value c. The corresponding empirical density curves of z are plotted in Figure 2b. We see that the analytical PDF of z fits well with empirical observations.

3.3 PDF model of the extracted signal

Recalling the definition of r given previously, it is immediate to write

$$p_R(r|b) = \begin{cases} \sum_{j=-\infty}^{\infty} p_Z(z_j - r|b, d), & r \in [-\Delta, \Delta) \\ 0, & \text{else,} \end{cases}$$

where $p_R(r|b)$ is the PDF of r conditioned on the transmitted symbol b, and z_j has the similar definition with y_k. Due to (5), the above equation becomes

$$p_R(r|b) = \begin{cases} \sum_j \sum_k P(y_k|b) p_v(\mu_{jk} - r), & r \in [-\Delta, \Delta) \\ 0, & \text{else} \end{cases} \quad (6)$$

with $\mu_{jk} = z_j - \rho y_k - c$.

Now, let us analyze the properties of $p_R(r)$. If ignoring the effect of d on $P_Y(y)$, in view of (6), we derive $p_R(r - \epsilon d|b, d \neq 0) \approx p_R(r|b, d = 0)$ with $\epsilon = \rho - 1$. This shows that for the case $d \neq 0$ the PDF $p_R(r|b)$ can be approximately obtained by translating $p_R(r|b, d = 0)$. Further, while $|\epsilon|$ is small enough for neglecting the term ϵd, we have the property $p_R(r|b, d \neq 0) \approx p_R(r|b, d = 0)$. That is, despite the choice of d, $p_R(r)$ approximately remains unchanged for small $|\epsilon|$. Similarly to $p_Z(z)$, by assuming the PDFs $p_X(x)$ and $p_v(v)$ are even, we obtain the mirror property $p_R(r - 2c|b = 0) = p_R(-r|b = 1)$ for $d = -\frac{\Delta}{2}$ and the symmetric property $p_R(r - 2c|b) = p_R(-r|b)$ for $d = 0$. At the same time, for any ϵ, we derive $p_R(r|b, \rho = 1 + \epsilon) = p_R(r|b, \rho = 1\text{-}\varepsilon)$ for $d = 0$ and $p_R(r|b = 0, \rho = 1 + \varepsilon) = p_R(r|b = 1, \rho = 1\text{-}\varepsilon)$ for $d = -\frac{\Delta}{2}$, where $p_R(r|b, \rho)$ denotes the conditional PDF of r given the transmitted symbol b and the scaling factor ρ. These properties of $p_R(r)$ are helpful for us to analyze the performance of DM.

Figure 3 plots the probability density curves of r and the corresponding empirical ones for zero-mean Gaussian host data with variance 255^2 and zero-mean Gaussian noise. As can be seen, the distribution curve of r is either dilated or compressed by the scale factor ρ, and the PDF $p_R(r)$ with $c \neq 0$ corresponds to $p_R(r)$ with $c = 0$ translated by the constant value c. The probability that the values of r are distributed around zero decreases as attacks become strong, which results in the increase of BER. Comparison of Figure 3a, b reveals the analytical PDF of r fits perfectly with its empirical distribution.

4 Performance analysis of DM against the composite attacks

As the previous literatures, the decoding bit error probability P_e is used as the final performance measurement.

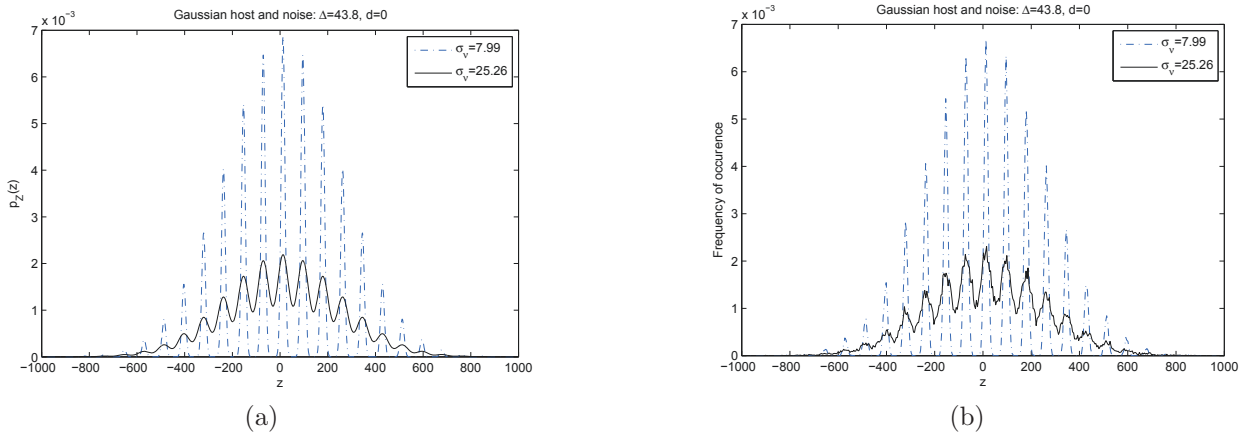

Figure 2 The distribution curves of z in presence of the composite attacks for different values of the noise variance σ_v when given ρ = 0.95 and c = 13: **(a)** the theoretical results, **(b)** the empirical results.

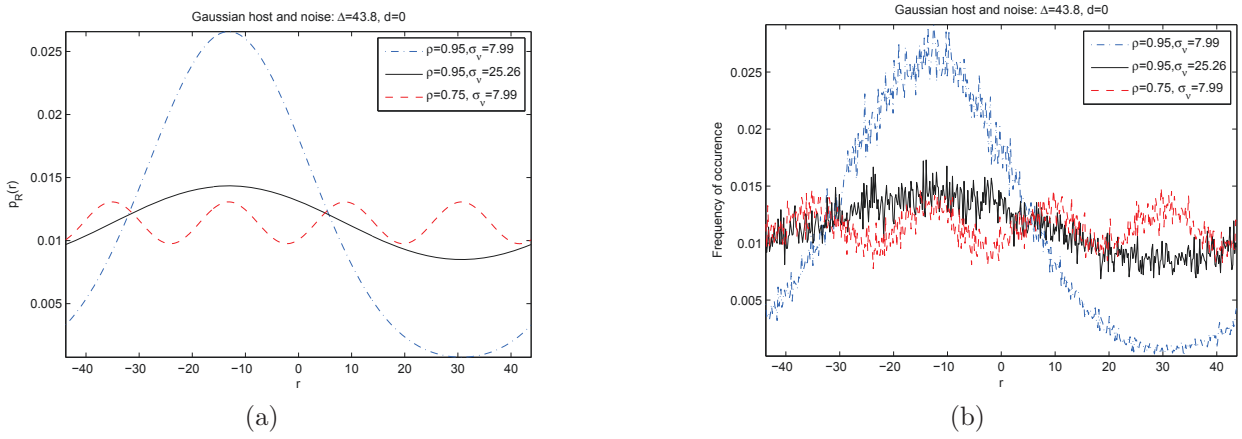

Figure 3 The distribution curves of *r* in presence of the composite attacks for different values of the noise variance σ_v and the scaling factor ρ when given $c = 13$: (a) the theoretical results, **(b)** the empirical results.

Assuming that the symbol b is sent, the bit error probability will be

$$P_e = P(||\vec{\boldsymbol{r}}|| > ||\Delta\vec{\mathbf{a}} - |\vec{\boldsymbol{r}}||\,|b) \tag{7}$$

where $|\boldsymbol{r}|$ denotes the vector of absolute values of components of $\boldsymbol{r}$. Defining $s \triangleq |\boldsymbol{r}|^T\mathbf{a}$, the above expression is equivalent to

$$P_e = \int_{L\Delta/2}^{L\Delta} p_S(s|b)ds. \tag{8}$$

To compute P_e, we need know the PDF $p_S(s)$ of s. The exact solution for $p_S(s)$ may be achieved by several means. One of the standard procedures is by performing multifold integral operation as

$$p_S(s|b) = \int_0^{(L-1)\Delta} \cdots \int_0^{2\Delta}\int_0^{\Delta} p_{|R_1|}(u_1|b)p_{|R_2|}(u_2-u_1|b) \ldots p_{|R_L|}\left(s - \sum_{i=1}^{L-1} u_i|b\right) du_1 du_2 \cdots du_{L-1}, \tag{9}$$

where $p_{|Rj|}(r_j|b) = p_{Rj}(r_j|b)+p_{Rj}(-r_j|b)$ and $p_{Rj}(r_j)$ is the PDF of the jth element of $\boldsymbol{r}$. The computation of $p_S(s)$ is feasible for a small L by (9). However, it becomes impractical as L increases. To solve the problem, it is nature to use mathematically tractable approximations. Let us assume that all components of $\boldsymbol{d}$ are equal, so that the vector $\boldsymbol{r}$ has i.i.d components. At this point, by the well known central limit theorem (CLT), s thus can be approximated by a Gaussian random variable, whose mean and variance are $L\mu_{|r|}$ and $L\sigma^2_{|r|}$. Using the derived PDF in (6), $\mu_{|r|}$ and $\sigma^2_{|r|}$ are represented as

$$\mu_{|r|} = \sum_j \sum_k P(y_k|b) \int_{-\Delta}^{\Delta} |u| p_v(\mu_{jk} - u)du \tag{10}$$

$$\sigma^2_{|r|} = \sum_j \sum_k P(y_k|b) \int_{-\Delta}^{\Delta} u^2 p_v(\mu_{jk} - u)du - \mu^2_{|r|}. \tag{11}$$

Then, the probability P_e is computed as

$$P_e \approx \Phi\left(\frac{\sqrt{L}(\Delta - \mu_{|r|})}{\sigma_{|r|}}\right) - \Phi\left(\frac{\sqrt{L}(\Delta/2 - \mu_{|r|})}{\sigma_{|r|}}\right) \tag{12}$$

where $\Phi(\cdot)$ stands for the cumulative distribution function (CDF) of the standard Gaussian distribution. It should be pointed out the CLT approximation to P_e is only valid for very large L. In reality, the condition is generally met in order to improve the watermarking robustness.

Now, we can observe several useful properties of P_e from the previous analysis. When $|\epsilon|$ is small enough, by the property $p_R(r|b, d \neq 0) \approx p_R(r|b, d = 0)$, it is easily understood that P_e approximately remains unchanged for different dither value. Therefore, without loss of generality, d is set to 0. Furthermore, for $d = 0$, if both $p_X(x)$ and $p_v(v)$ are even, using the property $p_R(r|b, \rho = 1 + \epsilon) = p_R(r|b, \rho = 1 - \epsilon)$, the same value of P_e is obtained for the cases of $\rho = 1 - \epsilon$ and $\rho = 1 + \epsilon$. As a result, the property of P_e also holds for $d \neq 0$ approximately.

4.1 Generalized Gaussian host and noise

Theoretically, P_e can be estimated only if the PDFs $p_X(x)$ and $p_v(v)$ are given. For the following analysis we consider a specific case where the host signal and attacking noise are statistically modeled by the GGD. The GGD

model is used because it includes a family of distributions and suitable for many practical applications. The PDF of the GGD is defined as

$$p(t) = \frac{\kappa\beta}{2\Gamma(\beta^{-1})} e^{-|\kappa(t-\mu)|^{\beta}}, \tag{13}$$

where $\kappa = \frac{1}{\sigma}\sqrt{\Gamma(3\beta^{-1})/\Gamma(\beta^{-1})}$, and $\Gamma(u) = \int_0^{\infty} t^{u-1}e^{-t}dt$ is the Gamma function. Thus, the distribution is completely specified by the mean μ, the standard deviation σ and the shape parameter β, and is denoted as GGD(β; μ, σ). The CDF of the GGD has the form [13]

$$\Psi(t) = \frac{1}{2} + sgn(t-\mu)\frac{\gamma(\beta^{-1}, |\kappa(t-\mu)|^{\beta})}{2\Gamma(\beta^{-1})}$$

where $\gamma(u_1, u_2) = \int_0^{u_2} t^{u_1-1}e^{-t}dt$ is the lower incomplete gamma function, and $sgn(\cdot)$ denotes the sign function, i.e.,

$$sgn(t) = \begin{cases} 1, & t \geq 0 \\ -1, & \text{else.} \end{cases}$$

Note that Gaussian and Laplacian distributions are just two special cases of the GGD with $\beta = 2$ and $\beta = 1$, respectively.

First, the PMF $P_Y(y)$ is calculated according to the distribution model of x. Given $p_X(x) \sim$ GGD(β_x; μ_x, σ_x) and the corresponding CDF $\Psi_x(x)$, in view of (4), we immediately write

$$P_Y(\gamma_k|b) = \Psi_x(\gamma_k + \Delta) - \Psi_x(\gamma_k - \Delta). \tag{14}$$

Then, the integration terms in (10) and (11) are derived for the case $p_v(v) \sim$ GGD(β_v; 0, σ_v). In appendix, we obtain

$$\begin{aligned} \int_{-\Delta}^{\Delta} |u| p_v(t-u)du = & \frac{\gamma(2\beta_v^{-1}, |\kappa(t-\Delta)|^{\beta_v}) - \gamma(2\beta_v^{-1}, |\kappa t|^{\beta_v})}{2\sigma_v^{-1}\sqrt{\Gamma(\beta_v^{-1})\Gamma(3\beta_v^{-1})}} \\ & + \frac{\gamma(2\beta_v^{-1}, |\kappa(t+\Delta)|^{\beta_v}) - \gamma(2\beta_v^{-1}, |\kappa t|^{\beta_v})}{2\sigma_v^{-1}\sqrt{\Gamma(\beta_v^{-1})\Gamma(3\beta_v^{-1})}} \\ & + (2\Psi_v(t) - \Psi_v(t-\Delta) - \Psi_v(t+\Delta))t \end{aligned} \tag{15}$$

and

$$\begin{aligned} \int_{-\Delta}^{\Delta} u^2 p_v(t-u)du = & \frac{\gamma(3\beta_v^{-1}, 0) - \gamma(3\beta_v^{-1}, |\kappa_v(t-\Delta)|^{\beta_v})}{2sgn(t-\Delta)\sigma_v^{-2}\Gamma(3\beta_v^{-1})} \\ & + \frac{\gamma(3\beta_v^{-1}, |\kappa_v(t+\Delta)|^{\beta_v}) - \gamma(3\beta_v^{-1}, 0)}{2sgn(t+\Delta)\sigma_v^{-2}\Gamma(3\beta_v^{-1})} \\ & - \frac{(\gamma(2\beta_v^{-1}, |\kappa(t+\Delta)|^{\beta_v}) - \gamma(2\beta_v^{-1}, |\kappa(t-\Delta)|^{\beta_v}))t}{\sigma_v^{-1}\sqrt{\Gamma(\beta_v^{-1})\Gamma(3\beta_v^{-1})}} \\ & + (\Psi_v(t+\Delta) - \Psi_v(t-\Delta))t^2 \end{aligned} \tag{16}$$

Now, the decoding bit-error probability can be estimated by computing (10), (11), and (12) with the use of (14), (15), and (16). Since the calculation of $p_Y(y)$ is relatively simple in (4), the above analysis can be easily extended for other host distributions. However, the derivation of the integration terms in (10) and (11) might become very complex for the noise v with other PDFs. Thus, they are computed numerically when needed.

4.2 Simulations on artificial signals

In order to verify the obtained theoretical results, we conduct a set of experiments on artificial signals. A set of 64000 random data, independently drawn from the GGD with zero mean and variance 255^2, are used as the host signal. A random message with equiprobable information bit is embedded using DM with $L = 64$, $\Delta = 7.79$, and $d = 0$. The noise signal is also generated according to the zero-mean GGD. We calculated the empirical BER under the composite attacks, and compared them to the theoretical values. The obtained results are summarized in Figure 4 for the case of Gaussian host and noise.

Figure 4a shows the BER of DM as a function of the scaling factor ρ while fixing the constant value c and the noise standard deviation σ_v. As can be seen, DM is definitely very sensitive to the scaling attack. The probability of error is unacceptably high when ρ movies beyond the range [0.98, 1.02]. The existence of noise and constant change causes the increase of BER further. And the effect of constant change becomes relatively distinct for strong noise. The theoretical approximation of BER agrees almost perfectly with the empirical results, particularly in the case of weak attacks. Figure 4a also demonstrates that the BER versus ρ curve is symmetric with respect to the point $\rho = 1$. Figure 4b depicts the plots of BER versus the constant value c while fixing the scaling factor ρ and the noise standard deviation σ_v. For small ρ and σ_v, the BER of DM starts to grow rapidly as long as the absolute value of c approaches to $\Delta/2$. The effect of c on performance decreases as ρ and σ_v increase. The estimated BERs are approximately equal to the empirical ones, but the estimation accuracy gets worse for a large c. At the same time, Figure 4b shows that the BER versus c curve is symmetrical around $c = 0$. Figure 4c plots the BER of DM as a function of the noise standard deviation σ_v while fixing the scaling factor ρ and the constant value c. Obviously, the BER increases as σ_v becomes large. The curve of BER versus σ_v seems to be translated due to the effect of valumetric scaling and constant change distortions. Similar to the previous tests, the theoretical BERs fit the empirical ones very well and the maximal difference between them is lower than 0.02.

In the sequel, the sensitivity of DM to statistical properties of the host and noise is investigated. We tested the performance of DM against valumetric scaling

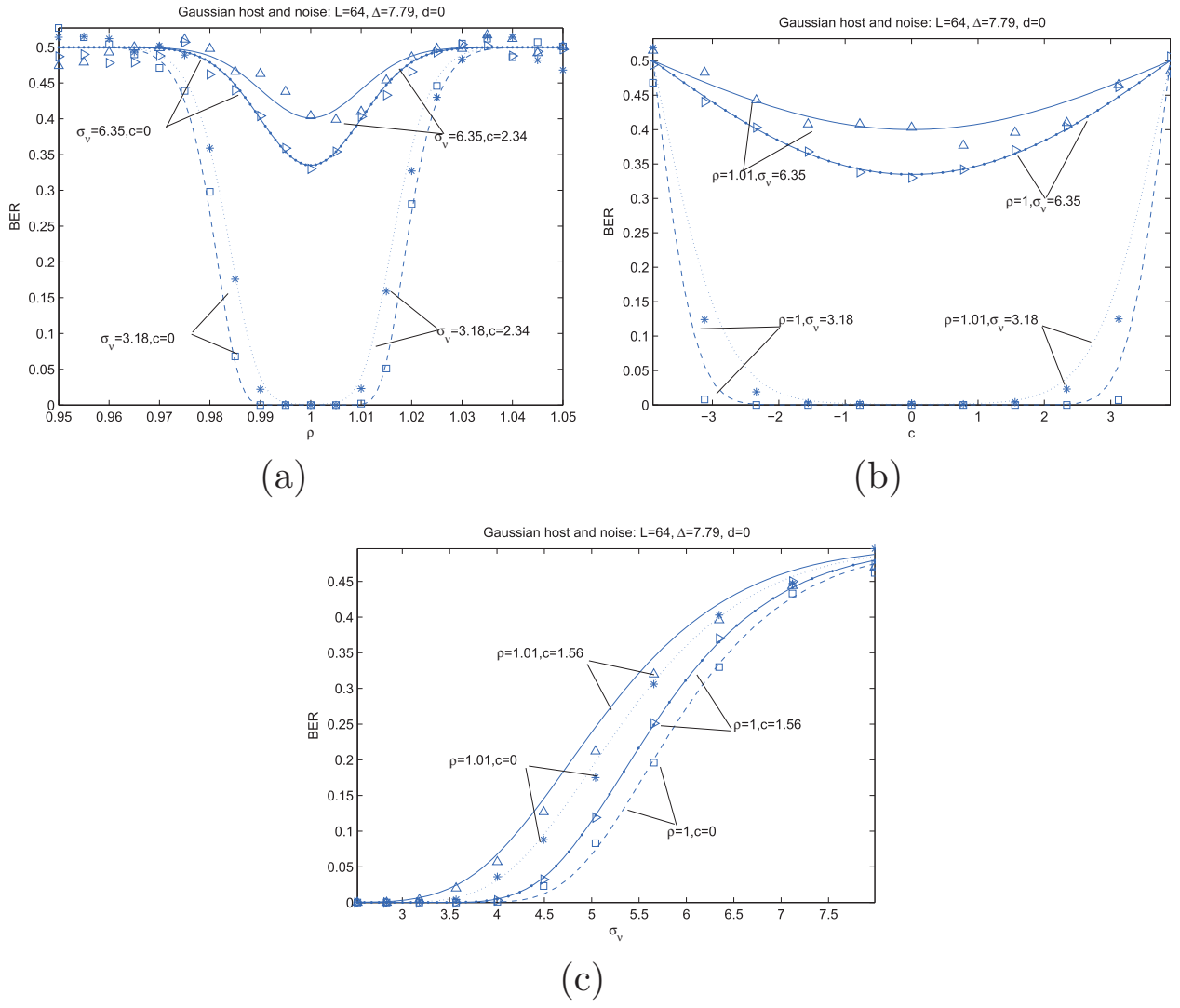

Figure 4 The performance of DM in the presence of the composite attacks, $z = \rho y + v + ca$: (a) BER versus ρ, while fixing c and σ_v, **(b)** BER versus c, while fixing ρ and σ_v, **(c)** BER versus σ_v, while fixing ρ and c. Lines and symbols stand for theoretical values and empirical data, respectively.

attacks for different host PDF shapes controlled by β_x. The results are displayed in Figure 5a. It is remarkable that the performance of DM increases significantly as β_x goes down. For a small β_x, the BER plot becomes relatively flat and the BER grows slowly. This behavior can be explained as follows. For the GGD, the smaller β_x is, the more impulsive the shape, and the heavier the tails. As a result, the probability that a big value of x presents over the range of interests decreases. Thus, the introduced distortion $(\rho - 1)\boldsymbol{y}$ by the scaling attack degrades for the same value of ρ, resulting in the decrease of BER. We also observe that the theoretical approximation agrees almost perfectly with the empirical results for the cases $\beta_x = 2, 8$, but does worse for $\beta_x = 1$. This is because the CLT approximation to BER may underestimate the importance of the tails of $p_X(x)$ with $\beta_x = 1$ and gives the smaller results than the true BERs [5]. However, in terms of constant change and additive noise, the performance of DM is insensitive to the shape parameter β_x, due to the fact that the two operations are independent from the watermarked signal. Hence, we just provide the results for the scaling attack herein. Then, we tested the performance of DM against additive noise with different PDF shapes controlled by β_v. The results are exihibited in Figure 5b. We observe that the BER of DM goes down as β_v increases for the same noise variance. Applying the same reasoning above here, we may understand that relatively serious distortions are introduced by the noise attack with a large β_v, and thus, the performance of DM worsens.

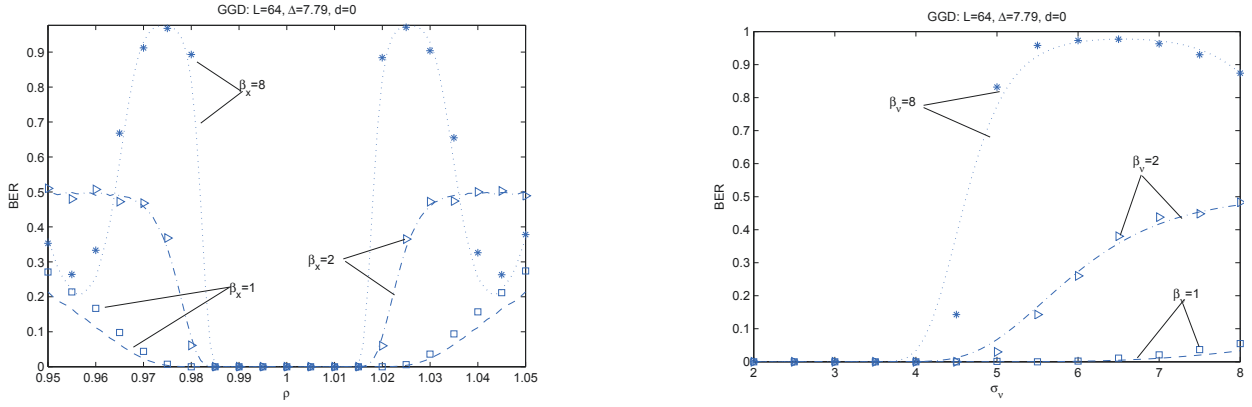

Figure 5 The sensitivity of DM to the shapes of the host and noise distributions: the performance of DM **(a)** against valumetric scaling for the GGD host signals and **(b)** against additive GGD noise. Lines and symbols stand for theoretical values and empirical data, respectively.

5 Normalized DM and its performance

The robustness improvement for DM is taken into account in this section. A novel normalized DM (NDM), is presented, which is theoretically invariant to valumetric scaling and constant change. On the other hand, the performance of NDM is theoretically evaluated in terms of null distortions and noise addition.

5.1 Normalized DM

The main idea of NDM is to construct a gain-invariant vector with zero mean for quantization. There are many strategies for the construction of such a vector. In the study, the vector is achieved in the way that the host vector subtracts its nonzero sample mean and then is divided by its sample standard deviation. The method is described in details as follows.

Let $\bar{u} \triangleq \boldsymbol{u}^T\mathbf{a}/L$ and $S_u^2 \triangleq ||\boldsymbol{u} - \bar{u}||^2/L$ denote the sample mean and variance of a L-dimensional vector $\boldsymbol{u}$, respectively. Watermark embedding is performed by

$$\boldsymbol{y}_j = \lambda_j S_{x_j} Q_{b_j}\left(\frac{\boldsymbol{x}_j - \bar{x}_j\mathbf{a}}{S_{x_j}}\right) + \eta_j\mathbf{a} \tag{17}$$

for $j = 1, \ldots, NR_m$, where the factors λ_j and η_j are determined by two specific distortion situations. For convenience, we define the normalized host vector as $\boldsymbol{x}'_j = (\boldsymbol{x}_j - \bar{x}_j\mathbf{a})/S_{x_j}$ and the error vector as $\boldsymbol{q}_{e_j} = Q_{b_j}(\boldsymbol{x}'_j) - \boldsymbol{x}'_j$. By (17), the sample variance of $\boldsymbol{y}_j$ satisfies $S_{y_j}^2 = \lambda_j^2 S_{x_j}^2(1 + S_{q_{e_j}}^2 + 2\boldsymbol{q}_{e_j}^T\boldsymbol{x}'_j/L)$. An appropriate strategy to choose λ_j is to let $S_{y_j}^2 = S_{x_j}^2$. Therefore, we have

$$\lambda_j = \left(1 + S_{q_{e_j}}^2 + 2\boldsymbol{q}_{e_j}^T\boldsymbol{x}'_j/L\right)^{-\frac{1}{2}}. \tag{18}$$

Then, η_j is obtained through minimizing the distance $||\boldsymbol{y}_j - \boldsymbol{x}_j||$. This leads to

$$\eta_j = \bar{x}_j - \lambda_j S_{x_j}\bar{q}_{e_j}. \tag{19}$$

At detection time, the received signal $\boldsymbol{z}$ is first normalized as done at the embedder's side and then the minimal distance decoder is applied. The modified detector is represented as

$$\widehat{b}_j = \arg\min_{-1,1}\left\|\frac{\boldsymbol{z}_j - \bar{z}_j\mathbf{a}}{S_{z_j}} - Q_{b_j}\left(\frac{\boldsymbol{z}_j - \bar{z}_j\mathbf{a}}{S_{z_j}}\right)\right\|. \tag{20}$$

Now, it is possible to simultaneously see why NDM is insensitive to valumetric scaling and constant change attacks: Substituting $\boldsymbol{z}_j = \mathrm{p}_j\boldsymbol{y}_j + c_j\mathbf{a}$ into (20), it can be readily verified that ρ_j and c_j cancel out in the expression, and consequently, the decision $\widehat{b}_j$ does not depend on ρ_j and c_j.

5.2 Performance analysis

Having known that NDM overcomes the main weakness of the conventional DM, we will evaluate the performance of NDM in terms of null distortions and noise addition. Performing the normalization operation on both sides of (17) and applying (18) and (19), we get

$$\frac{\boldsymbol{y}_j - \bar{y}_j\mathbf{a}}{S_{y_j}} = \lambda_j\left(Q_{b_j}\left(\frac{\boldsymbol{x}_j - \bar{x}_j\mathbf{a}}{S_{x_j}}\right) - \bar{q}_{e_j}\mathbf{a}\right). \tag{21}$$

The above equation indicates that NDM introduces two extra operations in the absence of channel noise: valumetric scaling with λ_j and constant change with $\bar{q}_{e_j}$. In other words, NDM can be regarded as DM undergoing valumetric scaling and constant change distortions. Thus, the theoretical performance of NDM for

null distortions is approximately determined by (10), (11), and (12) as the noise standard deviation σ_v approaches zeros.

To evaluate the effect of λ_j and $\bar{q}_{e_j}$ in (21) on the performance of NDM, we introduce the document-to-watermark ratio (DWR), defined as $\zeta_j \triangleq LS_{x_j}^2/||\boldsymbol{y}_j - \boldsymbol{x}_j||^2$ for the jth subvector. Combining (17), (18) and (19), λ_j can be rewritten as

$$\lambda_j = (1 - \frac{1}{2\zeta_j})/(1 + \frac{\boldsymbol{q}_{e_j}^T \boldsymbol{x}_j'}{L}). \tag{22}$$

For small Δ, it has been shown [14] that each element of the error vector $\boldsymbol{q}_e$ obeys independently a uniform distribution over the interval [-Δ, Δ) and $\boldsymbol{q}_e$ is statistically independent from $\boldsymbol{x}_j'$. Applying the properties, it is easy to derive that $\boldsymbol{q}_{e_j}^T \boldsymbol{x}_j'/L$ in (22) has zero mean and variance $\Delta^2/(3L)$. Thus, λ_j tends to 1 - $0.5/\zeta_j$ as $L \to \infty$ or $\Delta \to 0$ (i.e., $\zeta_j \to \infty$). Figure 6 plots the curves of the true average error $|\lambda_j - 1|$ versus ζ_j for different L, as well as the limit $0.5/\zeta_j$. Notably, the gap between the factor λ_j and its limit becomes smaller and smaller as L and DWR increase. Over the most interesting range of ζ_j from 25 dB to 50 dB, the value of $|\lambda_j - 1|$ is less than 0.01 for all the values of L tested. From Figure 4a, it is observed that the valumetric scaling with $|\lambda_j - 1| < 0.01$ affects the performance of NDM so less that no decoding error is made.

As to the constant change with $\mathrm{E}\{\boldsymbol{q}_{e_j}\}$, we have a sufficient condition that $\left|\mathrm{E}\{\boldsymbol{q}_{e_j}\}\right| < \Delta/2$ for making no error. Considering the statistical properties of $\boldsymbol{q}_{e_j}$, it is possible to resort to the CLT to show that for large L, the sample mean $\bar{q}_{e_j}$ can be accurately modeled by a Gaussian PDF with zero mean and variance $\Delta^2/(3L)$. Thus, the probability that $|\bar{q}_{e_j}| < \Delta/2$ holds can be computed as

$$\mathrm{P}(|\bar{q}_{e_j}| < \Delta/2) \approx 1 - 2\Phi(-\sqrt{3L/4}). \tag{23}$$

Since the probability in (23) approaches the value of 1 as L increases, NDM can present a zero probability of error as the original DM for large L in the absence of channel noise. Figure 7 shows the plots of the BER as a function of L. As shown in Figure 7, the probability of error sharply decreases to 0 as L increases. And the agreement of theoretical results with simulations is excellent.

Next, we will analyze the performance of NDM in noise channel. The received signal $\boldsymbol{z}_j$ has the form $\boldsymbol{z}_j$ =

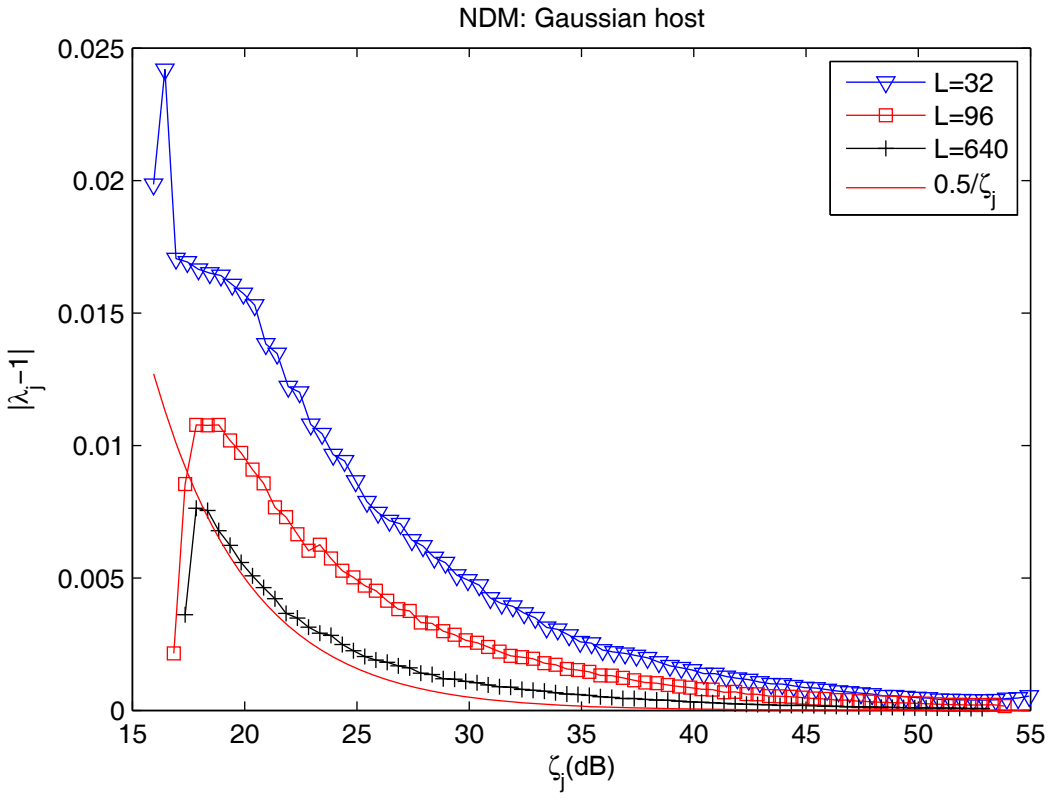

Figure 6 The average empirical values of $|\lambda_j - 1|$ and the corresponding limit as a function of DWR.

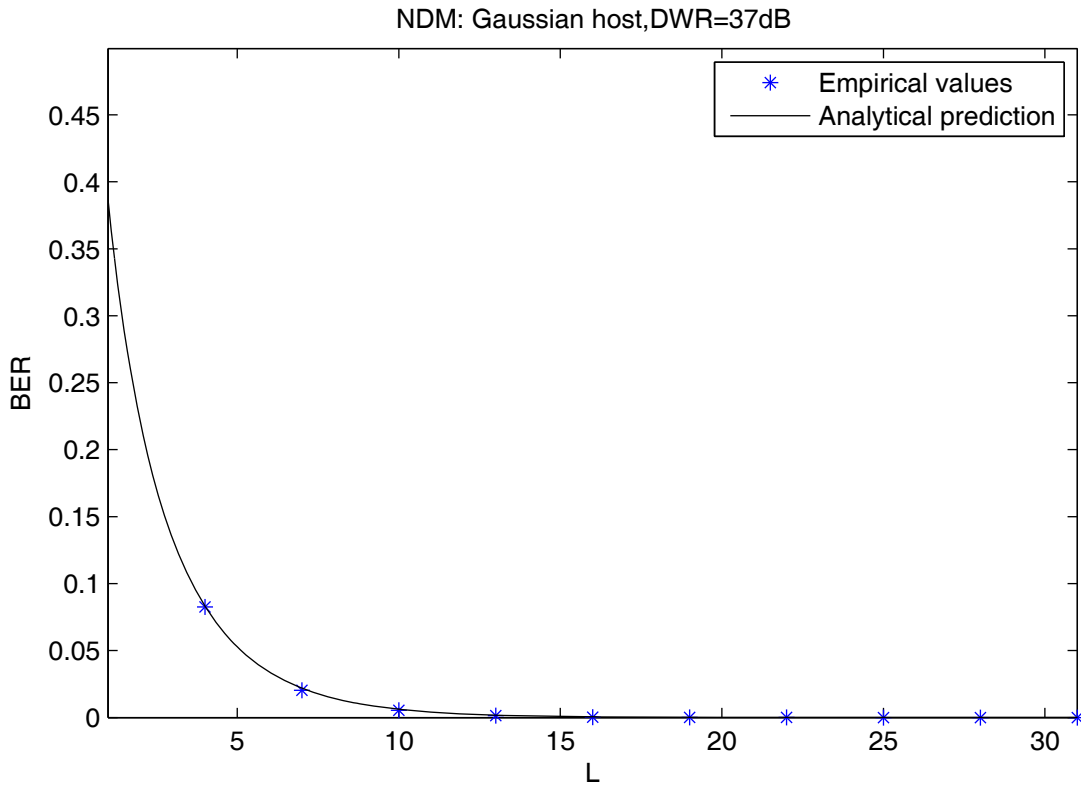

Figure 7 BER versus the number of host samples to conceal one bit information.

$\boldsymbol{y}_j + \boldsymbol{v}_j$, where $\boldsymbol{v}_j$ is an unknown noise source with zero sample mean and $\boldsymbol{v}_j$ and $\boldsymbol{y}_j$ are orthonormal. Since NDM is invariant to valumetric scaling and constant change attacks, it's sufficient to consider the case. To measure the impact of the noise, we will follow the popular watermark-to-noise ratio (WNR), defined as $\xi_j \triangleq ||\boldsymbol{y}_j - \boldsymbol{x}_j||^2/||\boldsymbol{v}j||^2$ for the jth subvector. Applying the normalization operation to $\boldsymbol{z}_j$ yields

$$\frac{z_j - \bar{z}_j\mathbf{a}}{S_{z_j}} = \lambda'_j Q_{b_j}\left(\frac{x_j - \bar{x}_j\mathbf{a}}{S_{x_j}}\right) + \nu'_j + \bar{q}'_{\mathrm{e_j}}\mathbf{a} \tag{24}$$

with $\lambda'_j = \lambda_j\sqrt{\frac{\zeta_j\xi_j}{\zeta_j\xi_j+1}}$, $\boldsymbol{v}'_j = \sqrt{\frac{\zeta_j\xi_j}{\zeta_j\xi_j+1}}\frac{\boldsymbol{v}_j}{s_{x_j}}$, and $\bar{q}'_{e_j} = -\lambda_j\bar{q}_{e_j}$. Expression (24) illustrates that NDM undergoes the composite attacks as considered in (3). Therefore, the previously obtained theoretical results can be used to predict the performance of NDM.

Generally speaking, the factor λ'_j in (24) is approximately equal to the value of 1 by the fact that $\zeta_j\xi_j \gg 1$ holds in practical applications. Figure 8 shows that the value of $|\lambda'_j - 1|$ is rather less even for serious distortions (e.g., WNR = -10 dB). On the other hand, for large L, the effect of $\bar{q}'_{e_j}$ in (24) can be neglected. Based on these two considerations, the increase of BER is mainly derived from the term $\boldsymbol{v}'_j$ in (24). As a result, we can draw the conclusion that NDM almost resists the same amount of noise as the original DM. Figure 9 illustrates the performance difference between NDM and DM under the additive noise attacks. As can be seen, NDM performs slightly worst than DM when the WNR is within the range [-1 dB, 3 dB], but outperforms it once WNR is lower than -1 dB. In principal, their performance is very close in this regard. Under the light of the above analysis we conclude that NDM achieves the performance approximately equal to DM, still keeping invariance against valumetric scaling and constant change attacks.

6 The improvement of NDM

The previous analysis shows that when $\lambda_j \neq 1$ and $\bar{q}_{e_j} \neq 0$ the two factors have the negative impact on the performance of NDM. Thus, the influence of them should be decreased or eliminated so as to obtain the improved performance. Based on this idea, we present the improved NDM (IM-NDM) in the sequel.

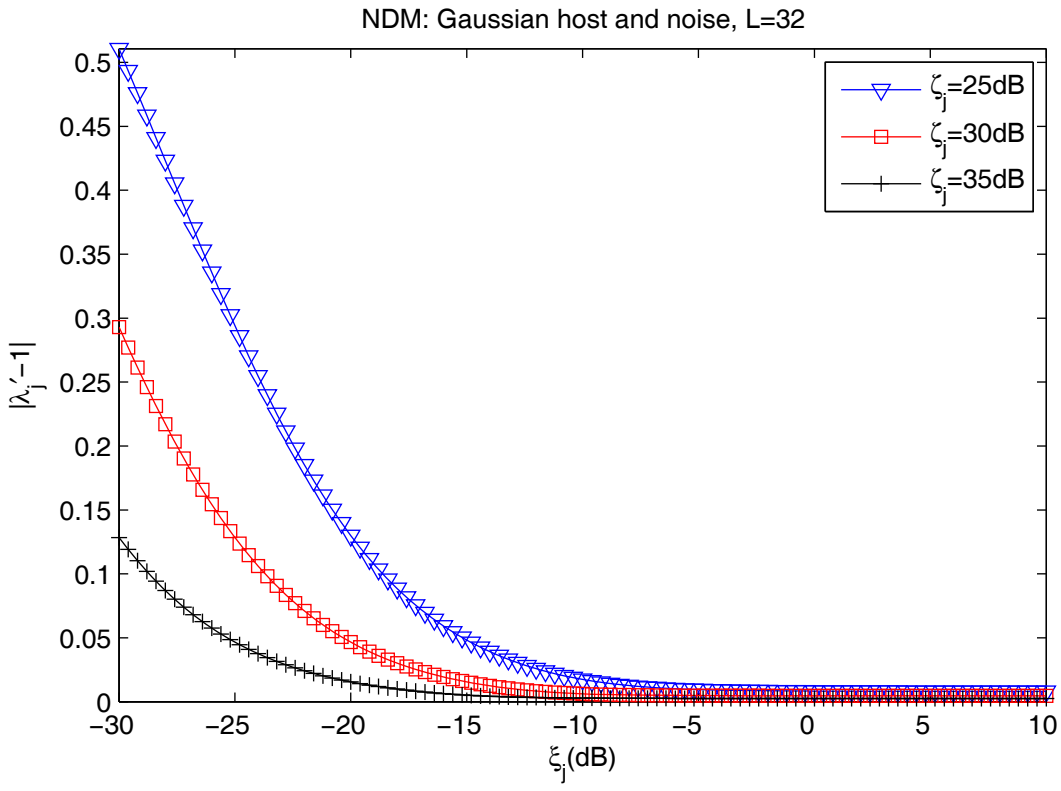

Figure 8 The average empirical values of $|\lambda_j' - 1|$ as a function of WNR for different values of DWR.

In IM-NDM, the watermarked vector is generated by weighting the quantization error signal and adding it back to the host signal. The modified embedder is expressed as

$$\boldsymbol{y}_j = \lambda_j S_{\boldsymbol{x}_j}(\boldsymbol{x}_j' + \boldsymbol{\alpha}_j \cdot \boldsymbol{q}_{e_j}) + \eta_j \mathbf{a}, \tag{25}$$

where $\boldsymbol{\alpha}_j$ denotes the weight vector whose element is between 0 and 1, and $\boldsymbol{\alpha}_j \cdot \boldsymbol{q}_{e_j}$ indicates that each dimension of $\boldsymbol{\alpha}_j$ is multiplied by the corresponding dimension of $\boldsymbol{q}_{e_j}$. Similarly to (18) and (19), it is derived that

$$\lambda_j = \left(1 + S^2_{q_{e_j}} + 2\boldsymbol{q}_{e_j}^T \boldsymbol{x}_j'/L\right)^{-\frac{1}{2}} \tag{26}$$

and

$$\eta_j = \bar{x}_j - \lambda_j S_{x_j} \bar{q}_{e_j}. \tag{27}$$

Note that NDM is a special case of IM-NDM with $\boldsymbol{\alpha}_j = \mathbf{a}$. The weight vector $\boldsymbol{\alpha}_j$ plays an important role in the performance of IM-NDM. Through a careful choice of $\boldsymbol{\alpha}_j$, the influence of both λ_j and $\bar{q}_{e_j}$ in (25) can be decreased (or even eliminated), and at the same time the distortion-compensation (DC) mechanism is introduced. The latter is proved to be an effective way to improve the performance of quantization-based watermarking [1].

By letting $\lambda_j = 1$ and $\eta_j = \bar{x}_j$, we have

$$\begin{aligned} (\boldsymbol{\alpha}_j \cdot \boldsymbol{q}_{e_j})^T (\boldsymbol{\alpha}_j \cdot \boldsymbol{q}_{e_j} + 2\boldsymbol{x}'_j) &= 0 \\ (\boldsymbol{\alpha}_j \cdot \boldsymbol{q}_{e_j})^T \mathbf{a} &= 0. \end{aligned} \tag{28}$$

Taking use of one solution of (28) in (25) allows us to eliminate the negative impact of λ_j and η_j. Obviously, it is easy to obtain one solution of (28) for one of the two equations in (28) is linear. If (28) has multiple solutions, an appreciate one should be chosen by the performance of IM-NDM. Obtaining the appropriate solution for $\boldsymbol{\alpha}_j$ and investigation of its effect on watermarking performance is beyond the scope of this article and is a good direction for future research. If (28) has none solution, $\boldsymbol{\alpha}_j$ should be chosen to minimize $|\lambda_j - 1|$ under the situation $\eta_j = \bar{x}_j$. This is a constraint optimization problem and can be solved using the Lagrangian multiplier method.

Figure 9 illustrates the performance of IM-NDM described above under the additive noise attack, together with DC-DM, and the distortion compensated NDM (DC-NDM), namely IM-NDM taking the same weight for each element of $\boldsymbol{q}_{e_j}$, where the DC value is set

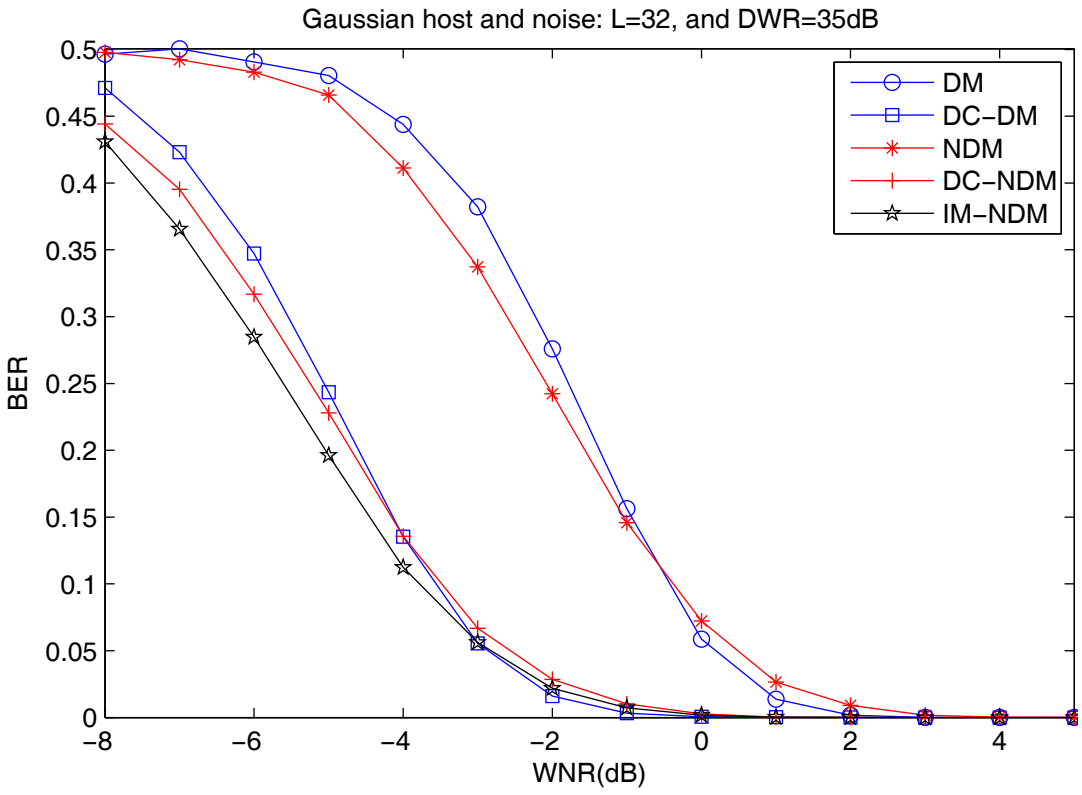

Figure 9 BER as a function of AWGN for Gaussian host.

to 0.66 for the latter two schemes. Obviously, DC-NDM almost presents the same robustness as DC-DM against weak attacks, and performs a little better facing very serious distortions (WNR <-4 dB). And they are noticeably outperformed by IM-NDM.

7 Experimental results

In this section, a series of experiments are conducted on real images to evaluate the validity of analytical derivations and performance of the proposed method.

7.1 Theoretical verification

In the experiments, we use three standard images, shown in Figure 10. The DM method is implemented in the spatial domain so as to observe its performance without the impact of transform operations. Specifically, all pixels of one image are rearranged in a vector as the host signal. A random binary message is embedded into the host vector by DM when given the quantization step Δ, the dither value d and the number of dimensions L. The watermarking algorithm is tested under the composite attacks of valumetric scaling with the factor ρ, constant change with the value c and additive noise v following the distribution GGD(β_v; 0, σ_v). The distribution parameters of image pixels used for the computation of theoretical BERs are displayed in Table 1, which are obtained by the maximum likelihood estimator [15]. The experimental results are summarized in Figure 11 for L = 32, Δ = 8, and d = 0.

Figure 10 Three standard test images: Crowd (right), Mandrill (middle), and Lena (left).

Table 1 Mean, variance, and shape parameter of image pixels

Image	μ_x	σ_x	β_x
Crowd	85.2	50.9	1.5
Mandrill	129.1	42.4	3.3
Lena	99.1	47.9	10.6

Figure 11a depicts the plots of BER as a function of the scaling factor ρ for each image. On the Crowd image, which has the smallest shape parameter among the tested images, DM achieves the best performance. This behavior is consistent with the results in Figure 5a. The shape parameter of the Lena image is larger than one of Mandrill, but better performance is achieved on the Lena image. This can be explained as follows. The valumetric scaling operation introduces the serious distortions on the Mandrill image with a distinctively large mean luminance. As a result, not only the performance gain caused by the host PDF shape cancels out but also the BER grows up. The analytical curves closely fit the empirical data for the Lena and Mandrill images. By contrast, the prediction accuracy becomes slightly worse for the Crowd image. That is mainly due to the fact that the GGD is a poor model for this image. Figure 11b illustrates the sensitivity of DM to the addition/subtraction of a constant luminance value while fixing ρ and σ_v. In the test, DM performs closely for all the test images. That is, the performance of DM with respect to constant change attack is insensitive to the statistical properties of host signal. It is remarkable that the empirical performance of DM is predicted by the theoretical results with a high degree of accuracy. The plots of BER versus the standard deviation σ_v are shown in Figure 11c for $\beta_v = 1$ and Figure 11d for $\beta_v = 8$ while fixing ρ and c. As to the attack, the obvious performance difference is observed between different images. The effect is actually caused by the valumetric scaling operation, and thus can be removed by setting $\rho = 1$. Comparing Figure 11c with Figure 11d, it becomes clear

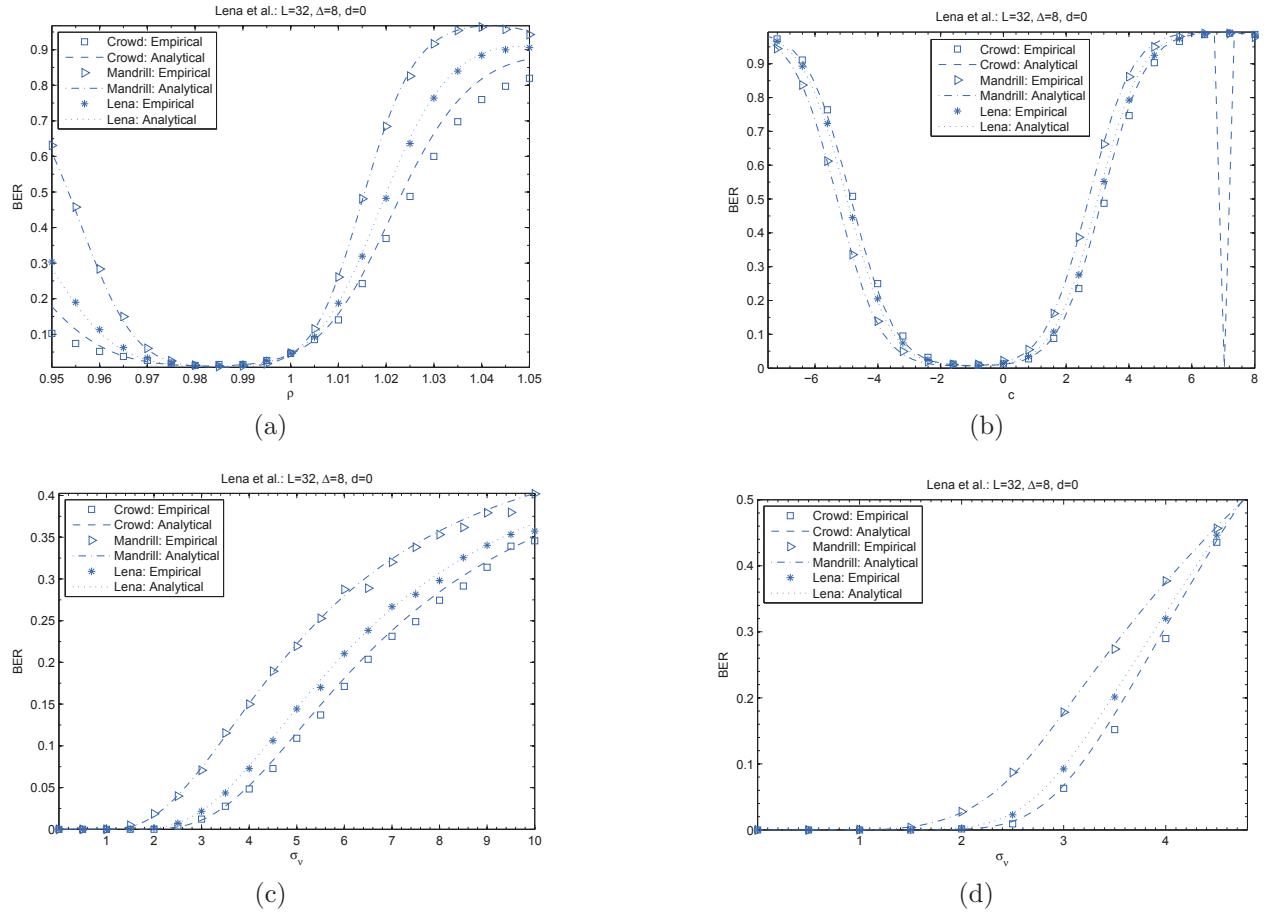

Figure 11 The performance of DM in the presence of the composite attack $z = \rho y + v + ca$: **(a)** BER versus ρ, while $c = \frac{\Delta}{8}$, $\sigma_v = \frac{\Delta}{4}$, and $\beta_v = 2$, **(b)** BER versus c, while $\rho = 1.01$, $\sigma_v = \frac{\Delta}{4}$, and $\beta_v = 2$, **(c)** BER versus σ_v , while $\rho = 1.01$, $c = \frac{\Delta}{8}$, and $\beta_v = 1$, **(d)** BER versus σ_v , while $\rho = 1.01$, $c = \frac{\Delta}{8}$ and $\beta_v = 8$.

that the additive noise with a flat PDF is a worst-case attack for DM. This agrees with the observation in Figure 5b. In the two cases, the predictions are desirable, but there are small discrepancies at some points.

7.2 Performance evaluation

We tested the performance of the proposed NDM in terms of imperceptibility and robustness and compared it with DM, DC-DM, Oostveen's method [10] and RDM [9]. Experiments were carried out on a database of 4000 images from the Corel database, each of dimension 256 × 384. The watermark embedding was performed in the spatial domain in order to see the sensitivity of the tested schemes to constant intensity change. Specifically, we divided the target image into nonoverlapping blocks of size 8×8 and extracted a total of 225 blocks with the highest local variance. Each of the extracted blocks was modulated with two random message bits, so a total of 450 bits can be embedded into one image. The DC value was set to 0.66 for DC-DM. The L_2 vector norm of 50th-order was used as the division function in RDM.

In the experiment on watermarking imperceptibility, the watermark energy induced by all the tested schemes is kept the same and in this case the watermarked images' quality is assessed with several objective image quality metrics. The weighted peak signal-to-noise ratio (wPSNR) and the total perceptual error (TPE) are used to measure the global image quality, as well as the number of blocks greater than the first local perceptual error threshold (NLPE1) and the second local perceptual error threshold (NLPE2) to measure the local image quality. The parameters for them take the default values as suggested in Checkmark [16]. Table 2 reports the experimental results averaged over all the test images when the DWR is fixed at 21 dB.

As shown in Table 2, among all the tested watermarking schemes, NDM and its improved version offer the highest wPSNR values (in dB), the smallest TPE, NLPE1, and NLPE2 values for the same watermark energy. They all indicate that the performance of NDM, in terms of imperceptibility, is better than that of other ones. This is because the adaptive quantization step size is chosen to be proportional to the local variance of the host image in NDM (see (17)). The image quality produced by IM-NDM degrades when compared with NDM. The situation also presents between DM and DC-DM. This is attributed to the fact that a large quantization step is used for watermark embedding with distortion compensation. Surprisingly, RDM manifests the worst performance in this regard.

Table 2 Watermark imperceptibility assessment in use of several image quality metrics

Metrics	DM	DC-DM	Oostveen's	RDM	NDM	IM-NDM
wPSNR (dB)	42.63	42.62	42.91	42.58	42.98	42.97
TPE	0.035	0.035	0.032	0.041	0.032	0.032
NLPE1	8.50	8.65	7.50	9.87	6.70	6.83
NLPE2	4.63	4.55	4.30	6.53	3.58	3.60

In what follows, the watermark robustness will be evaluated with respect to some typical image processing operations. The watermarked images were produced by the tested schemes when fixing DWR at 21 dB. All the given BERs are averaged over the test set of images, except otherwise indicated.

Figure 12 shows the robustness to amplitude scaling for all schemes. Clearly, except the conventional DM and DC-DM, the others manifest strong robustness against this attack. Particularly, the lowest values of BER are achieved by IM-NDM over the whole range of scaling factor ρ tested. However, when ρ exceeds 1.2, the robustness of IM-NDM goes down slightly. That can be attributed to the increasing rounding and clipping distortions.

Figure 13 illustrates the sensitivity of all schemes to the addition/subtraction of a constant luminance value c. It can be seen that both DM and DC-DM are very fragile to constant change attack. The BER of them sharply increases to 1 when c gets close to 10 or -10. Although Oostveen's method and RDM perform better than the original QIM schemes, they are still sensitive to this kind of attack. Our methods are evidently more robust in this regard than other ones. They are almost invariant to constant change and approximately keep the BER of 0 over the range of c tested.

The robustness to AWGN is shown in Figure 14 for each watermarking scheme. In this regard, NDM clearly outperforms Oosteen' method and RDM. Comparing with DM, NDM achieves higher BER for weak noise. This can be explained by the fact that the introduced noise causes the errors in the estimation of the quantization step size for NDM. However, as the noise becomes strong, the BER of DM grows rapidly and is finally lower than one of NDM. The situation is in accordance with the analytical results in Section 5.2. Note that IM-NDM behaves like NDM but presents the improved performance.

The robustness of NDM against AWGN was also tested on the Lena image to verify the analytical derivations for NDM. Since the performance of NDM depends on the local variance of the host image, the empirical BER can not be accurately predicted by exploiting the information from a certain image block. Thus, for the computation of the theoretical BER, we chose three image blocks with different variance: the middle one is around the average variance over those image blocks for watermark embedding and other two ones are respectively a little larger and smaller than it. The theoretical

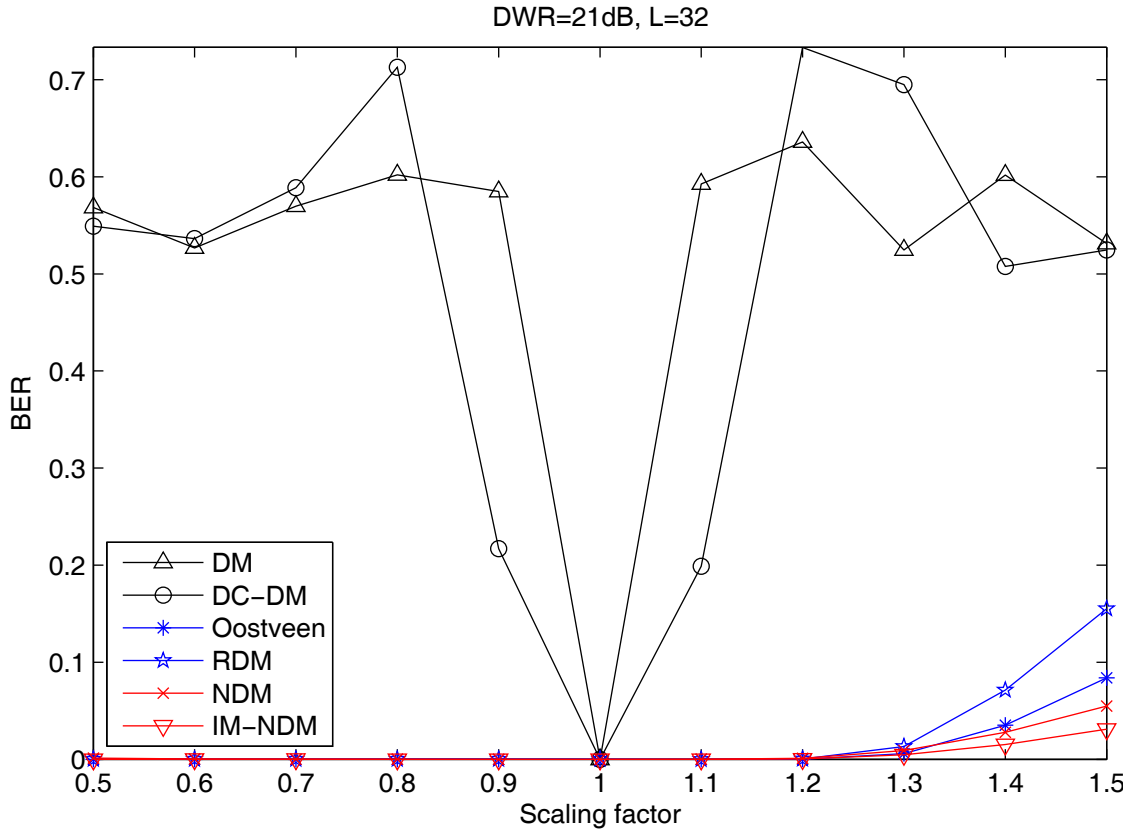

Figure 12 BER as a function of amplitude scaling.

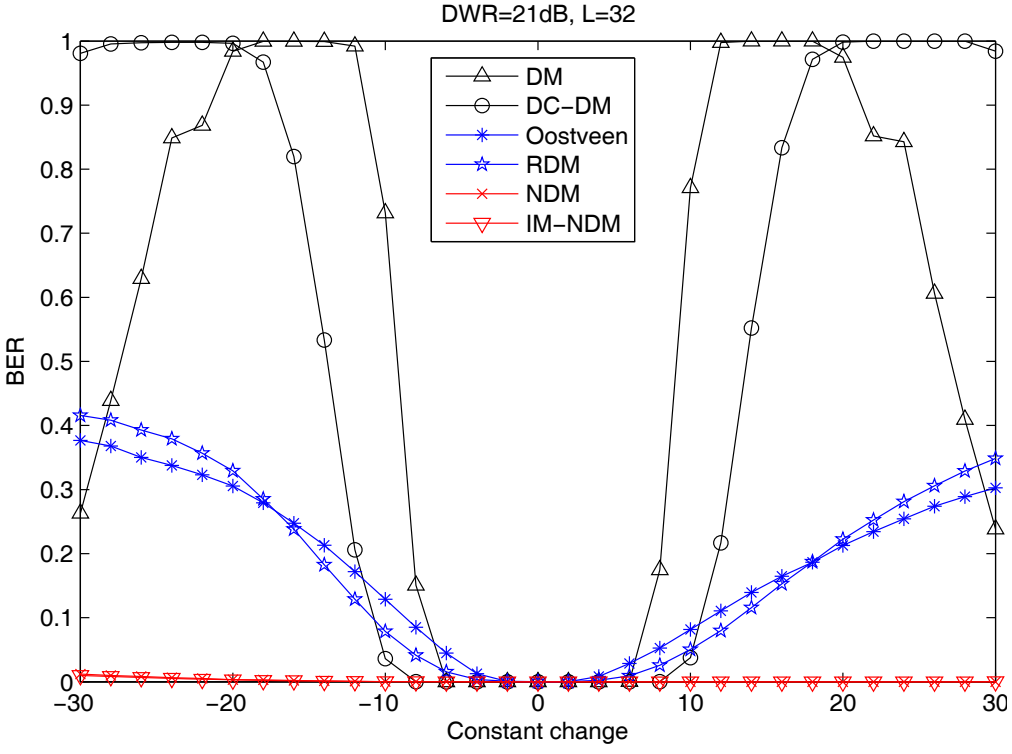

Figure 13 BER as a function of constant luminance change.

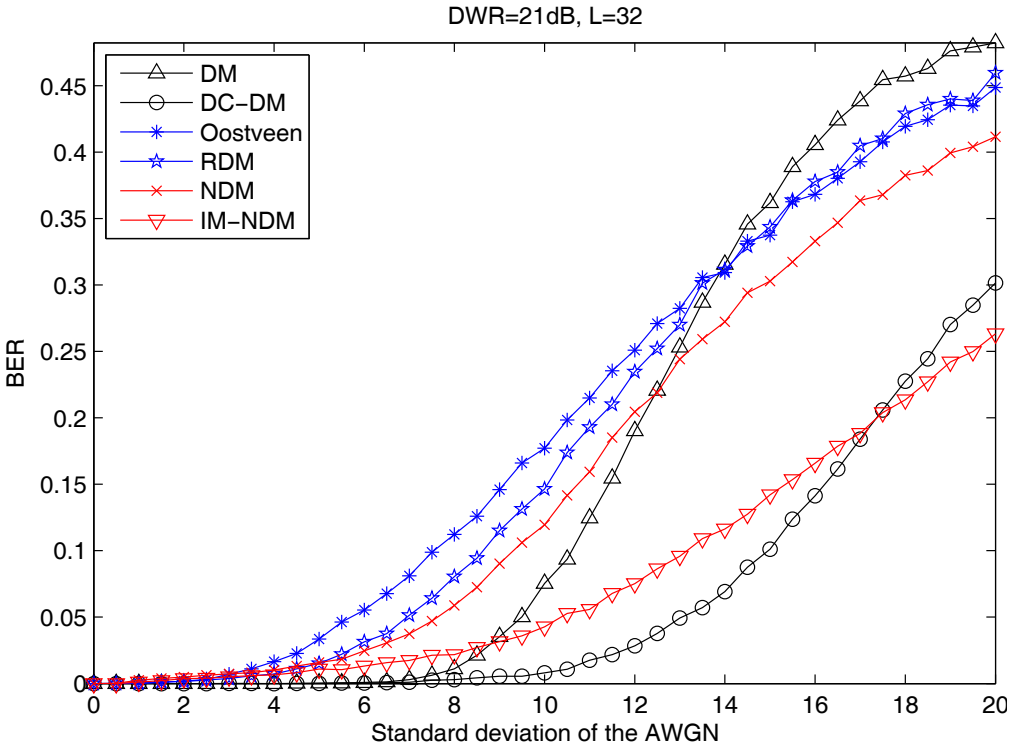

Figure 14 BER as a function of AWGN.

and empirical results are depicted in Figure 15. As can be seen, the upper analytical curve relatively fits well to empirical observations in the weak noise case, and other two curves respectively do well for the moderate noise and strong noise cases respectively. In principle, the theoretical results are effective for real image.

The sensitivity to JPEG compression is investigated in Figure 16. In this test, NDM performs a little worse than DM. IM-NDM improves the robustness of NDM, but still falls behind DC-DM. It is worth seeing that RDM has superior performance with respect to JPEG compression. That can be explained by the nature of JPEG compression. Unlike the AWGN, JPEG compression is an image-dependent processing operation. The goal of it is to reduce an image file size without noticeable image quality degradation. Thus, the perceptually unrelevant data are removed from an image after compression. The test results of image quality reveals that RDM modifies the image data to be easily noticed more largely than other ones, so that it is impaired less by compression. The situation is opposite for NDM. If the perceptual quality is set to be same for all the tested schemes, it is reasonable to believe that NDM will manifest better performance.

NDM is just a basic watermarking algorithm like DM. The above tests allow us to evaluate its performance baseline and the implementation is coarse. If one wants to design a NDM based watermarking scheme for practical applications, some effective technologies on performance improvement should be carried out, such as the choice of transform domain, the use of error-correction coding, etc. Several image-adaptive DM algorithms are presented by exploiting the characteristics of the human visual system in [11]. The same ideas can be straightforwardly applied to improve the performance of NDM. Recently, a new Logarithmic QIM is developed by introducing the μ-Law concept in [4]. NDM can also attempt to use the concept for the improvement of performance.

8 Conclusion

The contribution of this article is twofold. First, we have been theoretically evaluated the performance of DM facing the combination of valumetric scaling, additive noise and constant change. The analyzes were developed under the assumptions that both the host vector and the noise vector have i.i.d components and the two vectors are independent. We accurately derived the general expressions of the PDFs of the watermarked signal, the attacked signal and the extracted signal. By these derived PDFs, the decoding error probability was generally expressed in closed form. The specific analytical results were presented for the case of generalized Gaussian host and noise. Moreover, the theoretical results can be easily extended by modeling the host and noise signals with other distributions.

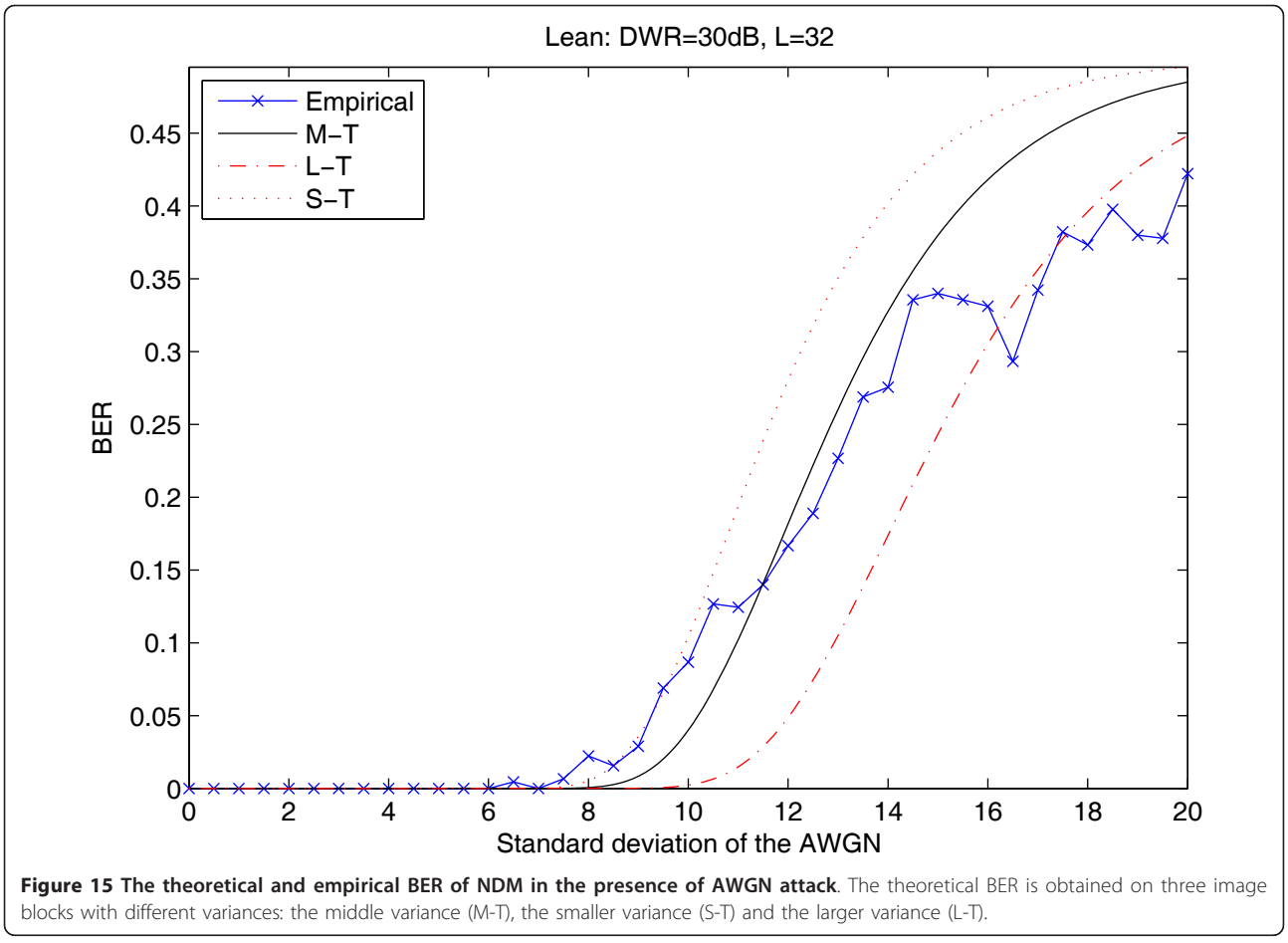

Figure 15 The theoretical and empirical BER of NDM in the presence of AWGN attack. The theoretical BER is obtained on three image blocks with different variances: the middle variance (M-T), the smaller variance (S-T) and the larger variance (L-T).

According to our analyzes, DM is largely vulnerable to valumetric scaling. And constant change and additive noise give rise to the relatively large performance loss of DM by combining them with valumetric scaling. Particularly, we have seen the effect of statistical properties of the host and noise signals on the performance of DM. The more impulsive the PDF shape of the host signal, the more robust DM is to valumetric scaling. The more flat the PDF shape of the noise source, the more sensitive DM is to additive noise. Simulations on artificial signals and real images show us that the bit-error probability is accurately predicted by the given theories for a wide range of host and noise PDF shapes. These can ultimately guide the design of efficient watermarking algorithms based on DM.

Second, a novel watermarking method, called NDM, has been developed. In the method, the normalized host signal vector is constructed for quantization. The NDM achieves its theoretical invariance to both valumetric scaling and constant change, but leads to small performance loss in the absence of channel noise. The BER of NDM against additive noise can be predicted by applying the presented theoretical results of DM. Further, the NDM is improved by weighting the quantization errors. Experiments on images demonstrate that the proposed method achieves better watermark imperceptibility and extremely strong robustness against valumetric scaling and constant change attacks comparing with the original QIM schemes and other improved versions.

Appendix

Here, we will derive the integration terms in (10) and (11) when the attacking noise obeys the distribution GGD(β_v; 0, σ_v). For this purpose, using a variable t instead of μ_{jk}, they are, respectively, rewritten as

$$\begin{aligned}\int_{-\Delta}^{\Delta} |u| p_v(t-u)du &= \int_{0}^{\Delta} u(p_v(t+u)+p_v(u-t))du \\ &= \int_{-t}^{\Delta-t} up_v(u)du + \int_{t}^{t+\Delta} up_v(u)du \\ &\quad + t\left(\int_{-t}^{\Delta-t} p_v(u)du - \int_{t}^{t+\Delta} p_v(u)du\right)\end{aligned} \tag{29}$$

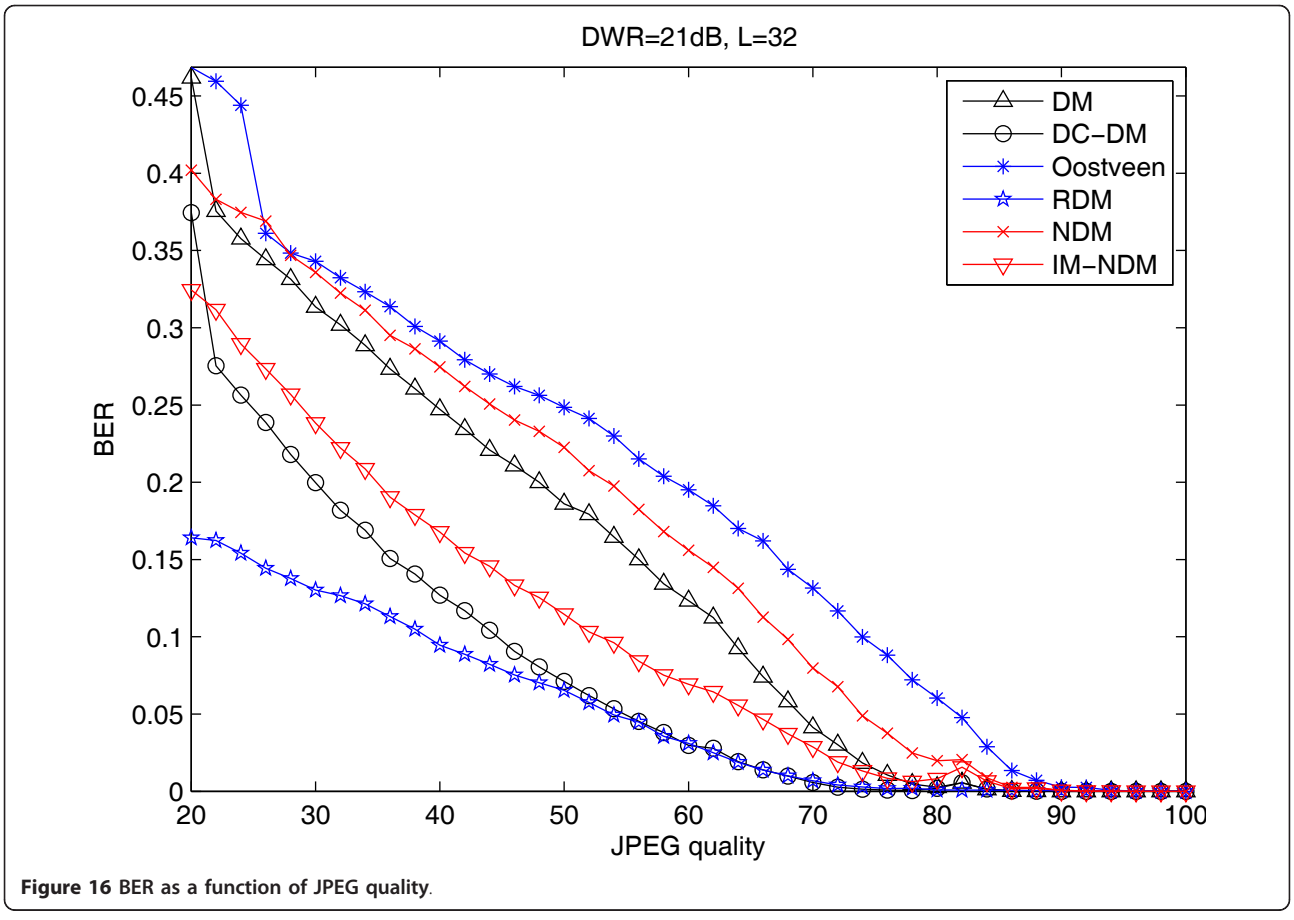

Figure 16 BER as a function of JPEG quality.

and

$$\int_{-\Delta}^{\Delta} u^2 p_v(t-u)du = \int_{t-\Delta}^{t+\Delta} (u^2 - 2tu + t^2)p_v(u)du = \int_{t-\Delta}^{t+\Delta} u^2 p_v(u)du - 2t\int_{t-\Delta}^{t+\Delta} up_v(u)du + t^2\int_{t-\Delta}^{t+\Delta} p_v(u)du \tag{30}$$

Thus, achieving the two integrations can be attributed to the computation of $I(t_1,\ t_2)$, defined as $I(t_1,\ t_2) \triangleq \int_{t_1}^{t_2} u^l p_v(u)du$ with l being an integer.

Considering the case of $t_1 \geq 0$ and $t_2 \geq 0$, we have

$$\begin{aligned} I(t_1,\ t_2) &= \frac{\kappa_v \beta_v}{2\Gamma(\beta_v^{-1})}\int_{t_1}^{t_2} u^l e^{-|\kappa_v u|^{\beta_v}}du \\ &= \frac{1}{2\Gamma(\beta_v^{-1})\kappa_v^l}\int_{(\kappa_v t_1)^{\beta_v}}^{(\kappa_v t_2)^{\beta_v}} u^{\frac{l+1}{\beta_v}-1}e^{-u}du \\ &= \frac{\gamma((l+1)\beta_v^{-1},(\kappa_v t_2)^{\beta_v}) - \gamma((l+1)\beta_v^{-1},(\kappa_v t_1)^{\beta_v})}{2\sigma_v^{-l}\sqrt{(\Gamma(\beta_v^{-1}))^{2-l}(\Gamma(3\beta_v^{-1}))^l}} \end{aligned} \tag{31}$$

where the first equality follows from (13) and the final equality follows from the definition of the lower incomplete gamma function.

In the case of $t_1 \leq 0$ and $t_2 \leq 0$, $I(t_1, t_2)$ has the form

$$\begin{aligned} I(t_1,\ t_2) &= \frac{\kappa_v \beta_v}{2\Gamma(\beta_v^{-1})}\int_{t_1}^{t_2} u^l e^{-(-\kappa_v u)^{\beta_v}}du \\ &= \frac{(-1)^{l+1}\kappa_v \beta_v}{2\Gamma(\beta_v^{-1})}\int_{-t_1}^{-t_2} u^l e^{-(-\kappa_v u)^{\beta_v}}du \\ &= \frac{\gamma((l+1)\beta_v^{-1},(-\kappa_v t_2)^{\beta_v}) - \gamma((l+1)\beta_v^{-1},(-\kappa_v t_1)^{\beta_v})}{2(-1)^{l+1}\sigma_v^{-l}\sqrt{(\Gamma(\beta_v^{-1}))^{2-l}(\Gamma(3\beta_v^{-1}))^l}} \end{aligned} \tag{32}$$

where the final equality follows from (31).

Last, while $t_1 \leq 0$ and $t_2 \geq 0$, it follows that

$$\begin{aligned} I(t_1,\ t_2) &= \frac{\kappa_v \beta_v}{2\Gamma(\beta_v^{-1})}\left(\int_{t_1}^{0} u^l e^{-(-\kappa_v u)^{\beta_v}}du + \int_{0}^{t_2} u^l e^{-(-\kappa_v u)^{\beta_v}}du\right) \\ &= \frac{\gamma((l+1)\beta_v^{-1},0) - \gamma((l+1)\beta_v^{-1},(-\kappa_v t_1)^{\beta_v})}{2(-1)^{l+1}\sigma_v^{-l}\sqrt{(\Gamma(\beta_v^{-1}))^{2-l}(\Gamma(3\beta_v^{-1}))^l}} \\ &\quad + \frac{\gamma((l+1)\beta_v^{-1},(\kappa_v t_2)^{\beta u}) - \gamma((l+1)\beta_v^{-1},0)}{2\sigma_v^{-l}\sqrt{(\Gamma(\beta_v^{-1}))^{2-l}(\Gamma(3\beta_v^{-1}))^l}} \end{aligned} \tag{33}$$

where the final equality is due to (31) and (32). Combining the three cases, a unified form of $I(t_1, t_2)$ is

$$I(t_1,\ t_2) = \frac{\gamma((l+1)\beta_v^{-1},0) - \gamma((l+1)\beta_v^{-1}, |\kappa_v t_1|^{\beta_v})}{2(sgn(t_1))^{l+1}\sigma_v^{-l}\sqrt{(\Gamma(\beta_v^{-1}))^{2-l}(\Gamma(3\beta_v^{-1}))^{l}}} + \frac{\gamma((l+1)\beta_v^{-1}, |\kappa_v t_2|^{\beta_v}) - \gamma((l+1)\beta_v^{-1},0)}{2(sgn(t_2))^{l+1}\sigma_v^{-l}\sqrt{(\Gamma(\beta_v^{-1}))^{2-l}(\Gamma(3\beta_v^{-1}))^{l}}} \quad (34)$$

By the formula (34) and the CDF of the GGD, (29) becomes

$$\int_{-\Delta}^{\Delta} |u| p_v(t-u)du = \frac{\gamma(2\beta_v^{-1}, |\kappa(t-\Delta)|^{\beta_v}) - \gamma(2\beta_v^{-1}, |\kappa t|^{\beta_v})}{2\sigma_v^{-1}\sqrt{\Gamma(\beta_v^{-1})\Gamma(3\beta_v^{-1})}} + \frac{\gamma(2\beta_v^{-1}, |\kappa(t+\Delta)|^{\beta_v}) - \gamma(2\beta_v^{-1}, |\kappa t|^{\beta_v})}{2\sigma_v^{-1}\sqrt{\Gamma(\beta_v^{-1})\Gamma(3\beta_v^{-1})}} + (2\Psi_v(t) - \Psi_v(t-\Delta) - \Psi_v(t+\Delta))t$$

and (30) becomes

$$\int_{-\Delta}^{\Delta} u^2 p_v(t-u)du = \frac{\gamma(3\beta_v^{-1},0) - \gamma(3\beta_v^{-1}, |\kappa_v(t-\Delta)|^{\beta_v})}{2sgn(t-\Delta)\sigma_v^{-2}\Gamma(3\beta_v^{-1})} + \frac{\gamma(3\beta_v^{-1}, |\kappa_v(t+\Delta)|^{\beta_v}) - \gamma(3\beta_v^{-1},0)}{2sgn(t+\Delta)\sigma_v^{-2}\Gamma(3\beta_v^{-1})} - \frac{(\gamma(2\beta_v^{-1}, |\kappa(t+\Delta)|^{\beta_v}) - \gamma(2\beta_v^{-1}, |\kappa(t-\Delta)|^{\beta_v}))t}{\sigma_v^{-1}\sqrt{\Gamma(\beta_v^{-1})\Gamma(3\beta_v^{-1})}} + (\Psi_v(t+\Delta) - \Psi_v(t-\Delta))t^2$$

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant No. 60803122, 61103018), by the Natural Science Foundation of Jiangsu Province (Grant No. BK2011442), by the Innovative Foundation of Yangzhou University (Grant No. 2011CXJ023), by the Opening Project of State Key Laboratory of Digital Publishing Technology, and by the Opening Project of State Key Laboratory of Software Development Environment (Grant No. SKLSDE-2011KF-08). The authors would like to thank the anonymous reviewers for their detailed comments that improved both the editorial and technical quality of this article substantially.

Abbreviations

AWGN: additive white Gaussian noise; CDF: cumulative distribution function; CLT: central limit theorem; DC: distortion-compensation; DC-DM: distortion compensated dither modulation; DC-NDM: distortion compensated NDM; DM: dither modulation; DWR: document-to-watermark ratio; GGD: generalized Gaussian distribution; i.i.d.: independently and identically distributed; IM-NDM: improved NDM; LQIM: logarithmic QIM; NDM: normalized dither modulation; NLPE1: number of blocks greater than the first local perceptual error threshold; NLPE2: number of blocks greater than the second local perceptual error thresh-old; PDF: probability distribution function; PMF: probability mass function; BER: bit error rate; QIM: quantization index modulation; r.v.: random variable; RDM: rational dithered modulation; SCS: scalar Costa scheme; STDM: spread transform dither modulation; TPE: total perceptual error; WNR: watermark-to-noise ratio; wPSNR: weighted peak signal-to-noise ratio.

Author details

[1]School of Information Engineering, Yangzhou University, Yangzhou 225009, China [2]State Key Laboratory of Digital Publishing Technology, Chengfu Road 298, Beijing 100871, China [3]State Key Laboratory of Software Development Environment, Beihang University, Beijing 100191, China

Competing interests

The authors declare that they have no competing interests.

Received: 20 June 2011 Accepted: 2 March 2012
Published: 2 March 2012

References

1. B Chen, GW Wornell, Quantization index modulation: a class of provably good methods fordigital watermarking and information embedding. IEEE Trans Inf Theory. **47**(4), 1423–1443 (2001). doi:10.1109/18.923725
2. JJ Eggers, R Bauml, R Tzschoppe, B Girod, Scalar costa scheme for information embedding. IEEE Trans Signal Process. **51**(4), 1003–1019 (2003). doi:10.1109/TSP.2003.809366
3. JP Boyer, P Duhamel, J Blanc-Talon, Performance analysis of scalar DC-QIM for zero-bit watermarking. IEEE Trans Inf Foren Secur. **2**(2), 283–289 (2007)
4. NK Kalantari, SM Ahadi, A logarithmic quantization index modulation for perceptually better data hiding. IEEE Trans Image Process. **19**(6), 1504–1517 (2010)
5. F Pérez-Gonzàlez, F Balado, JRH Martin, Performance analysis of existing and new methods for data hiding with known-host information in additive channels. IEEE Trans Signal Process. **51**(4), 960–980 (2003). doi:10.1109/TSP.2003.809368
6. F Bartolini, M Barni, A Piva, Performance analysis of ST-DM watermarking in presence of nonadditive attacks. IEEE Trans Signal Process. **52**(10), 2965–2974 (2004). doi:10.1109/TSP.2004.833868
7. F Pérez-Gonzàlez, C Mosquera, Quantization-based data hiding robust to linear-time-invariant filtering. IEEE Trans Inf Foren Secur. **3**(2), 137–152 (2008)
8. JH Conway, NJA Sloane, *Sphere Packings, Lattices, and Groups*, (Springer, New York, 1988)
9. F Pérez-Gonzàlez, C Mosquera, M Barni, A Abrardo, Rational dither modulation: a high-rate data-hiding method invariant to gain attacks. IEEE Trans Signal Process. **53**(10), 3960–3975 (2005)
10. JC Oostveen, AAC Kalker, M Staring, Adaptive quantization watermarking, in *Proc of SPIE: Security, Steganography, and Watermarking of Multimedia Contents VI*, San Jose, CA, **5306**, 296–303 (2004)
11. Q Li, IJ Cox, Using perceptual models to improve fidelity and provide resistance to valumetric scaling for quantization index modulation watermarking. IEEE Trans Inf Forens Secur. **2**(2), 127–139 (2007)
12. A Papoulis, *Probability, Random Variables, and Stochastic Processes*, (McGraw-Hill, New York, 1991)
13. N Saralees, A generalized normal distribution. J Appl Stat. **32**(7), 685–694 (2005). doi:10.1080/02664760500079464
14. L Schuchman, Dither signals and their effect on quantization noise. IEEE Trans Commun Technol. **CT-12**, 162–165 (1964)
15. MN Do, M Vetterli, Wavelet-based texture retrieval using generalized gaussian density and Kullback-Leibler distance. IEEE Trans Image Process. **11**(2), 146–158 (2002). doi:10.1109/83.982822
16. S Voloshynovskiy, S Pereira, V Iquise, T Pun, Attack modelling: towards a second generation watermarking benchmark. Signal Process. (Special Issue) **81**(6), 1177–1214 (2001)

doi:10.1186/1687-6180-2012-53
Cite this article as: Zhu and Ding: **Performance analysis and improvement of dither modulation under the composite attacks.** *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:53.