Exploring QIM based Anti-Collusion Fingerprinting for Multimedia

Ashwin Swaminathan, Shan He and Min Wu

Department of Electrical and Computer Engineering
University of Maryland, College Park, U.S.A

Abstract

Digital fingerprinting is an emerging technology to protect multimedia from unauthorized use by embedding a unique fingerprint signal into each user's copy. A robust embedding algorithm is an important building block in order to make the fingerprint resilient to various distortions and collusion attacks. Spread spectrum embedding has been widely used for multimedia fingerprinting. In this paper, we explore another class of embedding methods – Quantization Index Modulation (QIM) for fingerprinting applications. We first employ Dither Modulation (DM) technique and extend it for embedding multiple symbols through a basic dither sequence design. We then develop a theoretical model and propose a new algorithm to improve the collusion resistance of the basic scheme. Simulation results show that the improvement algorithm enhances the collusion resistance, while there is still a performance gap with the existing spread spectrum based fingerprinting. We then explore coded fingerprinting based on spread transform dither modulation (STDM) embedding. Simulation results show that this coded STDM based fingerprinting has significant advantages over spread spectrum based fingerprinting under blind detection.

Keywords: Quantization Index Modulation, Multimedia Fingerprinting, Collusion Resistance.

I. Introduction

With the rapid development of the internet and multimedia processing techniques, the protection of multimedia content becomes increasingly important. Digital fingerprinting is an emerging technology to protect the multimedia from unauthorized redistribution [1]. It embeds a unique ID into each user's copy, which can be extracted to help identify culprits when an unauthorized leak is found. A powerful, cost-effective attack is the collusion attack from a group of users, where the users combine their copies of the same content to generate a new version. If designed improperly, the fingerprints can be attenuated or even removed by collusion attacks.

This research was supported by U.S. Office of Naval Research N000140510634 (Young Investigator Award). The authors can be contacted by email at {ashwins, shanhe, minwu}@eng.umd.edu

Several techniques have been proposed in the literature to provide collusion resistance in multimedia fingerprinting systems [1] [2]. One category is orthogonal fingerprinting in which each user is assigned a spread spectrum [3] sequence as fingerprint, and the sequence is chosen so that they are mutually orthogonal to each other. Another category of approaches employs explicit code constructions [4] [5]. Most of these techniques use spread spectrum techniques to embed the fingerprints.

An important alternative to spread spectrum embedding is Quantization Index Modulation (QIM) [6]. In QIM, the host data is quantized using multiple quantizers, the index of which is chosen based on the message to be embedded. In this paper, we explore the possibility of employing QIM for anti-collusion fingerprinting applications. Specifically, we employ dither modulation (DM) for the fingerprint embedding. We have observed that the existing DM algorithm primarily focusses on embedding binary bits. We first construct a basic embedding scheme to resist collusion attacks by extending the existing DM to embed multiple symbols, and study its performance. To better understand the results, we introduce a general theoretical model and analyze the collusion resistance of DM based fingerprinting. From our theoretical analysis, we infer that fingerprint sequences with low correlation have better collusion resistance. We then design a new algorithm to construct dither sequences so that the resulting fingerprints have low correlation and are approximately orthogonal. We demonstrate through simulations that our proposed method performs better than the basic scheme, and compare the results with those obtained using spread spectrum based fingerprinting. Our results show that the fingerprint correlation is not easy to control through QIM embedding and hence it does not perform as well as the spread spectrum based fingerprinting even under non-blind detection.

Spread Transform Dithered Modulation (STDM) is an alternative robust quantization-based embedding approach whereby a random unitary transformation is applied to the signals before quantization. With this embedding method, every bit of information can be spread over the signal. This would have similar effect as spread spectrum based embedding, and the quantization operation during the embedding would bring benefits in blind detection scenario. Meanwhile, we notice that the existing QIM embedding techniques are well defined for embedding binary bits, and our results on DM based fingerprinting show that it is non-trivial to extend these methods to embed non-binary symbols. In principle, it would be possible to construct binary fingerprint sequences employing collusion-secure codes such as Boneh-Shaw's [2] for fingerprinting multimedia. However, since these codes are designed without considering the embedding issues explicitly, they are often too long to be reliably embedded [2], and/or are unable to resist a nontrivial number of colluders [7]. For example, to attain moderate levels of collusion resistance such as to resist 10 colluders out of 1000 users, the Boneh-Shaw code requires a long codeword at least on the order of 10^6 bits. Such high payloads often exceed the embedding capacity for most multimedia data under stringent robustness requirements. In this paper, we propose to use non-binary fingerprint code, such as traceability code employed in [8] and map each symbol to a binary codeword through an efficient construction for embedding.

II. Background on Fingerprinting and QIM

A. Spread Spectrum based Fingerprinting

Spread spectrum embedding has been widely used for multimedia fingerprinting [4] [8] [9]. One typical example is orthogonal fingerprinting, whereby mutually orthogonal spreading sequences are generated as fingerprint for each user. Another way to construct fingerprint is to employ a coding step, such as error

correcting code (ECC), and map symbols in the alphabet to orthogonal sequences [5]. The i^{th} user's fingerprinted copy $\mathbf{y}_i$ is obtained by adding his/her fingerprint sequence $\mathbf{u}_i$ to the host signal $\mathbf{x}$, i.e.

$$\mathbf{y}_i = \mathbf{x} + \mathbf{u}_i. \tag{1}$$

After the fingerprinted copies reach end users, some users may mount collusion attacks and try to remove the traces of the embedded fingerprint. Averaging collusion plus additive noise is mostly studied in the literature [1] [9] and a number of non-linear collusions have been shown to be well approximated by this model [10]. Under averaging collusion, the resulting signal, $\mathbf{z}$, is the average of c colluders' fingerprinted copy:

$$\mathbf{z} = \frac{1}{c}\sum_{i \in S_c} \mathbf{y}_i + \mathbf{n}, \tag{2}$$

where S_c is the colluder set containing c colluders, $\mathbf{n} = [n_1, n_2, \ldots, n_N]^T$ is additive noise that models additional distortions applied on the colluded signal, and N is the length of the fingerprint sequence. For simplicity, we assume $\mathbf{n}$ follows an $i.i.d.$ Gaussian distribution.

The goal of the detector is to catch at least one of the colluders with a high probability given the suspicious copy, $\mathbf{z}$. As the host signal can be made available to detectors in many fingerprinting applications, we subtract the host signal from the suspicious copy to obtain a test signal. Match filter detector is then employed to find the colluder; that is, we correlate the test signal with each of the N_u spreading sequences (one for each user) and identify the sequence that gives the maximum correlation. The detection statistic for the i^{th} user is defined as

$$T_i = \frac{(\mathbf{z} - \mathbf{x})^T \mathbf{u}_i}{\sqrt{||\mathbf{u}_i||^2}}, \tag{3}$$

and the $\hat{m}^{th}$ user is declared as a colluder if

$$\hat{m} = \arg\max_{i=1,2,\ldots,N_u} T_i. \tag{4}$$

B. Quantization Index Modulation(QIM)

1) Dither Modulation(DM): In quantization based methods, the host data is quantized using multiple quantizers and the index of the quantizer is chosen based on the message to be embedded [6]. A simple way to build multiple quantizers is by dither modulation (DM). Specifically, for a host-signal $\mathbf{x}$, the embedding function for hiding binary messages can be written as

$$\mathbf{q}_{x_i} = Q_\Delta(\mathbf{x} + \mathbf{d}_i) - \mathbf{d}_i \quad \forall i \in \{0, 1\}, \tag{5}$$

where $Q_\Delta(.)$ represents the quantization function with step size Δ and $\mathbf{d}_i$ represents the dither sequence that is used to perturb the host signal before quantization. One possible way to construct the dither sequence is by first choosing one dither vector (say $\mathbf{d}_0$) as $i.i.d.$ random variables following a uniform distribution over $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$ and then the second one can be obtained using [6]

$$d_{1k} = \begin{cases} d_{0k} + \frac{\Delta_k}{2} & \text{if} \quad d_{0k} < 0, \\ d_{0k} - \frac{\Delta_k}{2} & \text{if} \quad d_{0k} \geq 0, \end{cases} \quad \forall k \in \{1, 2, \ldots, N\}, \tag{6}$$

where $\mathbf{d}_i = [d_{i1}, d_{i2}, \ldots, d_{iN}]^T$.

It has been shown in [6] [11] that the rate-distortion and robustness tradeoff can be improved in the basic QIM method by compensation and other postprocessing operations. In the distortion compensated QIM (DC-QIM), a fraction of the quantization error is added back to the original signal. Thus, the watermarked image can be represented as

$$\mathbf{y}_i = \alpha \left(Q_{\underset{\alpha}{\Delta}}(\mathbf{x} + \mathbf{d}_i) - \mathbf{d}_i \right) + (1-\alpha)\mathbf{x} \quad \forall i \in \{0,1\}, \tag{7}$$

where the constant α can be chosen appropriately to maximize the Signal-to-Noise Ratio (SNR) [6] or to maximize embedding capacity [11] [12].

2) Spread Transform Dither Modulation (STDM): Another robust way to implement QIM is STDM. Instead of applying scalar DM directly on each component of host signal, STDM first applies a random unitary transformation by projecting the host signal $\mathbf{x}$ onto a random direction, $\mathbf{u}$. The projection values, $x^{(p)} = \mathbf{u}^T\mathbf{x}$, are then quantized using DM to obtain the watermarked signal [6], i.e.

$$\mathbf{y} = \mathbf{x} + (y^{(p)} - x^{(p)})\mathbf{u}, \tag{8}$$

$$y^{(p)} = Q_{\triangle}(x^{(p)} + d_b) - d_b, \quad b \in \{0,1\}, \tag{9}$$

where b is the message bit to be embedded. Typically, the projection direction $\mathbf{u}$ is randomly generated according to a secret key, and therefore we do not need to introduce uncertainty in the choice of dither sequence d_b. In our implementation, we choose the dither sequences to be deterministic. Due to random projections, only the noise in the direction of $\mathbf{u}$ would affect performance. Thus, the STDM provides a higher effective Watermark to Noise Ratio (WNR), and is more robust against additive noise attacks [6].

During the detection, the test signal $\mathbf{z} = \mathbf{y} + \mathbf{n}$, is projected onto vector $\mathbf{u}$ to get $z^{(p)} = \mathbf{z}^T\mathbf{u}$. The embedded bit is determined as

$$\hat{m} = \begin{cases} \arg\min_{b=0,1} \|z^{(p)} - (Q_{\triangle}(x^{(p)} + d_b) - d_b)\| & \text{for non-blind detection,} \\ \arg\min_{b=0,1} \|z^{(p)} - (Q_{\triangle}(z^{(p)} + d_b) - d_b)\| & \text{for blind detection.} \end{cases} \tag{10}$$

III. Dither Modulation based Fingerprinting

A. Extending QIM to Fingerprinting

It is known in the recent literature that lattice-based quantizers can be used to embed multiple alphabets [13], but they generally have a very high computational complexity. To overcome this problem, we consider a simple extension of the DM scheme for embedding multiple symbols, i.e. use mutually orthogonal dither sequences for each user. Specifically, we construct N_u random dither sequences $\mathbf{d}_i$ following an $i.i.d.$ Gaussian distribution such that $E(\mathbf{d}_i^T\mathbf{d}_j) = 0 \ \forall i,j \in \{1,...,N_u\}$ and $i \neq j$. The fingerprinted copies are then obtained using equation (7).

When the content owner obtains the suspicious copy $\mathbf{z}$, he/she can apply maximum likelihood detection, which would involve an exhaustive search over $O(2^{N_u})$ different colluder combinations. Although this detector is optimal in minimizing the probability of detection error, its complexity is very high and grows exponentially with the number of users. Therefore, in our implementation, we apply the minimum-distance detection as used in the QIM literature [6] to find one of the colluders. More specifically, the $\hat{m}^{th}$ user is declared a colluder if

$$\hat{m} = \arg\min_{k=1,2,...N_u} ||\mathbf{z} - \mathbf{y}_k||^2. \tag{11}$$

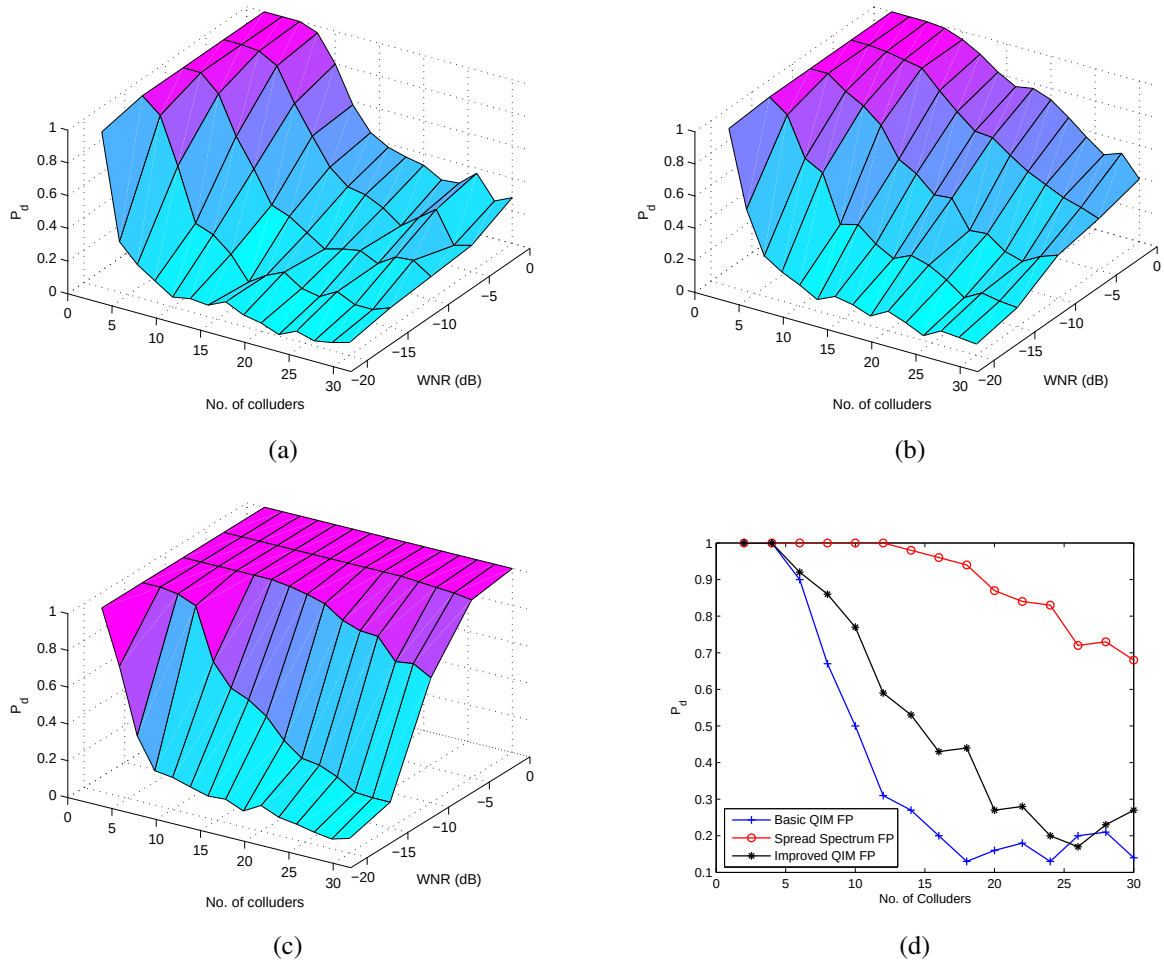
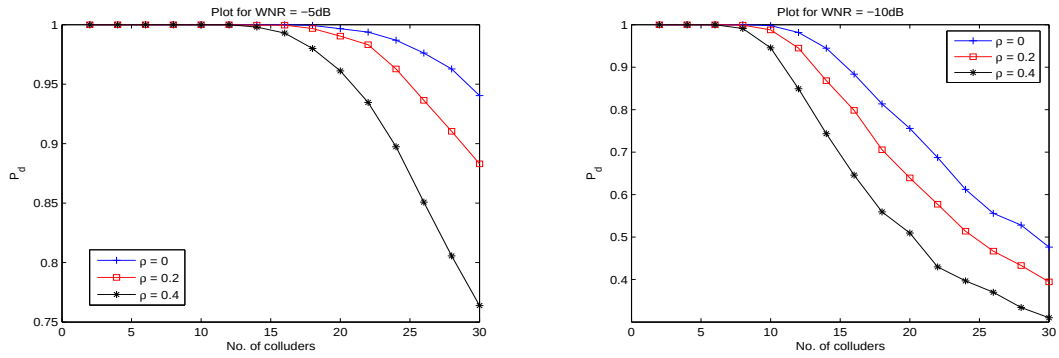


Fig. 1. Comparison on the performance of QIM-based and spread spectrum based fingerprinting: (a) Basic QIM-based Fingerprinting; (b) Improved QIM-based Fingerprinting; (c) Spread spectrum based Fingerprinting; (d) Results under WNR$=-10$dB.

This detector also provides a fair comparison with the spread spectrum based fingerprinting employing match filter detection of equation (3) [9].

We simulated this basic scheme for $N_u = 1024$ users under averaging collusion on a 256×256 size Lena image with the PSNR of the fingerprinted image with respect to the original set to 42dB. The embedding was done in the block DCT domain and the quantization step sizes were chosen according to the JPEG quantization table. We examine the probability of catching one colluder, P_d, at different watermark-to-noise-ratio (WNR), and the results are shown in Fig. 1(a). For comparison purposes, we show in Fig. 1(c) the performance of a spread spectrum based fingerprinting under the same conditions. From the results, we observe that the basic QIM-based fingerprinting can only resist about half dozen colluders at moderate to high WNRs, while spread spectrum based fingerprinting can resist more than 30 colluders with high probability in the same WNR range. To facilitate the analysis of results, we build a theoretical model to study the detection performance in the next subsection.



(a) WNR = -5dB (b) WNR = -10dB

Fig. 2. Theoretical results on probability of correct detection with respect to number of colluders for different ρ values

B. Theoretical Analysis of QIM-based Fingerprinting

Without loss of generality, we assume the first c colluders perform averaging collusion as formulated in equation (2). Then, the probability of catching one colluder, P_d, is

$$P_d = Pr(\min(X_1, X_2, \ldots, X_c) < \min(X_{c+1}, X_{c+2}, \ldots, X_{N_u})), \tag{12}$$

where $X_k = ||\mathbf{z} - \mathbf{y}_k||^2$ is the detection statistic for user k. We can show that for a system with totally N_u users and c colluders, $\mathbf{X} = [X_1, X_2, \ldots, X_{N_u}]^T$ approximately follows a multi-variate Gaussian distribution with mean and covariance matrix given by:

$$m_k = E(X_k) = \begin{cases} \left(\frac{c-1}{c}\right)\frac{\Lambda}{2} + N\sigma_n^2 & \text{if } \ 1 \le k \le c, \\ \left(\frac{c+1}{c}\right)\frac{\Lambda}{2} + N\sigma_n^2 & \text{if } \ c+1 \le k \le N_u, \end{cases} \tag{13}$$

$$R(i,j) = cov(X_i, X_j) = 2\sigma_n^4 \left[N + \left(\frac{\Lambda}{\sigma_n^2}\right) \frac{P(i,j)}{c} \right], \tag{14}$$

where N is the length of fingerprint sequence, σ_n^2 is variance of the additive noise, and Λ is the average mean square difference between two fingerprinted copies. The matrix $P_{(N_u \times N_u)}$ is given by

$$P(i,j) = \begin{cases} c-1 & \text{if } \ 1 \le i,j \le c & \text{and } \ i = j, \\ -1 & \text{if } \ 1 \le i,j \le c & \text{and } \ i \ne j, \\ 0 & \text{if } \ 1 \le i \le c & \text{and } \ j > c, \\ 0 & \text{if } \ 1 \le j \le c & \text{and } \ i > c, \\ c+1 & \text{if } \ c < i,j \le N_u & \text{and } \ i = j, \\ 1 & \text{if } \ c < i,j \le N_u & \text{and } \ i \ne j. \end{cases} \tag{15}$$

The detailed derivation is described in Appendix A. We remark that the above theoretical framework is general and applicable to any fingerprinting scheme as long as the distance between any pair of fingerprints is identical. Thus, this model can help explain the results obtained by spread spectrum techniques as well.

From equation (13), we notice that the difference between the means of X_k for the colluders and the innocent users is $\Delta m = \frac{\Lambda}{c}$. Thus, we infer that the average performance in terms of probability of catch one colluder would improve as the distance between two fingerprint sequences, Λ, is increased, or the

number of colluders c is decreased. This result can also be interpreted in terms of the correlation between the fingerprint sequences $\rho = 1 - \frac{\Lambda}{2W}$ (W is the average energy of the fingerprint). In Fig. 2, we show the probability of correct decision P_d for different values of the correlation parameter ρ by numerically evaluating the theoretical model. We observe from the plot that the performance of the fingerprinting system increases when ρ reduces (or Λ increases). Based on this principle, we examine the correlation for basic QIM-based fingerprinting. We observe that the main reason for our basic QIM construction not performing well compared to the spread spectrum case is because the resulting correlation value $\rho = 0.45$ was much higher than that of the spread spectrum based fingerprinting (close to zero). Therefore, in order to improve the collusion resistance of QIM-based fingerprinting, we need to carefully select dither sequences so that the resulting fingerprint sequences have low correlation. In the next section, we propose a new technique that will help reduce the correlation and improve the detection performance.

C. Improved Dither Sequence Construction for QIM-based Fingerprinting

According to the theoretical model, for best results, the dither sequences should be constructed so as to make the final fingerprints have as low correlation as possible. The problem can be formulated as

$$\min \quad (Q_\Delta(\mathbf{x}+\mathbf{d}_i) - \mathbf{x} - \mathbf{d}_i)^T (Q_\Delta(\mathbf{x}+\mathbf{d}_j) - \mathbf{x} - \mathbf{d}_j), \quad \forall\ i,j \in \{1,2,\ldots,N_u\}, i \neq j, \tag{16}$$

subject to the fairness constraints that the fingerprint energies for different users are equal, i.e.

$$(Q_\Delta(\mathbf{x}+\mathbf{d}_i) - \mathbf{x} - \mathbf{d}_i)^T (Q_\Delta(\mathbf{x}+\mathbf{d}_i) - \mathbf{x} - \mathbf{d}_i) = W, \quad \forall i = 1,2,\ldots,N_u. \tag{17}$$

Let $\Delta = [\Delta_1, \Delta_2, \ldots, \Delta_N]^T$, where Δ_k is the step size of the uniform quantizer in the k^{th} component. We can show that the quantization operation (for the mid-raiser quantizer) is given by

$$Q_\Delta(\mathbf{x}+\mathbf{d}_i) = \mathbf{a} + \frac{1}{2}\Delta \otimes \mathbf{Y}_i, \tag{18}$$

where $\mathbf{a} = [a_1, a_2, \ldots, a_N]^T$, $\mathbf{Y}_i = [Y_{i1}, Y_{i2}, \ldots, Y_{iN}]^T$ and $\Delta \otimes \mathbf{Y}_i = [\Delta_1 Y_{i1}, \Delta_2 Y_{i2}, \ldots, \Delta_N Y_{iN}]^T$. The corresponding k^{th} element in the vector can be represented as

$$a_k = \begin{cases} t_k\Delta_k & \text{if} \quad t_k\Delta_k \leq x_k < (t_k+0.5)\Delta_k, \\ (t_k+1)\Delta_k & \text{if} \quad (t_k+0.5)\Delta_k \leq x_k < (t_k+1)\Delta_k; \end{cases} \tag{19}$$

$$Y_{ik} = \begin{cases} -1 & \text{if} \quad \frac{-\Delta_k}{2} \leq d_{ik} < (a_k - x_k), \\ 1 & \text{if} \quad (a_k - x_k) \leq d_{ik} < \frac{\Delta_k}{2}. \end{cases} \tag{20}$$

Here, we assume that $-\frac{\Delta_k}{2} \leq d_{ik} < \frac{\Delta_k}{2}$. Note that a_k is a multiple of the quantization step size Δ_k, that is closest to the host data sample x_k. Further, the value of a_k is independent of the choice of the dither sequence. The term $\frac{1}{2}\Delta_k Y_{ik}$ denotes the residue term that would choose one among the two nearby quantization points based on the value of the dither sequence.

By substituting equations (19) and (20) back into the minimization problem, and using the Lagrange multipliers to incorporate the equal-energy constraints we obtain an equivalent cost function$-J$ given by

$$\begin{aligned} J &= (\mathbf{a} - \mathbf{x} + \frac{1}{2}\Delta \otimes \mathbf{Y}_i - \mathbf{d}_i)^T (\mathbf{a} - \mathbf{x} + \frac{1}{2}\Delta \otimes \mathbf{Y}_j - \mathbf{d}_j) \\ &+ \nu_1 \left((\mathbf{a} - \mathbf{x} + \frac{1}{2}\Delta \otimes \mathbf{Y}_i - \mathbf{d}_i)^T (\mathbf{a} - \mathbf{x} + \frac{1}{2}\Delta \otimes \mathbf{Y}_i - \mathbf{d}_i) - W \right) \\ &+ \nu_2 \left((\mathbf{a} - \mathbf{x} + \Delta \otimes \mathbf{Y}_j - \mathbf{d}_j)^T (\mathbf{a} - \mathbf{x} + \frac{1}{2}\Delta \otimes \mathbf{Y}_j - \mathbf{d}_j) - W \right), \end{aligned} \tag{21}$$

where ν_1 and ν_2 are Lagrange multiplier constants. Setting the gradient of J with respect to both the dither vectors to zero, we get a set of linear equations solving which we obtain

$$\mathbf{d}_i = \frac{1}{2}\Delta \otimes \mathbf{Y}_i - \kappa_i(\mathbf{x} - \mathbf{a}), \tag{22}$$

where κ_i are scalars chosen so that the total energy of the fingerprint is equal to W. In our implementations, we first choose the vectors $\mathbf{Y}_i \in \{-1, 1\}^N$ for each user. The dither sequences are then generated according to equation (22). In the next section, we present the results for this scheme and compare it with our basic dither based approach presented earlier in Section III-A.

D. Results and Discussions

To examine the effectiveness of the proposed improvement algorithm, we apply the constructed dither sequences on the Lena image with the same parameter settings as in Section III-A; that is, 256×256 Lena image fingerprinted with a PSNR of 42dB and $N_u = 1024$ users. The results are shown in Fig. 1(b) alongside the corresponding plots for the basic QIM scheme and spread spectrum fingerprinting. For better illustration, we compare the performance of the three schemes at WNR = -10dB in Fig. 1(d). We observe that the improved scheme performs much better than the basic scheme. This gain can be attributed to the reduced average correlation among fingerprint sequences in the improved scheme (around 0.1), compared to a high value of 0.45 in the basic scheme. We also observe that our improved scheme still does not perform as well as the traditional spread spectrum based scheme. A closer examination shows that the variance of the correlation statistic for QIM based fingerprinting is larger (ρ values range from -0.15 to 0.2), while the spread spectrum based fingerprinting has correlation ranging from -0.04 to 0.04. Owing to the nonlinear quantization operation employed in QIM, it is not easy to control the correlation between the fingerprints for a total of N_u users as in spread spectrum based fingerprinting.

From the results, we can see that in the proposed DM based fingerprinting it is non-trivial to construct dither sequences to get fingerprint sequences with low correlation.Since the QIM has been well studied for embedding binary bits, a natural way of using QIM for fingerprinting is to embed binary fingerprint codeword. In the mean time, we observe that STDM based embedding has an effect of spreading the embedded bit over the host signal. By choosing mutually orthogonal projection vectors for different bits, we can achieve an effect similar to the overlapped spread spectrum embedding. Taking these two factors into consideration, we explore the STDM based coded fingerprinting in the next section.

IV. STDM BASED ECC FINGERPRINTING

ECC based fingerprinting with spread spectrum embedding has been shown very promising in providing an excellent trade-off between the collusion resistance and detection efficiency [8]. In this section, we explore the performance of STDM based ECC fingerprinting. For embedding, we propose to map each symbol to a binary codeword that is constructed to well separate q symbols.

A. Fingerprint Embedding and Detection

To embed a q-ary fingerprint codeword with length L_1, we partition the host signal into L_1 segments. In each segment, we choose a simplex code $S(L_2, m, D)$ to represent each of the q symbols. A simplex code of dimension m has $q = 2^m$ codewords, each of length $L_2 = 2^m - 1$ and provides an equal distance

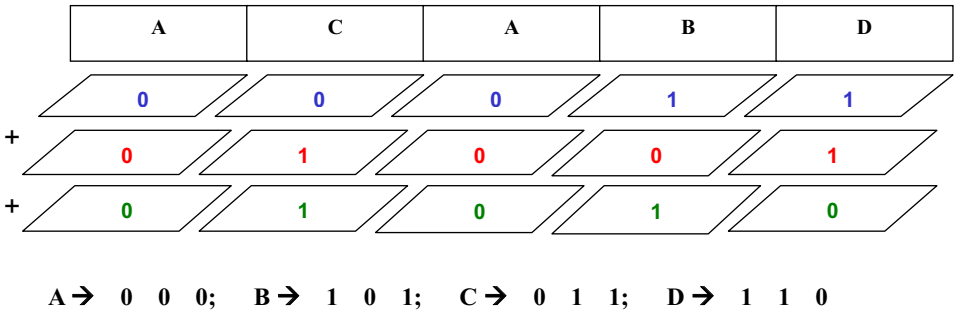


Fig. 3. An example of embedding q-ary codeword using STDM.

of $D = 2^{m-1}$. Simplex code has good properties such as a large relative distance (> 0.5) and a non-trivial code rate; and thus it can support a large alphabet size q with better separation among symbols. The binary simplex codeword for each symbol is embedded into a segment of the host signal through STDM, where we project the host signal to L_2 mutually orthogonal random directions and quantize the resulting projection values. This has an overall effect of overlapped embedding that the bits representing one symbol are added on top of each other and spread over the segment. An illustration is shown in Fig. 3.

In the implementation, to get a better robustness, we propose to spread each bit of the simplex codeword to multiple bits by mapping bit 1 to a l-bit random binary sequence $\mathbf{s}$ and 0 to $\bar{\mathbf{s}}$, the bit-wise flipped version of $\mathbf{s}$. The spreading factor l can be adjusted to tradeoff the perceptibility and robustness. For clear presentation, we shall use "logic bit" to refer to the bit in the simplex codeword, and "bit" refers to the bit in the spreading sequence $\mathbf{s}$. Every bit in the sequence $\mathbf{s}$ is embedded into the same segment using STDM by projecting the signal onto a random direction. As a result, a total of $L_2 l$ bits for all the L_2 logic bit in a simplex codeword are superposed and spread over one segment of host signal. The k^{th} fingerprinted segment $\mathbf{y^{(k)}}$ can be represented as

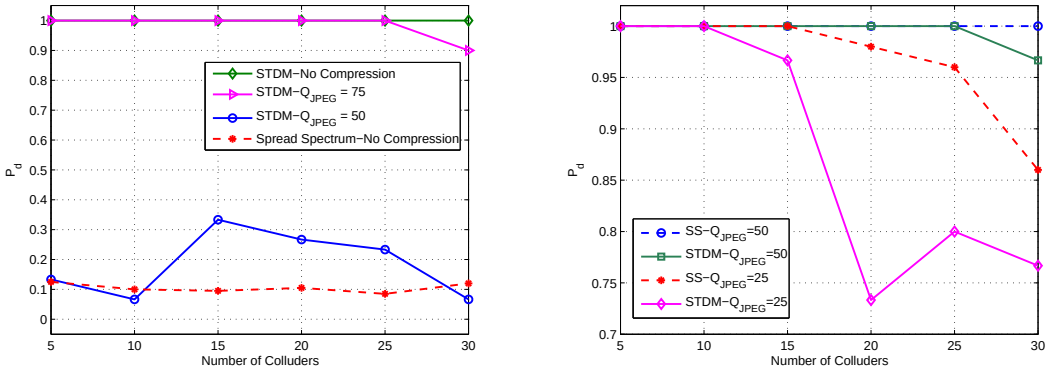
$$\mathbf{y^{(k)}} = \mathbf{x^{(k)}} + \sum_{i=1}^{L_2} \sum_{j=1}^{l} (y_{ij}^{(p)} - x_{ij}^{(p)}) \mathbf{u_{ij}}, \tag{23}$$

where $\mathbf{u_{ij}}$ is the projecting direction for the i^{th} bit in logic bit j's spreading sequence; $x_{ij}^{(p)}$ is the projection of the k^{th} segment host signal $\mathbf{x^{(k)}}$ on $\mathbf{u_{ij}}$, and $y_{ij}^{(p)}$ is obtain by quantizing $x_{ij}^{(p)}$ using equation (8).

During the detection, we first calculate the distance information for each bit according to equation (10). Then we add all these distance information from every bit of each user's fingerprint codeword. The user who has the smallest distance with the test signal is declared as colluder.

B. Results and Discussions

We test the performance of the proposed STDM based ECC fingerprinting on a 256×256 Lena image. We choose Reed-Solomon code (14, 2, 13) as the fingerprint code with code length 14, dimension 2, and alphabet size 16. Each of the 16 symbols is mapped to a binary codeword of a simplex code (15, 4, 8). Mutually orthogonal spreading sequences are chosen for projecting the input data and the resulting values are quantized using the binary dither modulation method. As mentioned earlier, we choose deterministic



(a) (b)

Fig. 4. Simulation results of STDM based and spread spectrum based ECC fingerprinting under averaging collusion and JPEG compression: (a) blind detection; (b) non-blind detection.

dither sequences $d_0 = 0$ and $d_1 = \triangle/2$ to maximize separation. The PSNR of the fingerprinted copy is set at 40.8 dB.

In Fig. 4, we show the probability of catching one colluder, P_d, under averaging collusion and additional JPEG compression. The results for the blind and non-blind scenarios are shown in Fig. 4(a) and (b) respectively. We notice that under moderate JPEG compression, the system is able to resist at least a few dozen users' collusion in both cases. When the JPEG quality factor reduces, the performance drops sharply in the case of blind detection even for a small number of colluders. This is expected because the projected point $z^{(p)}$ moves outside the correct decoding region when a large JPEG quantization step size is used. This leads to wrong estimates of the true projection points $x^{(p)}$, eventually resulting in a large probability of decoding error. On the other hand, in the case of non-blind detection, the projected point $z^{(p)}$ provides some information for correct decoding. Therefore, the performance of non-blind detection degrades gracefully as the JPEG quality factor reduces and the number of colluders increases.

To facilitate comparison, we also implement the spread spectrum based ECC fingerprinting with the same Reed-Solomon code, i.e. each symbol is mapped to an orthogonal spreading sequences before embedding [8]. Match filter detection is employed for catching one colluder. Also in Fig. 4, we show the results for spread spectrum based fingerprinting under both the blind detection and non-blind detection. Under blind detection, we notice that the spread spectrum based fingerprinting performs much worse than STDM based scheme even without further compression. On the other hand, the spread spectrum based ECC fingerprinting performs a little better than STDM based scheme under non-blind detection. This is because the spread spectrum based scheme employs orthogonal modulation to embed each symbol, while STDM based scheme use a simplex code, which does not perform as well as orthogonal modulation in separating different symbols. Overall, the proposed STDM based fingerprinting shows significant advantages over spread spectrum based fingerprinting under blind detection and slightly reduced performance under non-blind detection.

V. Conclusions

In this paper, we consider using Quantization Index Modulation as an alternative to the popular spread spectrum technique for fingerprinting applications. We first present a dither sequence construction to extend the existing QIM technique to embed multiple symbols. We develop a generalized theoretical model and show that the performance of the fingerprinting scheme can be improved by reducing the effective correlation between the fingerprints. We propose an improved QIM fingerprinting scheme that reduces the overall correlation and thus achieves better performance. The comparison results with spread spectrum based fingerprinting show that the fingerprint correlation is not easy to control through QIM embedding, and thus it does not perform as well as spread spectrum based fingerprinting.

We then explore the capability of QIM for coded fingerprinting. In particular, we use spread transform dither modulation (STDM) to embed the fingerprint code, where each q-ary symbol is mapped to a binary simplex code for embedding. The results show significant advantage of STDM based embedding over spread spectrum based embedding under blind detection and slightly reduced performance under non-blind detection. This suggests that STDM is a promising embedding technique for fingerprinting under blind detection scenarios, which we plan to explore further in future work.

Appendix A: General Theoretical Model

In this appendix, we present the theoretical analysis on the performance of fingerprinting schemes employing minimum distance decoding. Let $\mathbf{x}$ denote the host signal and $\mathbf{y}_i$ represent user i's fingerprinted copy. In the case of QIM, $\mathbf{y}_i$ is obtained by quantizing the host signal $\mathbf{x}$ as shown in equation (7). Under the averaging collusion model, the received signal, $\mathbf{z}$, is given by

$$\mathbf{z} = \frac{1}{c}\sum_{i=1}^{c}\mathbf{y}_i + \mathbf{n}, \tag{24}$$

where $\mathbf{n}$ denotes the additive noise used to model any further processing. Here, we assume, without loss of generality, that the first c colluders participate in collusion. The decoder applies minimum distance decoding as given in equation (11) to find one of the colluders.

The probability of catching one colluder, P_d, is given by

$$P_d = Pr(\min(X_1, X_2, \ldots, X_c) < \min(X_{c+1}, X_{c+2}, \ldots, X_{N_u})), \tag{25}$$

where $X_k = ||\mathbf{z} - \mathbf{y}_k||^2$ is the detection statistic for user k. Substituting for $\mathbf{z}$ from equation (24), we get

$$X_k = ||\mathbf{n} + \frac{\alpha}{c}\sum_{i=1}^{c}\left(\mathbf{q}_{x_i} - \mathbf{q}_{x_k}\right)||^2. \tag{26}$$

The detection statistic X_k is a random variable and its distribution would depend on the noise statistics. The mean of X_k can be obtained as

$$m_k = E(X_k) = \mathbf{s}_k^T\mathbf{s}_k + trace(\Sigma_n), \tag{27}$$

where Σ_n is the covariance matrix of the noise variable $\mathbf{n}$ and $\mathbf{s}_k = \frac{\alpha}{c}\sum_{i=1}^{c}\left(\mathbf{q}_{x_i} - \mathbf{q}_{x_k}\right)$ gives the average difference between the quantization points among the users in collusion set. In a similar note, the $(i,j)^{th}$ element of the covariance matrix R of the detection statistics X_k can be expressed as

$$R(i,j) = cov(X_i, X_j) = {\mathbf{w}_n^{(4)}}^T\mathbf{1}_N - trace(\Sigma_n^T\Sigma_n) + 2{\mathbf{w}^{(3)}}^T(\mathbf{s}_i + \mathbf{s}_j) + 4\mathbf{s}_i^T\Sigma_n\mathbf{s}_j, \tag{28}$$

where $\mathbf{1}_N$ denotes a column vector with all N elements as 1 and $\mathbf{w}_n^{(l)}$ is a $N \times 1$ vector in which the i^{th} element represents the l^{th} order moment of the corresponding noise component n_i.

If we assume that the noise $\mathbf{n}$ is Gaussian with zero mean and variance σ_n^2, then the detection statistic would follow the chi-square distribution [14] and its mean and variance can be simplified as

$$m_k = \mathbf{s}_k^T \mathbf{s}_k + N\sigma_n^2, \tag{29}$$

$$R(i,j) = 2N\sigma_n^4 + 4\sigma_n^2 \mathbf{s}_i^T \mathbf{s}_j. \tag{30}$$

We remark that as the length of the fingerprint N is increased, the detection statistic can be well approximated as a multi-variate Gaussian distribution [14]. Substituting for $\mathbf{s}_i$, we obtain

$$\mathbf{s}_i^T \mathbf{s}_j = \left(\frac{\alpha}{c}\right)^2 \sum_{l=1}^{c} \sum_{k=1}^{c} (\mathbf{q}_{x_l} - \mathbf{q}_{x_i})^T (\mathbf{q}_{x_k} - \mathbf{q}_{x_j}), \tag{31}$$

which can be further reduced to give

$$\mathbf{s}_i^T \mathbf{s}_j = \frac{P(i,j)\Lambda}{2c}. \tag{32}$$

Here $P(i,j)$ is given as in equation (15) and Λ is the average mean squared difference between any two fingerprinted copies,

$$\Lambda = \frac{2}{N_u(N_u-1)} \sum_{i=1}^{N_u} \sum_{j=1, j\neq i}^{N_u} ||\mathbf{y}_i - \mathbf{y}_j||^2. \tag{33}$$

Substituting for $\mathbf{s}_i^T \mathbf{s}_j$ from equation (32) into equations (29) and (30), we obtain the desired expressions as given in equations (13) and (14).

REFERENCES

[1] M. Wu, W. Trappe, Z. J Wang, and K. J. R. Liu, "Collusion Resistant Fingerprinting for Multimedia," *IEEE Signal Processing Magazine*, Vol. 21, No. 2, pp 15–27, March 2004.

[2] D. Boneh and J. Shaw, "Collusion-secure Fingerprinting for Digital Data," *IEEE Transactions on Information Theory*, Vol. 44, No. 5, pp. 1897–1905, September 1998.

[3] I. Cox, J. Killian, F. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Transactions on Image Processing*, Vol. 6, No. 12, pp. 1673–1687, December 1997.

[4] W. Trappe, M. Wu, Z. J. Wang, and K. J. R. Liu, "Anti-collusion Fingerprinting for Multimedia," *IEEE Transactions on Signal Processing, Special issue on Signal Proceedings for Data Hiding in Digital Media & Secure Content Delivery*, Vol. 51, No. 4, pp.1069–1087, April 2003.

[5] S. He and M. Wu, "Performance Study of ECC-based Collusion-resistant Multimedia Fingerprinting," in *Proceedings of the 38th Conference on Information Sciences and Systems (CISS)*, pp. 827–832, Princeton, NJ, March 2004.

[6] 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, May 2001.

[7] A. Barg, G. R. Blakley, and G. Kabatiansky, "Digital Fingerprinting Codes: Problem Statements, Constructions, Identification of Traitors," *IEEE Transactions of Information Theory*, Vol. 49, No. 4, pp. 852–865, April 2003.

[8] S. He and M. Wu, "Improving Collusion Resistance of Error Correcting Code Based Multimedia Fingerprinting," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*, Vol. 2, pp. 1029–1032, Philadelphia, PA, March 2005.

[9] Z. J. Wang, M. Wu, H. Zhao, W. Trappe, and K. J. R. Liu, "Anti-Collusion Forensics of Multimedia Fingerprinting Using Orthogonal Modulation," *IEEE Transactions on Image Processing*, Vol. 14, No. 6, pp. 804–821, June 2005.

[10] H. V. Zhao, M. Wu, Z. J. Wang, and K. J. R. Liu, "Forensic Analysis of Nonlinear Collusion Attacks for Multimedia Fingerprinting," *IEEE Transactions on Image Processing*, Vol. 14, No. 5, pp.646–661, May 2005.

[11] J. J. Eggers, R. Bauml, R. Tzschoppe, and B. Girod, "Scalar Costa Scheme for Information Embedding," *IEEE Transactions on Signal Processing*, Vol. 51, No. 4, pp. 1003–1019, April 2003.

[12] P. Moulin and R. Koetter, "Data-Hiding Codes," *Proceedings of the IEEE*, Vol. 93, No. 12, pp. 2083–2126, December 2005.

[13] Q. Zhang and N. Boston, "Quantization Index Modulation using the E_8 lattice," in *Proceedings of the 41th Annual Allerton Conference on Communication, Control and Computing*, Allerton, IL, USA, October 2003.

[14] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw Hill publications, 2002.