# Constructing practical Fuzzy Extractors using QIM

Ileana Buhan, Jeroen Doumen, Pieter Hartel, Raymond Veldhuis,
EEMCS Faculty, University of Twente,
{Ileana.Buhan, Jeroen.Doumen, Pieter.Hartel, Raymond.Veldhuis}@utwente.nl

**Abstract**

Fuzzy extractors are a powerful tool to extract randomness from noisy data. A fuzzy extractor can extract randomness only if the source data is discrete while in practice source data is continuous. Using quantizers to transform continuous data into discrete data is a commonly used solution. However, as far as we know no study has been made of the effect of the quantization strategy on the performance of fuzzy extractors. We construct the encoding and the decoding function of a fuzzy extractor using quantization index modulation (QIM) and we express properties of this fuzzy extractor in terms of parameters of the used QIM. We present and analyze an optimal (in the sense of embedding rate) two dimensional construction. Our 6-hexagonal tiling construction offers $(\frac{\log_2 6}{2} - 1) \approx 0.3$ extra bits per dimension of the space compared to the known square quantization based fuzzy extractor.

## 1 Introduction

A fuzzy extractor is a procedure to extract cryptographic keys from noisy data composed typically from two functions. The first is the encoder which takes a noise free feature vector and an independently generated secret and outputs a public sketch. The second is the decoder which takes as input a noisy feature vector and a public sketch and outputs the secret unless the noise exceeds a given threshold. Three parameters are important when extracting secrets from noisy data. Reliability represents the probability of an identification error, embedding rate is the number of bits that are embedded in each component of a feature vector and leakage quantifies the amount of secret information leaked by publishing the sketch. The problem is that to the best of our knowledge this relationship has not been formalized, yet to be able to achieve the best tradeoff between the parameters for a specific application, such a formalization is essential. Once it has been decided which is the most important parameter the formalization helps to find the optimal setting of the other, related parameters.

There is a strong resemblance between *cs*-fuzzy extractors (where the *cs* denotes that we start from a continuous source) and watermarking schemes. During watermark encoding, secret information (the watermark) is embedded into a host signal. Without the host signal it should be hard to find or alter the watermark hidden in the cover. If we consider the feature vector in a biometric system as the host signal and the secret key to be the watermark we observe a similarity between fuzzy extractors and watermarking. However they are not exactly the same: the fuzzy extractor should hide the host signal, while a watermarking scheme should publish a signal close to the host.

Quantization Index Modulation (`QIM`) is a class of data hiding codes used for the construction of optimal watermaking schemes [5]. A `QIM` is an ensemble of quantizers, where the number of quantizers in the ensemble determines the number of distinct possible watermarks. In this context watermarking refers to modulating an index or a sequence of indices with the information that is hidden and then quantizing the space with the indexed quantizer. The quantization function divides a continuum into decision regions and labels each decision region with one reconstruction point. A quantizer is specified by the set of its reconstruction points and by the partition of the continuum into decision regions.

**Contribution.** Our contribution is to show that by using a `QIM` to construct a *cs*-fuzzy extractor it is possible to develop a deep understanding of the tradeoffs between the three properties of a *cs*-fuzzy extractor (i.e reliability, rate and leakage). Our approach is intuitive because it allows modelling the properties of a *cs*-fuzzy extractor in terms of properties of the `QIM` . In our construction reliability is determined by the size and shape of the decision regions. The number of quantizers in the ensemble determines the embedding rate. The distances between neighboring reconstruction points determines the security of a *cs*-fuzzy extractor. Thus optimizing reliability, rate and security can be seen as maximizing the size of the decision regions, maximizing the number of quantizers in the ensemble while keeping the distance between the centroids of different quantizers as small as possible. In this sense an optimal *cs*-fuzzy extractor can be modelled as a dual optimum sphere covering and sphere packing problem. As a result properties of the *cs*-fuzzy extractor can be improved by using higher-dimensional constructions, rather than just stacking one-dimensional constructions as is common in the literature.

The rest of the paper is organized as follows. Related work is discussed in section 2. Section 3 contains notations and fundamental definitions of the `QIM` and fuzzy extractor. In section 4 we construct a *cs*-fuzzy extractor in terms of a `QIM` and study its properties. Section 5 contains two practical constructions for the quantization based *cs*-fuzzy extractor . We compare the properties of these construction with the existing square lattice packing.

## 2 Related work

Our work combines results from the area of data hiding, signal processing and randomness extraction from noisy data.

Uniformly reproducible randomness is the main ingredient of a good cryptographic system. Good quality uniform random sources are rare compared to the more common non-uniform sources. Biometric data is easily accessible, high entropy data. However it is not uniformly distributed and its randomness cannot be exactly reproduced. Depending on the source properties several constructions were proposed. Dodis et al [6] consider discrete distributed noise and propose fuzzy extractors and secure sketches for different error models. These models are not directly applicable to continuously distributed sources. Linnartz et al. [11] construct shielding functions for continuously distributed data and propose a practical construction which can be considered a one-dimensional `QIM`. The same approach is taken by Li et al [10] who propose quantization functions for extending the scope of secure sketches to continuously distributed data. Buhan et al [2] analyze the achievable performances of such constructions given the quality of the source in terms of $\mathrm{FRR}$ and $\mathrm{FAR}$ .

The process of transforming a continuous distribution to a discrete distribution influences the performance of fuzzy extractors and secure sketches. Quantization is the process of replacing analog samples with approximate values taken from a finite set of allowed values. The basic theory of one-dimensional quantization is reviewed by Gersho [7]. The same author investigates [8] the influences of high dimensional quantization on the performance of digital coding for analogue sources. `QIM` constructions are used by Chen and Wornell [4] in the context of watermarking. The same authors introduce dithered quantizers [3]. Moulin and Koetter [12] give an excellent overview of `QIM` in the general context of data hiding. Barron et al [1] develop a geometric interpretation of conflicting requirements between information embedding and source coding with side information.

## 3 Fundamentals

**Notation.** With capital letters we denote random variables, with small letters we denote realizations of random variables, while calligraphic letters are reserved for sets and Greek letters are used to describe properties. Let $\mathcal{U}^k$ be a $k$-dimensional continuous space endowed with a metric $d$ and with background distribution $P_{\mathcal{U}^k}$. Let $X$ be a $k$-dimensional random vector sampled from $\mathcal{U}^k$ with joint density $P_x = p(x_1, x_2, \ldots x_k)$. For optimal encoding-decoding performance during encoding we use the best representative of distribution $P_x$, the estimated mean denoted with $\mathbf{E}[P_x]$. Let $\mathcal{M}$ be a set of labels, and $|\mathcal{M}| = N$. By $P_l$ we denote the uniform distribution of all sequences of length $l$. The min-entropy or the predictability of $X$ denoted by

$\mathbf{H}_\infty(X)$ is defined as minus the logarithm of the most probable element in the distribution: $\mathbf{H}_\infty(X) = -\log_2(\max_x P(X = x))$. The min-entropy represents the number of nearly uniform bits that can be extracted from the variable $X$. By $\mathbf{H}(A|B)$ we denote the conditional entropy which shows the number of bits of randomness remaining in $A$ when $B$ is made public. By $I(A;B)$ we denote the Shannon mutual information. The Kolmogorov distance or *statistical distance* between two probability distributions $A$ and $B$ is defined as: $SD(A,B) = sup_v|Pr(A = v) - Pr(B = v)|$.

**Quantization.** A quantizer is a function $Q : \mathcal{U}^k \rightarrow \mathcal{C}$ that maps each point in $\mathcal{U}^k$ into one of the reconstruction points in a set $\mathcal{C} = \langle c_1, c_2, \ldots \rangle$ where each $c_i \in \mathcal{U}^k$ such that $Q(x) = \mathrm{argmin}_{c_i \in \mathcal{C}}\, d(x, c_i)$ (the function $\mathrm{argmin}$ returns the argument instead of the actual minimum).

An $N$ point $\mathtt{QIM} : \mathcal{U}^k \times \mathcal{M} \rightarrow \mathcal{C}_{\mathtt{QIM}}$ is a set of quantizers $\{Q_1, Q_2, \ldots Q_N\}$, that maps $x \in \mathcal{U}^k$ into one of the reconstruction points of the quantizers in the set. The quantizer is chosen by the input value $m \in \mathcal{M}$ such that $\mathtt{QIM}(x, m) = Q_m(x)$. The set of all reconstruction points is $\mathcal{C}_{\mathtt{QIM}} = \bigcup_{m \in \mathcal{M}} C_m$ where $C_m$ is the set of reconstruction points of quantizer $Q_m$. The Voronoi cells of points in this set are called decision regions $\Omega(c_m^i)$.

A dithered quantizer is a special type of $\mathtt{QIM}$ for which all decision regions of all quantizers are congruent polytopes (generalization of a polygon to higher dimensions). Each quantizer in the ensemble can be obtained by shifting the reconstruction points of any other quantizer in the ensemble. The shifts correspond to dither vectors. The number of dither vectors is equal to the number of quantizers in the ensemble.

We define the minimum distance, $\delta_{\min}$, between centroids of the same quantizer as:

$$\delta_{\min} = \min_{m,n \in \mathcal{M}} \min_{i \in \mathcal{C}_m, j \in \mathcal{C}_n} ||c_m^i - c_n^j||,$$

so spheres with radius $\delta_{\min}/2$ and centers in $\mathcal{C}_{\mathtt{QIM}}$ are disjoint. Let $\zeta_m$ be the smallest radius circle such that circles centered in the centroids of quantizer $Q_m$ with radius $\zeta_m$ cover the universe $U^k$. We define the covering distance $\lambda_{\max}$ as:

$$\lambda_{\max} = \max_{m \in \mathcal{M}} \zeta_m,$$

so spheres with radius $\lambda_{\max}$ and centers in $\mathcal{C}_i$ cover the universe $\mathcal{U}^k$.

**Fuzzy extractors** For modelling the process of randomness extraction from noisy data, Dodis et al. [6] define the notion of a fuzzy extractor. Enrollment is performed by a function $\mathtt{Enc}$, which on input of the noise free biometric $x$ and the binary string $m$, will compute a public string $w$. The binary string $m$ can be extracted from the biometric data itself [13] or can be generated independently [11]. During authentication, the function $\mathtt{Dec}$ takes as input a noisy measurement $x'$ and the public $w$ and it will output

the binary string $m$ if $x$ and $x'$ are close enough. For a discrete source, the formal definition of a fuzzy extractor may be found in Dodis et al [6].

# 4 Constructing *cs*-fuzzy extractor using a `QIM`

In this section we propose a general approach to extract cryptographic keys from noisy data represented in a continuous domain. The first step is to recall the extension of a fuzzy extractor to a *cs*-fuzzy extractor introduced by Buhan et al [2]. We make the assumption that the random binary string $m$ is not extracted from the random vector $x$ but generated independently.

**Definition 1 (*cs*-Fuzzy Extractors)** *A* cs*-fuzzy extractor scheme is a tuple* $(\mathcal{U}^k, \mathcal{M}, \mathcal{W}, \mathtt{Enc}, \mathtt{Dec})$*, where* $\mathtt{Enc} : \mathcal{U}^k \times \mathcal{M} \rightarrow \mathcal{W}$ *is an encoder and* $\mathtt{Dec} : \mathcal{U}^k \times \mathcal{W} \rightarrow \mathcal{M}$ *is a decoder.*
*We say the scheme is* $\rho$-reliable *for the distribution* $X$ *on* $U^k$ *if*

$$P(\mathtt{Dec}(x, \mathtt{Enc}(E[X], m)) = m | X = x) \geq \rho,$$

*for all* $m \in \mathcal{M}$*. We say the scheme is* $\epsilon$-secure *if for any* $x$ *we have that*

$$SD[\langle M, W\rangle, \langle P_{\mathcal{M}}, W\rangle] \leq \epsilon,$$

*where the joint distribution* $\langle M, W\rangle$ *is induced by the tuple* $(m, \mathtt{Enc}(x, m))$ *and* $P_{\mathcal{M}}$ *is uniformly distributed over the labels* $\mathcal{M}$*.*

As discussed in the previous section, we construct a *cs*-fuzzy extractor using a `QIM`. We will assume $\mathcal{U}^k \subseteq \Re^k$. Our construction works as follows:

**Definition 2 (`QIM`-Fuzzy Extractor)** *A* `QIM`*-Fuzzy Extractor is a* cs*-fuzzy extractor where the encoder and decoder are defined as*

$$\mathtt{Enc}(x, m) = \mathtt{QIM}(x, m) - x,$$

*and* `Dec` *is the minimum distance Euclidian decoder:*

$$\mathtt{Dec}(y, w) = \widetilde{Q}(y + w),$$

*where*

$$\widetilde{Q} : \mathcal{U}^k \rightarrow \mathcal{M}, \widetilde{Q}(y) = \underset{m \in \mathcal{M}}{\operatorname{argmin}}\, d(y, C_m).$$

Intuitively, our construction, is a generalization of the scheme of Linnartz and Tuyls [11]. Figures 1 and 2 illustrate the encoding respectively the decoding functions for a `QIM` ensemble of three quantizers $\langle Q_\circ, Q_+, Q_\star\rangle$. During encoding the secret $m \in \{\circ, \star, +\}$ selects a quantizer, say $Q_\circ$. The selected quantizer finds the centroid $Q_\circ(x)$ closest to $x$ and the encoder returns the difference between the two as $w$, with $|w| \leq \lambda_{\max}$. Decoding $w$ and $y$ should return $\circ$ if $y$ is drawn from $P_x$, however this happens only if $y + w$ is close to $Q_\circ(x)$ or in other words if $y + w$ is in the decision region of the chosen centroid (gray area in figure 2). Errors occur if $(y+w) \notin \Omega(Q_\circ(x))$, thus the size of $\Omega(Q_\circ(x))$ parametrized by $\delta_{\min}$ determines the probability of errors.

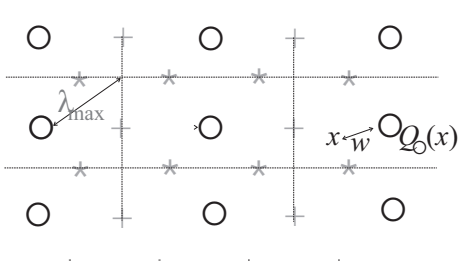


Figure 1: *Encoding with a* QIM

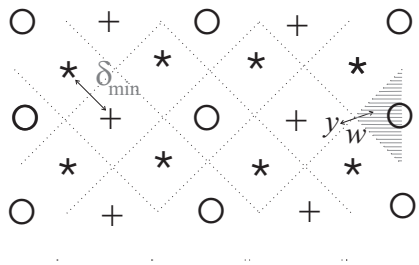


Figure 2: *Decoding with a* QIM

## 4.1 Performance criteria for *cs*-fuzzy extractor

In the following we express the properties of a *cs*-fuzzy extractor in terms of the used quantizers. The proofs for the theorem and lemmas in this section are given in appendix A.

### 4.1.1 Embedding Rate

The embedding rate or simply rate of a *cs*-fuzzy extractor represents the number of bits that can be embedded per dimension. The number of quantizers in the ensemble, $N$ and the dimensionality of the space $k$, gives the embedding rate, written as $l = \frac{\log_2 N}{k}$.

Depending on the method of quantization and the background distribution $P_{\mathcal{U}^k}$, the a-posteriori randomness in $\mathcal{M}$ can change. This remark is rather subtle. During encoding each $m$ is drawn uniformly at random from $\mathcal{M}$. However when the decoder map is published some labels may become more probable then others if $P_{\mathcal{U}^k}$ is not uniform. $\mathbf{H}_\infty(\mathcal{M})$ measures the randomness remaining in $\mathcal{M}$ after publishing the decoding map. In all cases we know that $\mathbf{H}_\infty(\mathcal{M}) \leq \log_2 N$. Buhan et al. [2] show that the min-entropy and the embedding rate determine an upper bound on $\epsilon$. We call effective embedding rate the min-entropy of the label distribution given the background distribution and the QIM construction.

### 4.1.2 Reliability

We link in the following lemma the reliability of a *cs*-fuzzy extractor to the geometric construction of a QIM. More precisely we link reliability to the size and shape of the decision regions.

**Lemma 1 (Bounds on $\rho$)** *The reliability of a* QIM*-Fuzzy Extractor for any random $X \in \mathcal{U}^k$ with joint density function $P_x$ and any secret $m \in \mathcal{M}$ can*

*be bounded as follows:*

$$\begin{aligned}\rho &\leq \int_{\bigcup_i \Omega(c_m^i)} P_x(y - \mathtt{Enc}(E[X], m))dy \\ \rho &\geq \int_{B(E[X], \frac{\delta_{\min}}{2})} P_x(x)dx,\end{aligned}$$

*where $B(x, r)$ is the sphere centered in $x$ with radius $r$.*

### 4.1.3 Security

We require that the *cs*-fuzzy extractor keeps the value of $E[X]$ secret. If compromised, noisy data characterized by $X$ cannot be used for generating secrets. When $X$ is biometric data, leaking the value of $E[X]$ means compromising the privacy of the biometric data. We measure the information leaked about $E[X]$ when publishing the sketch by the Shannon mutual information $I(X; W)$. A good *cs*-fuzzy extractor should leak as little information as possible about $E[X]$. Lemma 2 links $I(X; W)$ to the covering distance. However it was shown by Tuyls et al. [14] that the sketch cannot be made independendent of $X$, thus $I(X; W)$ cannot be zero. Lemma 4 gives a lower bound on the covering distance in terms of minimum distance, number of quantizers and dimension of the space.

**Lemma 2** *For a* `QIM`*-Fuzzy Extractor the amount of information leaked when publishing the encoder output for any random $X$ on $\mathcal{U}^k$ is bounded by above by the covering distance as: $I(X; W) \leq \log_2 \lambda_{\max}$.*

Another problem is leaking information about the secret $m \in \mathcal{M}$. This problem was extensively studied in the context of digital watermarking and information embedding [1, 3, 5], where the solution of dither modulated quantizers surfaced. In this case they will also hide the key perfectly, as shown in the next lemma.

**Lemma 3** *Our* `QIM`*-Fuzzy Extractor construction perfectly hides the key (i.e. $\epsilon = 0$), when the* `QIM` *is a set of dithered quantizers and a uniformly random point $x \in \mathcal{U}^k$ is encoded.*

## 4.2 Optimizing *cs*-fuzzy extractor

Optimizing a fuzzy extractor means increasing the reliability and the embedding rate while keeping the size of the sketch as small as possible. The constraint on both the sketch size and the reliability and the requirement that from any location in the space it should be possible to chose any label is similar to a simultaneous sphere covering and sphere packing problem. The sphere covering is induced by the encoder: from any point in the space it should be possible to find any label at a distance at most $\lambda_{\max}$, so we need a

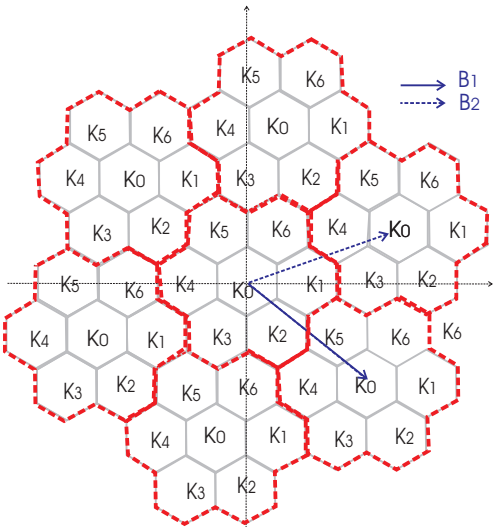


Figure 3: *Decoding of 7-hexagonal tiling*

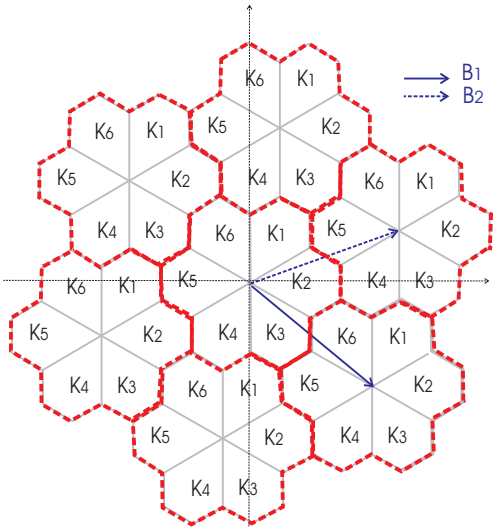


Figure 4: *Decoding of 6-hexagonal tiling*

covering of the space with spheres of radius $\lambda_{\max}$. We have a sphere packing problem at the decoder side since spheres centered in the reconstruction points with radius $\frac{\delta_{\min}}{2}$ cannot overlap. In this setting we obtain an optimum embedding rate by having a dense sphere packing. A good `QIM` construction will maximize both $\delta_{\min}$ and $N$ while keeping $\lambda_{\max}$ to a minimum. These two radii can be linked as follows.

**Lemma 4** *The covering distance of a* `QIM` *ensemble, defined as above is lower bounded by:*

$$\lambda_{\max} \geq \sqrt[k]{N}\frac{\delta_{\min}}{2}$$

*where $k$ represents the dimension of the space and $N$ is the number of different quantizers.*

Assuming a spherically symmetric background distribution (which is weaker then the often made gaussian assumption), there is only so much different equiprobable labels one can achieve:

**Theorem 1 (Optimal high dimensional packing.)** *Assume the background distribution to be spherically symmetrical. If one wants to achieve equiprobable labels given this distribution, the number of labels in a $k$-dimensional* `QIM` *is upper bounded by the kissing number $\tau(k)$.*

Combined with known bounds on the kissing number [9, 15], we arrive at the following somewhat surprising conclusion:

**Corrolary 1** *Assuming a spherically symmetrical distribution on $\mathcal{U}^k$ and equiprobably labels, for a* `QIM`*-Fuzzy Extractor the best rate is attained by quantizing two dimensions at a time, leading to*

$$N(k) = 6^{\lfloor \frac{k}{2} \rfloor} 2^{(k-2\lfloor \frac{k}{2} \rfloor)}$$

*different labels.*

# 5 Practical constructions

In this section we present two constructions for *cs*-fuzzy extractors in two dimensional space using a dithered `QIM`. We choose a hexagonal lattice for the `QIM`, since this gives both a smallest circle covering (for the encoder) and a densest circle packing (for the decoder). The first construction has a rate of $\frac{\log_2 7}{2}$ bits. The scheme is optimal from the reliability point of view. However, in this scheme keys are not equiprobable if the distribution isn't flat enough. The second construction fixes this problem, but has a slightly lower rate of $\frac{\log_2 6}{2}$ bits. Reconstruction points of all quantizers are shifted versions of some base quantizer $Q_0$. A dither vector $\overrightarrow{v_m}$ is defined for each possible $m \in \mathcal{M}$. The *tiling polytope* is the repeated structure in the space that is obtained by decoding to the closest reconstruction points. It follows from the definition that the tiling polytope contains exactly one decision region of each quantizers in the ensemble.

## 5.1 7-Hexagonal Tiling

The first construction is a dithered `QIM` defined as an ensemble of 7 quantizers. Decision regions for this tiling are regular hexagons. A tiling polytope is a union of 7 hexagons. In figures 3, 4 the tiling polytopes are delimited by the red dotted line. The reconstruction points of the base quantizer, $Q_0$ are defined by the lattice spanned by the vectors $\overrightarrow{B_1} = (5, \sqrt{3})q$, $\overrightarrow{B_2} = (4, -2\sqrt{3})q$, where $q$ is the scaling factor of the lattice. In figure 3 these points are labelled $k_0$. The other reconstruction points of quantizers $Q_i, i = 1, \ldots, 6$ are obtained by shifting the base quantizer with the dither vectors $\{\overrightarrow{v_1}, \cdots, \overrightarrow{v_6}\}$ such that $Q_i(x) = Q_0(\overrightarrow{x} + \overrightarrow{v_i})$. The values for these dithered vectors are: $\overrightarrow{v_1} = (2, 0)$, $\overrightarrow{v_2} = (-3, \sqrt{3})$, $\overrightarrow{v_3} = (-1, -\sqrt{3})$, $\overrightarrow{v_4} = (-2, 0)$, $\overrightarrow{v_5} = (3, -\sqrt{3})$ and $\overrightarrow{v_6} = (1, \sqrt{3})$. Encoding and decoding works as in our construction. The decoding is shown graphically in figure 3.

## 5.2 6-Hexagonal Tiling

This construction eliminates the middle hexagon, to make all keys equiprobable (see Theorem 1). The embedding rate is $\frac{\log_2 6}{2}$ bits. The tiling polytope is formed by 6 decision regions and thus there are only 6 dither vectors,

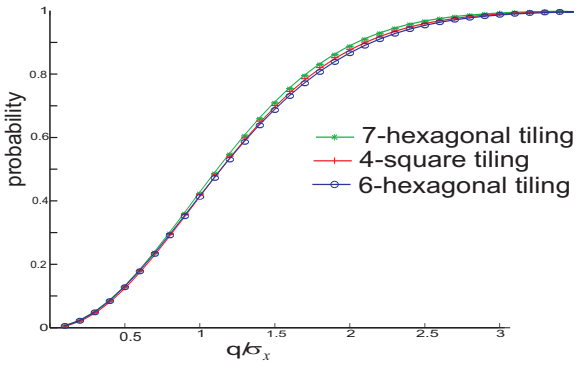


Figure 5: *Reliability three* QIM*-fuzzy extractor constructions*

see figure 4. The same dither vectors, $\{\overrightarrow{v_1}, \cdots, \overrightarrow{v_6}\}$ are used to construct the quantizers, but the basic quantizer $Q_0$ is not used itself. The encoding-decoding functions are defined as in the previous section 5.1.

## 5.3 Performance comparison

We compare the two constructions proposed above, 7-hexagonal tiling figure 3, and 6-hexagonal tiling figure 4, in terms of reliability, embedding rate and leakage with the scalar quantization scheme introduced by Linnartz et al. [11] on each dimension separately (we will refer to this as 4-square tiling).

To perform the comparison we consider identically and independently distributed (i.i.d) Gaussian sources. We assume the background distribution $P_{\mathcal{U}^2}$ to have mean $(0, 0)$ and standard deviation $\sigma_{\mathcal{U}^2}$. Without loss of generality we assume that for any random $X \in \mathcal{U}^2$, the probability distribution $P_x$ has mean $E[X]$ drawn from $P_{\mathcal{U}^2}$, and standard deviation $\sigma_x$.

To evaluate reliability we compute probabilities associated to equal area decision regions, with the reconstruction point centered in the mean $E[X]$ of distribution $P_x$. The curves in figure 5 where obtained by progressively increasing the area of the decision regions. The size of decision region is controlled by the scaling factor of the lattice, $q$. The best performance is obtained by the hexagonal decision regions. This is because the regular hexagon best approximates a circle, the optimal geometrical form. However, differences between reliability of the three QIM *cs*-fuzzy extractor are not spectacular.

We measure the effective embedding rate by calculating the min-entropy given the background distribution. The min-entropy associated to the labels distribution is compared in figure 6 among 7-hexagonal tiling, 6-hexagonal tiling and 4-scalar tiling. Maximizing the min-entropy means minimizing the probability for an attacker to guess the key correctly on her first try. The min-entropy of the 7-hexagonal tiling decreases rapidly with the increase of the lattice scaling factor $q$ relative to $\sigma_{\mathcal{U}^2}$. While for a small lattice scaling fac-

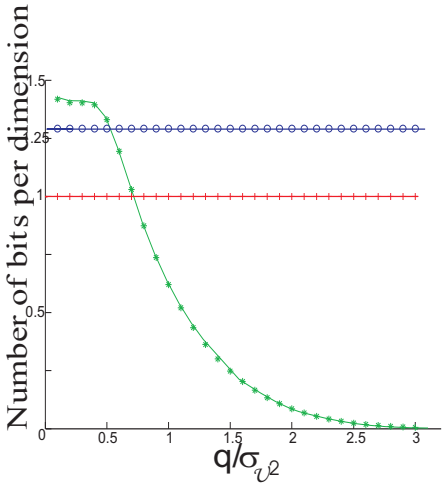


Figure 6: $\mathbf{H}_\infty(\mathcal{M})$ *evaluation for the three* `QIM` *based* cs*-fuzzy extractors constructions*

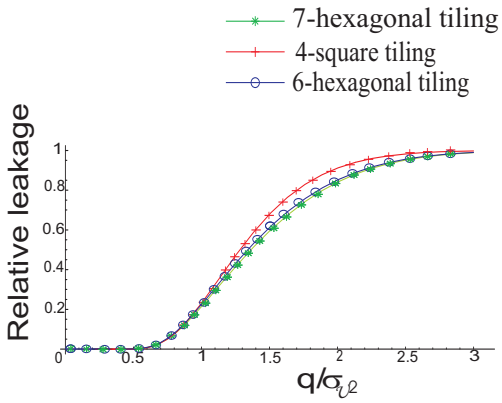


Figure 7: $I(\mathcal{M};\mathcal{W})$ *evaluation for the three* `QIM` *based* cs*-fuzzy extractors*

tor $q$ one can approximate the background distribution as uniform, with the increase in scaling the center hexagon has a substantially higher probability associated and thus one label is more likely then the others. The 6-hexagonal tiling construction eliminates the middle hexagon and as a result all labels become equiprobable, at the cost of a somewhat lower reliability.

Finally, we evaluate the leakage when publishing the helper data. While in the theoretical section we defined security of a `QIM` based *cs*-fuzzy extractor in terms of statistical distance, in practice one learns more from looking at the closely related leakage. Leakage is defined as $I(M;W)$, the mutual information between the key distribution (assumed to be uniform) and the helper data distribution (induced by the key and background distributions). It can be interpreted as the amount of key bits one reveals by publishing the helper data. Unlike in Lemma 3, our $x$ is not distruted uniformly. Since publishing the helper data effectively means that the original $x$ was that vector plus a centroid, one should concentrate on the distribution of $x$. As long as it can be approximated as uniform, the leakage is 0 (as proven in Lemma 3).

# 6 Conclusions

We use `QIM` to construct the encoding and decoding functions of a *cs*-fuzzy extractor . We describe the rate-leakage tradeoff as a simultaneous sphere-packing sphere-covering problem and we show that quantizing dimensions in pairs gives the highest rate. We give two explicit two-dimensional constructions, which perform better then the existing stacked one-dimensional

4-square tiling. We show that 6-hexagonal tiling realizes the optimal two dimensional quantization. Using the 6-hexagonal construction we obtain $k(\frac{log_2 6}{2} - 1)$ more bits compared to the 4-tiling scheme.

# References


[1] RJ Barron, B. Chen, and GW Wornell. The duality between information embedding and source coding with side information and some applications. *Information Theory, IEEE Transactions on*, 49(5):1159–1180, 2003.

[2] I. R. Buhan, J. M. Doumen, P. H. Hartel, and R. N. J. Veldhuis. Fuzzy extractors for continuous distributions. In R. Deng and P. Samarati, editors, *Proceedings of the 2nd ACM Symposium on Information, Computer and Communications Security (ASIACCS), Singapore*, pages 353–355, New York, March 2007. ACM.

[3] B. Chen and G.W. Wornell. Dither modulation: a new approach to digital watermarking and information embedding. *Proceedings of SPIE Vol. 3657: Security and Watermarking of Multimedia Contents*.

[4] B. Chen and GW Wornell. Quantization index modulation: a class of provably good methods fordigital watermarking and information embedding. *Information Theory, IEEE Transactions on*, 47(4):1423–1443, 2001.

[5] B. Chen and G.W. Wornell. Quantization Index Modulation Methods for Digital Watermarking and Information Embedding of Multimedia. *The Journal of VLSI Signal Processing*, 27(1):7–33, 2001.

[6] Y. Dodis, L. Reyzin, and A. Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In Christian Cachin and Jan Camenisch, editors, *EUROCRYPT*, volume 3027 of *Lecture Notes in Computer Science*, pages 523–540. Springer, 2004.

[7] A. Gersho. Principles of quantization. *Circuits and Systems, IEEE Transactions on*, 25(7):427–436, 1978.

[8] A. Gersho. Asymptotically optimal block quantization. *Information Theory, IEEE Transactions on*, 25(4):373–380, 1979.

[9] GA Kabatyanskii and VI Levenshtein. Bounds for packings on a sphere and in space. *Probl. Inform. Transm*, 14(1):1–17, 1978.

[10] Q. Li, Y. Sutcu, and N. Memon. Secure sketch for biometric templates. In *ASIACRYPT*, pages 99–113, 2006.

[11] J.P. Linnartz and P. Tuyls. New shielding functions to enhance privacy and prevent misuse of biometric templates. In Josef Kittler and Mark S. Nixon, editors, *AVBPA*, volume 2688 of *Lecture Notes in Computer Science*, pages 393–402. Springer, 2003.

[12] P. Moulin and R. Koetter. Data-hiding codes. *Proceedings of the IEEE*, 93(12):2083–2126, 2005.

[13] P. Tuyls, A. Akkermans, T. Kevenaar, G. Schrijen, A. Bazen, and R. Veldhuis. Practical biometric authentication with template protection. In Takeo Kanade, Anil K. Jain, and Nalini K. Ratha, editors, *AVBPA*, volume 3546 of *Lecture Notes in Computer Science*, pages 436–446. Springer, 2005.

[14] P. Tuyls and J. Goseling. Capacity and Examples of Template-Protecting Biometric Authentication Systems. *Biometric Authentication Workshop*, pages 158–170, 2004.

[15] K. Zeger and A. Gersho. Number of nearest neighbors in a Euclidean code. *Information Theory, IEEE Transactions on*, 40(5):1647–1649, 1994.

## APPENDIX A

**Lemma 1.** ***Proof*:** We can write the first relation as:

$$P(\texttt{Dec}(x', \texttt{Enc}(x, m)) = m) = \sum_{i \in I} \int_{\Omega(c_m^i)} P(x)dx$$

where $x' \in X$. We have that $(\forall)m \in \mathcal{M}$:

$$\rho \leq \sum_{i \in I} \int_{\Omega(c_m^i) P(x)dx}$$

We have equality when probability associated to the sum of all decision regions of all quantizers is equal. In other words if probability associated to all codewords is equal.
The second relation is straightforward. Reliability is at least the sum of all balls of radius $\frac{\delta_{\min}}{2}$ inscribed in the decision regions. Thus the size of this ball determines reliability. The shape of the decision region that inscribes the ball is important as well.

**Lemma 2.** ***Proof*:**

$$I(X; \mathcal{W}) = H(X) - H(X|\mathcal{W}) \leq H(X) - H(X) + \log_2 |\mathcal{W}| = \log_2 \lambda_{\max}$$

**Lemma 3.** ***Proof*:** The proof is immediate due to the property of the dither-modulated quantizers to make the published sketch independent of the embedded secret. As a consequence no information is leaked as long as $P_{\mathcal{X}}$ is uniform. Since the `QIM` is dithered, all individual quantizers in the ensemble are just $v_i$ translations of each other. In particular, we have that $\texttt{Enc}(x, m_i) = \texttt{Enc}(x + \delta_j - \delta_i, m_j)$. As long as $P_x$ is distributed uniformly, the output of the encoder function is independent of the used label, and hence $\epsilon = 0$.

**Lemma 4.** ***Proof*:** As noted above, all spheres with radius $\delta_{\min}/2$ centered in the centroids of the whole ensemble are disjoint. Each collection of spheres with radius $\lambda_{\max}$ centered in the centroids of an individual quantizer gives a covering of the space $\mathcal{U}^k$. Therefore, a sphere with radius $\lambda_{\max}$, regardless of its center, contains at least the volume of $N$ disjoint spheres of radius $\delta_{\min}/2$, one for each quantizer in the ensemble. Comparing the volumes, we have that

$$s_k \lambda_{\max}^k \geq s_k N (\frac{\delta_{\min}}{2})^k$$

where $s_k$ is a constant only depending on the dimension.

**Theorem 1.** ***Proof sketch*:** Our reliability constraints imply that we use a densest sphere packing for the decoder. If we want to achieve a maximum number of equiprobable labels (without sacrificing too much reliability), the best construction is to center the distribution in one sphere, and give each touching sphere a different label. Note that disregarding this "first" ring of spheres doesn't help to embed more labels in general, since there generally are multiple distances with only $\tau(k)$ different spheres at that distance.

**Corollary 1.** ***Proof*:** Known upper bounds on the kissing number in $k$ dimensions [9] state $\tau(k) \leq 2^{0.401k(1+o(1))}$. This means that $N(k) \geq \tau(k)$ in all dimensions, since $N(k) \approx 2^{1.3k}$ and small dimensions can easily be verified by hand. Also note that $N(k_1 + k_2) \leq N(k_1)N(k_2)$. Thus quantizing dimensions pairwise gives the biggest number of equiprobable keys for any spherically symmetric distribution.