# Quantizer characteristics important for Quantization Index Modulation

Hugh Brunk

Digimarc, 19801 SW 72nd Ave., Tualatin, OR 97062

## ABSTRACT

Quantization Index Modulation (QIM) has been shown to be a promising method of digital watermarking. It has recently been argued that a version of QIM can provide the best information embedding performance possible in an information theoretic sense. This performance can be demonstrated via random coding using a sequence of vector quantizers of increasing block length, with both channel capacity and optimal rate-distortion performance being reached in the limit of infinite quantizer block length. For QIM, the rate-distortion performance of the component quantizers is unimportant. Because the quantized values are not digitally encoded in QIM, the number of reconstruction values in each quantizer is not a design constraint, as it is in the design of a conventional quantizer. The lack of a rate constraint in QIM suggests that quantizer design for QIM involves different condiderations than does quantizer design for rate-distortion performance. Lookabaugh[1] has identified three types of advantages of vector quantizers vs. scalar quantizers. These advantages are called the space-filling, shape, and memory advantages. This paper investigates whether all of these advantages are useful in the context of QIM. QIM performance of various types of quantizers is presented and a heuristic sphere-packing argument is used to show that, in the case of high-resolution quantization and a Gaussian attack channel, only the space-filling advantage is necessary for nearly optimal QIM performance. This is important because relatively simple quantizers are available that do not provide shape and memory gain but do give a space-filling gain.

**Keywords:** digital watermarking, informatioin embedding, quantization index modulation, high resolution quantization theory, capacity

## 1. INTRODUCTION

Digital watermarking is both an active area of research and a growing source of commercial activity. Early research focused upon watermarking as a solution to verifying ownership of images and other media, but watermarking has since found its way into a variety of applications. Some uses for digital watermarking include counterfeit prevention, broadcast monitoring, brand protection, and enabling of content on the internet.

In these and other applications, digital watermarking allows both the unobtrusive embedding of information into media as well as the capability to reliably recover the information, possibly after application specific degradation of the watermarked media. There is a natural tension between requirements of being able to embed a large amount of information, requirements for the embedded information to imperceptibly alter the original source, and requirements of being able to recover the information after severe distortions to the watermarked source. Information theory has been applied to the watermarking problem and has been used to determine the space of requirements over which we can hope to build real systems. Early works which have been found fundamental to watermarking include Gel'fand[2] and Costa.[3] More recently, Chen[4] and Moulin[5] have sought to provide answers to specific watermarking problems.

Quantization index modulation (QIM)[4] was proposed by Chen and Wornell as an efficient method of digital watermarking. QIM is used in the information theoretic proofs of watermarking capacity. This capacity can be demonstrated via random coding using a sequence of vector quantizers of increasing block length, with both channel capacity and optimal rate-distortion performance being reached in the limit of infinite quantizer block length. It is tempting to investigate the utility of QIM in practical watermarking situations as well. It is possible that different considerations should be used in the design of quantizers for QIM than for the design of quantizers for compression. Because the quantized values are not digitally encoded in QIM, the number of reconstruction values in each quantizer is not a design constraint, as it is in the design of a conventional quantizer. The lack of a rate constraint in QIM suggests that quantizer design for QIM involves different considerations than does quantizer design for rate-distortion

E-mail: hbrunk@digimarc.com

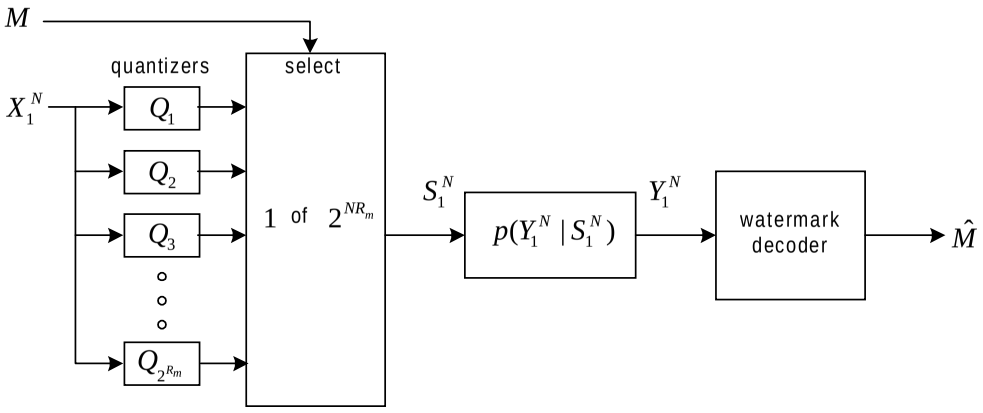


**Figure 1.** Block diagram of QIM watermarking.

performance. Lookabaugh[1] has identified three types of advantages of vector quantizers vs. scalar quantizers. These advantages are called the space-filling, shape, and memory advantages. The space-filling advantage is due to the increasing efficiency with which polytopes can be used to fill space as the dimension increases. The shape advantage involves the marginal distribution of the source and the tendency for the typical source distribution to be concentrated in a small volume in high dimensional space. Finally, the memory advantage exploits dependencies between samples of the source. This paper investigates the space-filling advantage in particular, and provides a heuristic arguement that the space-filling advantage alone is sufficient to achieve watermark capacity in some situations. QIM performance of some relatively simple quantizers with a space filling gain and with a shape gain are compared and found to support the idea that space filling gain is most important for QIM applications of quantizers.

The following section provides some notation and a brief review of QIM. Then section 3 investigates the construction of vector quantizers for QIM which have a space-filling gain and no shape or memory gain. Section 4 presents QIM performance results, and the final section presents conclusions.

## 2. NOTATION AND REVIEW OF QIM

The basic watermarking task is to communicate a watermark message $M$ by modifying a host signal $X_1^N = (X_1, X_2, X_3, \ldots, X_N)$ to produce the watermarked signal $S_1^N$. The watermarked signal then undergoes possible distortions over a watermark signal channel $p(Y_1^N|S_1^N)$ resulting in the received signal $Y_1^N$. The definition of the host signal is quite general. Each sample of the host signal could be a pixel from an image, a sample from a subband decomposition of an image, or an audio sample. In some cases it will be helpful to consider the noise introduced by the watermark signal channel as additive with variance $\sigma^2$. Throughout this paper the host signal, watermarked signal, and received signal will be assumed to be taken from $\mathbb{R}^N$. An embedding rate $R_m$ is used to embed the message in the host signal; accordingly, the message $M \in \{0, 1, 2, \ldots, 2^{NR_m} - 1\}$. It will be assumed that the distortion introduced by the watermark signal channel is independent of the message, and that the host signal is also independent of the message. The goal in the design of the watermarking system is to minimize the error probability in transmitting the message subject to a constraint $D_m$ on the distortion introduced by the watermarked signal. The distortion measure used here will be the per letter squared error:

$$D(S_1^N, X_1^N) = \frac{1}{N} \sum_{i=1}^{N} (S_i - X_i)^2. \tag{1}$$

### 2.1. QIM

Figure 1 shows how QIM can be used to implement the watermarking task. The watermark encoder associates a separate quantizer $\mathcal{Q}$ with each of the possible $2^{NR_m}$ messages. It chooses the quantizer $\mathcal{Q}_M$ corresponding to the

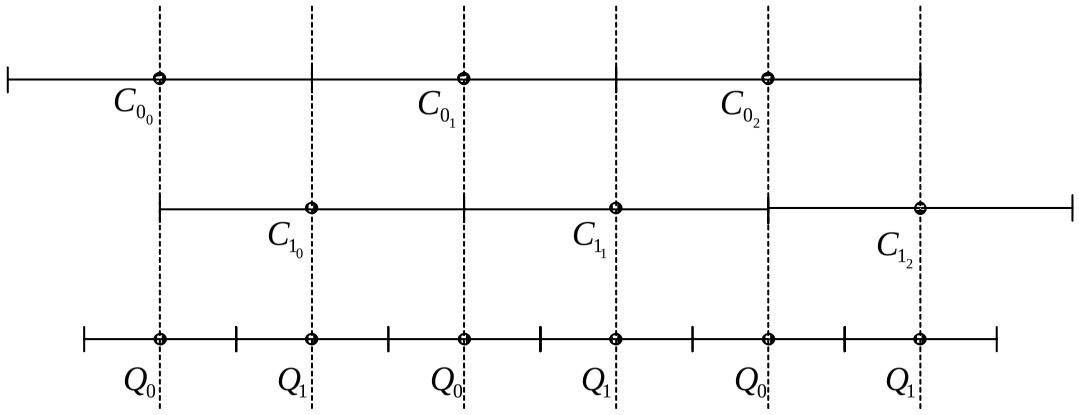


**Figure 2.** Quantizers used for dithered uniform scalar QIM.

input message $M$ and uses this quantizer to quantize the host signal, producing the watermarked signal $S_1^N$. After this signal is passed through the channel, the watermark decoder's task is to estimate the message $M$ based upon the received signal $Y_1^N$. This is done by quantizing $Y_1^N$ using each of the $2^{NR_m}$ quantizers available to the encoder. The corresponding distortion is calculated for each quantizer, and the quantizer resulting in the lowest distortion is assumed to be the one most likely to have been used by the encoder. Then the decoder chooses the message $\hat{M}$, which corresponds to this quantizer, as its estimate of the original message $M$. The watermark encoder can be defined by

$$\mathcal{F}(m, \mathbf{x}) = \mathcal{Q}_{\mathbf{m}}(\mathbf{x}), \tag{2}$$

and the decoder by

$$\mathcal{G}(\mathbf{y}) = \arg \min_{\mathbf{0} \leq \mathbf{m} \leq \mathbf{2^{NR_m}-1}} \mathbf{D}(\mathbf{y}, \mathcal{Q}_{\mathbf{m}}(\mathbf{y})) \tag{3}$$

Figure 2 illustrates dithered uniform scalar QIM, proposed by Chen and Wornell.[4] The message consists of a single bit that causes the encoder to choose between quantizer $Q_0$, shown at top, and quantizer $Q_1$, shown below. The decoder effectively uses the quantizer shown at the bottom of the figure to decide between the two possible message values.

### 2.2. Distortion compensated QIM

Chen and Wornell propose[4] a postprocessing technique for QIM that they call *distortion compensated* QIM. They associate a scale factor $\frac{\Delta}{\alpha}$ with each quantizer and use the parameter $\alpha$ to add part of the quantization error back to the watermarked source:

$$\mathcal{F}_{DC}(m, \mathbf{x}) = \mathcal{Q}_{\mathbf{m}}(\mathbf{x}, \mathbf{\Delta}/\alpha) + (\mathbf{1} - \alpha)[\mathbf{x} - \mathcal{Q}_{\mathbf{m}}(\mathbf{x}, \mathbf{\Delta}/\alpha)]. \tag{4}$$

The parameter $\alpha$ is restricted to $[0, 1]$, ensuring that $D(\mathbf{y}, \mathcal{Q}_{\mathbf{m}}(\mathbf{y}, \mathbf{\Delta}/\alpha))$ is independent of $\alpha$. Chen and Wornell showed that the watermark performance can be optimized by choosing $\alpha$ so that

$$\alpha = \frac{\text{DNR}}{\text{DNR} + 1}, \tag{5}$$

where DNR$= D_m/\sigma^2$.

## 3. DEVELOPMENT OF SPACE-FILLING QUANTIZERS FOR QIM

### 3.1. Two dimensions

Figure 3 shows how dithered uniform scalar QIM can be extended to QIM using optimal uniform two dimensional quantizers. In the one dimensional case, the Figure shows a region of $\mathbb{R}^1$ which is partitioned into two decision

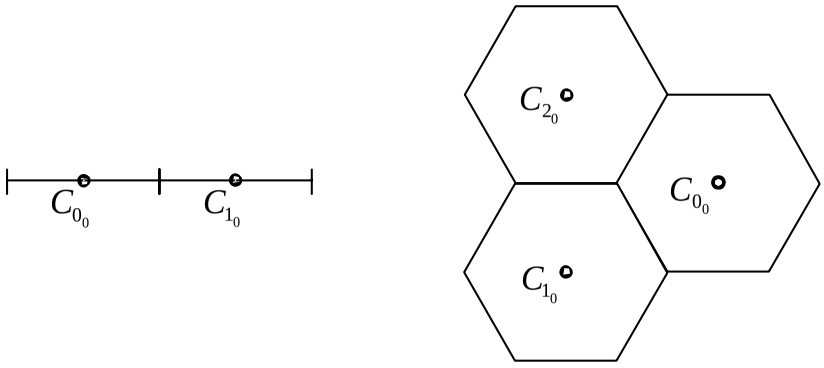


**Figure 3.** Extension of scalar dithered QIM (left) to two dimensions (right).

regions, one for each possible message value. Each decision region contains a code value at its centroid and is a polytope with minimal second moment about its centroid (in this case, trivially, an interval). By tiling $\mathbb{R}^1$ with the two decision regions and codevalues, it is possible to construct the quantizers shown in Figure 2.

The right side of Figure 3 shows a similar diagram in two dimensions. Here the optimal space-filling polytope is the hexagon; correspondingly, a region of $\mathbb{R}^2$ is partitioned into three hexagons, each with a code value at its centroid. Tiling $\mathbb{R}^2$ with this region results in a hexagonal lattice; each lattice point is assigned to one of three quantizers, and each of the three quantizers themselves have code values in a hexagonal sublattice of the original lattice. If code value $C_{0_0}$ is chosen to lie on the origin, then the lattice of all codevalues consists of

$$\{\mathbf{p} : \mathbf{p} = \mathbf{Um}, \mathbf{m} \in \mathbb{Z}^2\}, \text{ where} \mathbf{U} = \begin{bmatrix} \mathbf{0} & \sqrt{\mathbf{3}}\boldsymbol{\Delta} \\ \mathbf{2}\boldsymbol{\Delta} & \boldsymbol{\Delta} \end{bmatrix}. \tag{6}$$

The lattices of the three quantizer's codevalues are given by

$$Q_i = \{\mathbf{p} : \mathbf{p} = \mathbf{Vm} + \mathbf{o_i}, \mathbf{m} \in \mathbb{Z}^2\}, \text{ where} \mathbf{V} = \begin{bmatrix} -\sqrt{\mathbf{3}}\boldsymbol{\Delta} & \sqrt{\mathbf{3}}\boldsymbol{\Delta} \\ \mathbf{3}\boldsymbol{\Delta} & \mathbf{3}\boldsymbol{\Delta} \end{bmatrix}, \mathbf{o_0} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \mathbf{o_1} = \begin{bmatrix} \mathbf{0} \\ \mathbf{2}\boldsymbol{\Delta} \end{bmatrix}, \mathbf{o_2} = \begin{bmatrix} \sqrt{\mathbf{3}}\boldsymbol{\Delta} \\ \boldsymbol{\Delta} \end{bmatrix}. \tag{7}$$

### 3.2. Higher dimensions

To get an idea of the performance which can be achieved from large dimensional quantizers having only a space-filling gain, the previous two dimensional case can be extended to an arbitrarily large number of dimensions $n$. To simplify the analysis, it will be convenient to apply some results of high resolution quantization theory, which assumes that quantizers have sufficient output levels so that the source probability distribution is essentially uniform within each quantization cell. The results of the analysis will be valid to the extent that the watermark distortion constraint $D_m$ is small in comparison to the variance of the watermark source.

Here it will be assumed that the watermark signal channel is an AWGN channel with variance $\sigma^2$. For large $n$ the channel noise will, with high probability, lie within an $n$-dimensional sphere of radius $\sqrt{n\sigma^2}$. Thus if the decision regions of the watermark decoder are spheres of that size centered about the corresponding quantizer codevalues, the probability of watermark decoding error will approach zero with increasing $n$. Maximizing the watermark message rate for a given watermark distortion constraint is equivalent to solving an $n$-dimensional sphere packing problem similar to a problem that can be used to show the capacity of the Gaussian channel.[6]

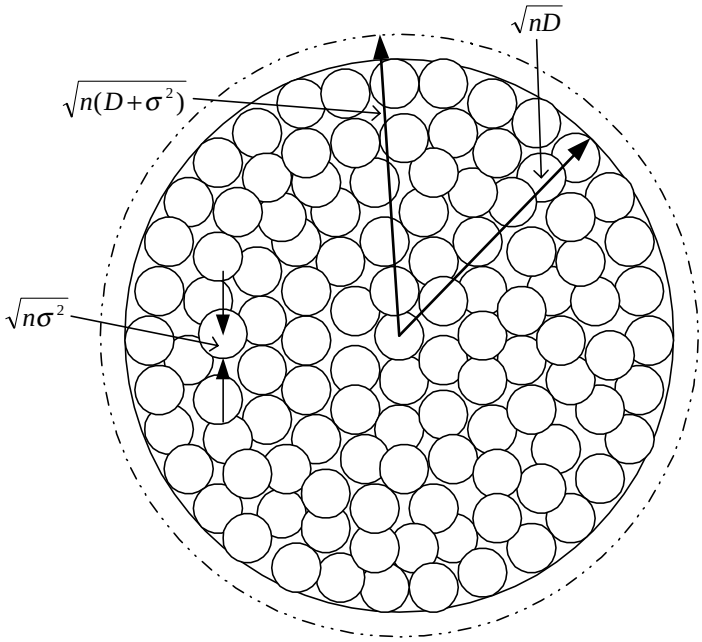


**Figure 4.** Sphere packing arrangement.

This is illustrated in Figure 4, which provides a two dimensional drawing of an $n$-dimensional problem. In the figure one decoding sphere for each possible message value is packed within a larger sphere. $\mathbb{R}^n$ can be tiled with this larger sphere to produce the set of all decoding regions used by the decoder. Each quantizer of the watermark encoder has quantization cells which are spheres of the same size as the larger sphere in Figure 4. These quantization cells are packed tightly to fill $\mathbb{R}^n$. Each of the quantizers is identical except for an offset in $\mathbb{R}^n$ corresponding to the location of the corresponding decoding sphere in Figure 4. The capacity of this idealized QIM is given by the ratio of the volumes of the smaller and larger $n$-dimensional spheres. The larger sphere must have a normalized second moment about its centroid equal to $nD_m$, the distortion constraint. This is given by $\frac{n}{n+2}r^2 = nD_m$. For large $n$, the radius of the larger sphere approaches $\sqrt{nD_m}$. Since the volume of an $n$-dimensional sphere is $k_n r^n$, the maximum number of decoding spheres that can fit inside the larger sphere is:

$$\frac{k_n(nD_m)^{n/2}}{k_n(n\sigma^2)^{n/2}} = 2^{\frac{n}{2}\log\frac{D_m}{\sigma^2}}. \tag{8}$$

This results in an operational watermark capacity for this space-filling QIM of $C = \frac{1}{2}\log\frac{D_m}{\sigma^2}$ bits per sample. If the source $X_1^N$ is iid Gaussian, the watermark capacity is given[4] by

$$C = \frac{1}{2}\log\left(\frac{D_m}{\sigma^2}+1\right) \quad \text{bits per sample.} \tag{9}$$

It is interesting to note that if the outer sphere with radius $\sqrt{n(D_m+\sigma^2)}$ shown in Figure 4 is packed with decoding spheres of radius $\sqrt{n\sigma^2}$, exactly enough decoding spheres fit to reach capacity; a QIM scheme using this larger sphere, however, would not satisfy the distortion constraint $D_m$.

### 3.3. Higher dimension distortion compensated QIM

The capacity of a higher dimensional distortion compensated QIM can be analyzed in a similar manner to the previous case. The larger sphere available for packing of decoding spheres has a radius of

$$r_{large} = \frac{\sqrt{nD_m}}{\alpha} = \frac{(D_m+\sigma^2)}{D_m}\sqrt{nD_m}. \tag{10}$$

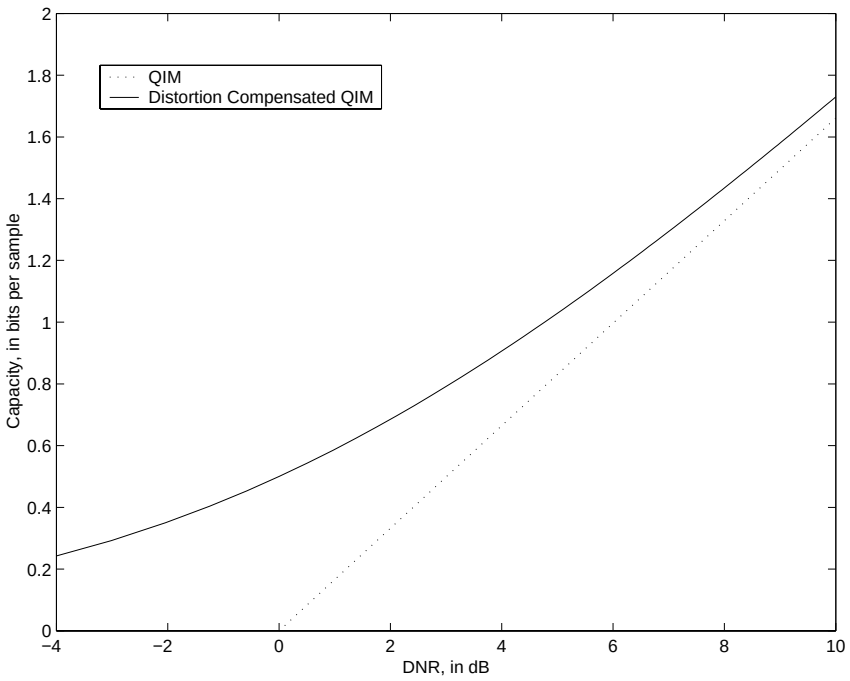


**Figure 5.** Asymptotic (in dimension and resolution) performance of space-filling QIM.

The distortion compensated watermarked signal will lie within a sphere of radius

$$r_{DC} = (1-\alpha) r_{large} = \frac{\sigma^2 \sqrt{n D_m}}{D_m} \tag{11}$$

centered about each quantizer reconstruction value. Each decoding sphere must contain the sum of the distortion compensated watermark signal and the channel noise. Since these two are independent, each decoding sphere must have a radius of

$$r_{decode} = \sqrt{\left(\frac{\sigma^4}{D_m^2}\right) n D_m + n\sigma^2} = \sqrt{n\sigma^2 \left(\frac{\sigma^2 + D_m}{D_m}\right)}. \tag{12}$$

Taking the ratio of the volumes of the two spheres gives a capacity of $C = \frac{1}{2} \log \frac{D_m}{\sigma^2} + 1$ bits per sample, exactly reaching the theoretical capacity.

Figure 5 compares the idealized high dimension space-filling QIM performance with and without distortion compensation.

## 4. PERFORMANCE EXAMPLES

To investigate how vector quantizers with different types of quantization gain perform in a QIM application, QIM using three types of quantizers were simulated. The first quantizer was the dithered scalar quantization proposed by Chen and Wornell. The second was the hexagonal lattice quantizer described in Section 3.1. The final type of quantizer simulated for QIM was a structured vector quantizer (SVQ).[7] In all cases the QIM used a blocklength of $N = 32$ samples in which either one of two messages (for dithered scalar QIM and SVQ QIM) or one of three messages (for 2-D hexagonal lattice QIM) was embedded.

### 4.1. Structured vector quantizers and QIM

SVQ exhibits no space filling gain, but does take advantage of the available shape gain. It is able to achieve the performance of an entropy coded quantizer, but at a fixed bit rate. The first step in designing an $n$-dimensional SVQ is to first design an $N$ level entropy coded scalar quantizer and to calculate the codeword lengths corresponding to each of the quantizer's levels. Then an $n$-dimensional product codebook, containing $N^n$ codevectors of length $n$ is formed. Each vector in the product codebook is assigned a length corresponding to the sum of the lengths of the corresponding scalar quantizer codeword outputs. Then a threshold $T \approx nH(X)$, where $H(X)$ is the entropy of the source is chosen. Finally each codevector with a length greater than $T$ is then removed from the product codebook. Source vectors of length $n$ are then quantized to the closest codevector in the remaining codebook. The number of bits required to represent each codevector is approximately $T$, which is less than $n \log_2 N$. As block length increases, the performance of the SVQ approaches the distortion and rate performance of the original entropy coded scalar quantizer.

For QIM two SVQs were designed, both with block lengths of 32 samples. The scalar quantizers used as the basis of these two SVQs were dithered uniform scalar quantizers. Logarithms of scalar quantizer probabilities were used instead of codeword lengths.

### 4.2. Results

Since the message size is different for the 2-D hexagonal lattice QIM than for the other two types of QIM, they cannot be directly compared by using the resulting probabilities of decoding error. Instead, the delivered channel capacities are derived from the decoding error probabilities and compared. The channel capacities[6] are given by:

$$C_{DITHER} = 1.0 + P_{err} \log (P_{err}) + (1 - P_{err}) \log (1 - P_{err}), \tag{13}$$

$$C_{SVQ} = 1.0 + P_{err} \log (P_{err}) + (1 - P_{err}) \log (1 - P_{err}), \tag{14}$$

$$C_{HEX} = log(3) + P_{err} \log (P_{err}/2) + (1 - P_{err}) \log (1 - P_{err}). \tag{15}$$

Figures 6 and 7 show the results for a unit variance Gaussian source with AWGN of 22dB. As expected, each type of QIM had better performance when distortion compensation was used. For both regular QIM and distortion compensated QIM, the hexagonal lattice based QIM performed the best, followed by the dithered scalar QIM, and the SVQ QIM had the poorest performance.

The loss in performance between the dithered scalar quantizer and SVQ is most likely due to a slight decline in the distortion characteristics of the SVQ, when compared with the original scalar quantizer it was designed from. In a compression application, this decline is more than compensated for by the lower rate requirement of the SVQ compared with the scalar quantizer. Since encoding rate is not important for QIM, this advantage does not accrue to the SVQ QIM.

The hexagonal lattice based QIM performed significantly better than the dithered scalar quantizer; in most cases the capacity of the hexagonal lattice QIM was approximately 150-200% of the capacity of the dithered scalar QIM.

## 5. CONCLUSIONS

The results of the previous section have shown that it is possible to improve QIM performance by using a quantizer which provides a space-filling gain; the results have also shown that quantizers which provide a shape gain do not necessarily improve QIM performance. The capacity arguement presented in section 3 implies that, asymptotically in block length and quantizer resolution, space filling gain alone may be enough to achieve watermarking capacity in distortion compensated QIM. In combination with the experimental results this provides motivation to further investigate quantizers with a space filling gain for use in QIM.

One direction for further study is a more rigorous development of the achievability of capacity by quantizers with only space-filling gain. Another is to investigate the degree to which the watermark capacity is independent of the source distribution.

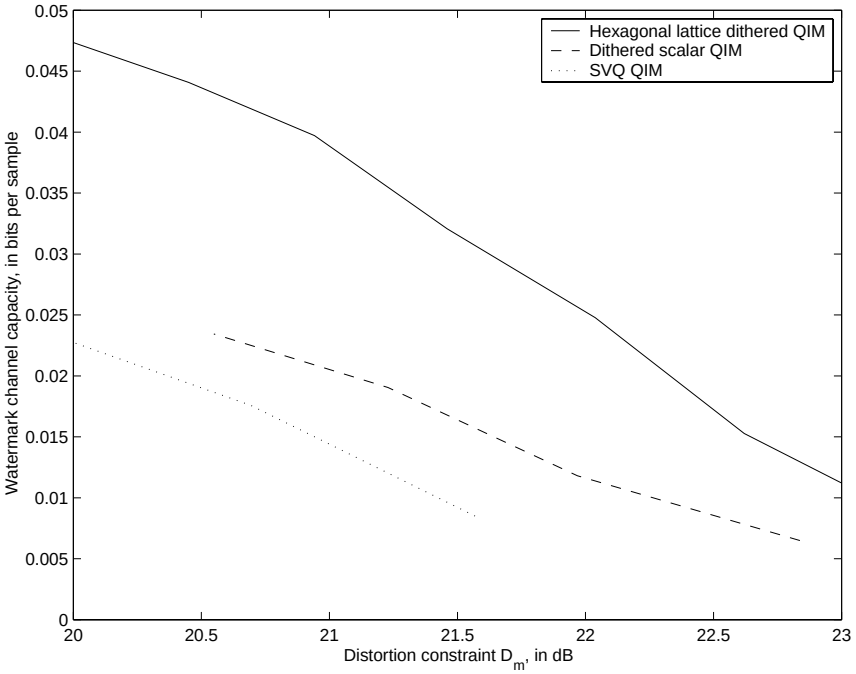


**Figure 6.** Performance of QIM with three different quantizer types.

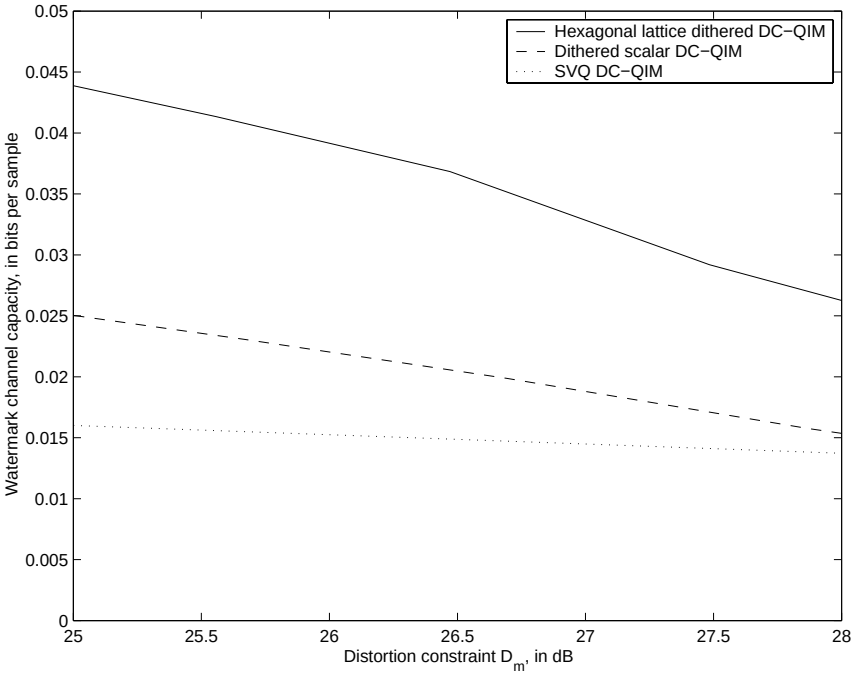


**Figure 7.** Performance of Distortion compensated QIM with three different quantizer types.

## REFERENCES


1. T. D. Lookabaugh and R. M. Gray, "High-resolution quantization theory and the vector quantizer advantage," *IEEE Trans. on Information Theory* **35**, pp. 1020–1033, Sept., 1989.
2. S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory* **9**, pp. 19–31, 1980.
3. M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. on Information Theory* **29**, pp. 439–441, May, 1983.
4. B. Chen and G. W. Wornell, "Preprocessed and postprocessed quantization index modulation methods for digital watermarking," in *Security and Watermarking of Multimedia Contents II*, P. W. Wong and E. J. Delp, eds., *Proc. SPIE* **3971**, pp. 48–59, 2000.
5. P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of watermarking," in *Security and Watermarking of Multimedia Contents II*, *Proc. ICASSP'00* , 2000.
6. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, 1991.
7. R. Laroia and N. Farvardin, "A structured fixed-rate vector quantizer derived from a variable-length scalar quantizer: Part 1 – memoryless sources," *IEEE Trans. on Information Theory* **39**, pp. 851–867, May, 1993.