de Groot *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:54
DOI 10.1186/s13634-016-0353-z

EURASIP Journal on Advances in Signal Processing

RESEARCH Open Access

Quantization in zero leakage helper data schemes

Joep de Groot[1,4], Boris Škorić[2], Niels de Vreede[3] and Jean-Paul Linnartz[1*]

Abstract

A helper data scheme (HDS) is a cryptographic primitive that extracts a high-entropy noise-free string from noisy data. Helper data schemes are used for preserving privacy in biometric databases and for physical unclonable functions. HDSs are known for the guided quantization of continuous-valued sources as well as for repairing errors in discrete-valued (digitized) sources. We refine the theory of helper data schemes with the zero leakage (ZL) property, i.e., the mutual information between the helper data and the extracted secret is zero. We focus on quantization and prove that ZL necessitates particular properties of the helper data generating function: (1) the existence of "sibling points", enrollment values that lead to the same helper data but different secrets and (2) quantile helper data.
We present an optimal reconstruction algorithm for our ZL scheme, that not only minimizes the reconstruction error rate but also yields a very efficient implementation of the verification. We compare the error rate to schemes that do not have the ZL property.

Keywords: Biometrics, Fuzzy extractor, Helper data, Privacy, Secrecy leakage, Secure sketch

1 Introduction

1.1 Biometric authentication: the noise problem

Biometrics have become a popular solution for authentication or identification, mainly because of their convenience. A biometric feature cannot be forgotten (like a password) or lost (like a token). Nowadays identity documents such as passports nearly always include biometric features extracted from fingerprints, faces, or irises. Governments store biometric data for forensic investigations. Some laptops and smart phones authenticate users by means of biometrics.

Strictly speaking, biometrics are not secret. In fact, fingerprints can be found on many objects. It is hard to prevent one's face or iris from being photographed. However, storing biometric features in an unprotected, open database. Introduces both security and privacy risks. Security risks include the production of fake biometrics from the stored data, e.g., rubber fingers [1, 2]. These fake biometrics can be used to obtain unauthorized access to services, to gain confidential information or to leave fake evidence at crime scenes. We also mention two privacy risks. (1) Some biometrics are known to reveal diseases and disorders of the user. (2) Unprotected storage allows for cross-matching between databases.

These security and privacy problems cannot be solved by simply encrypting the database. It would not prevent *insider attacks*, i.e., attacks or misuse by people who are authorized to access the database. As they legally possess the decryption keys, database encryption does not stop them.

The problem of storing biometrics is very similar to the problem of storing passwords. The standard solution is to store *hashed* passwords. Cryptographic hash functions are one-way functions, i.e., inverting them to calculate a secret password from a public hash value is computationally infeasible. Even inside attackers who have access to all the hashed passwords cannot deduce the user passwords from them.

Straightforward application of this hashing method to biometrics does not work for biometrics, however. Biometric measurements are noisy, which causes (small) differences between the digital representation of the enrollment measurement and the digitized measurement during verification. Particularly if the biometric value lies near a quantization boundary, a small amount of noise can

*Correspondence: j.p.linnartz@tue.nl
[1] Signal Processing Systems group, Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB, Eindhoven, The Netherlands
Full list of author information is available at the end of the article

flip the discretized value and trigger an avalanche of bit flips at the output of the hash.

1.2 Helper data schemes

The solution to the noise problem is to use a helper data scheme (HDS) [3, 4]. A HDS consists of two algorithms, `Gen` and `Rep`. In the enrollment phase, the `Gen` algorithm takes a noisy (biometric) value as input and generates not only a secret but also public data called *helper data*. The `Rep` algorithm is used in the verification phase. It has two inputs: the helper data and a fresh noisy (biometric) value obtained from the same source. The `Rep` algorithm outputs an estimator for the secret that was generated by `Gen`.

The helper data makes it possible to derive the (discrete) secret reproducibly from noisy measurements, i.e., to perform error correction, while not revealing too much information about the enrollment measurement. The noise-resistant secret can be hashed as in the password protection scheme.

1.3 A two-stage approach

We describe a commonly adopted two-stage approach for real-valued sources, as for instance presented in ([5], Chap. 16). The main idea is as follows. A first-stage HDS performs quantization (discretization) of the real-valued input. Helper data is applied in the "analog" domain, i.e., before quantization. Typically, the helper data consists of a 'pointer' to the center of a quantization interval. The quantization intervals can be chosen at will, which allows for optimizations of various sorts [6–8].

After the first stage, there is typically still some noise in the quantized output. A second-stage HDS employs digital error correction techniques, for instance the code offset method (also known as Fuzzy Commitment) [3, 9] or a variant thereof [10, 11].

Such a two-stage approach is also common practice in communication systems that suffer from unreliable (wireless) channels: the signal conditioning prior to the quantization involves optimization of signal constellations and multidimensional transforms. The discrete mathematical operations, such as error correction decoding, are known to be effective only for sufficiently error-free signals. According to the asymptotic Elias bound ([12], Chap. 17), at bit error probabilities above 10 % one cannot achieve code rates better than 0.5. Similarly, in biometric authentication, optimization of the first stage appears essential to achieve adequate system performance. The design of the first stage is the prime motivation, and key contribution, of this paper.

Figure 1 shows the data flow and processing steps in the two-stage helper data scheme. In a preparation phase preceding all enrollments, the population's biometrics are studied and a transform is derived (using well known techniques such as principal component analysis or linear discriminant analysis [13]). The transform splits the biometric vector $\underline{x}$ into scalar components $(x_i)_{i=1}^M$. We will refer to these components x_i as features. The transform ensures that they are mutually independent, or nearly so.

At enrollment, a person's biometric $\underline{x}$ is obtained. The transform is applied, yielding features $(x_i)_{i=1}^M$. The `Gen` algorithm of the first-stage HDS is applied to each feature independently. This gives continuous helper data $(w_i)_{i=1}^M$ and short secret strings $s_1, \ldots, s_M$ which may or may not have equal length, depending on the signal-to-noise ratio of the features. All these secrets are combined into one high-entropy secret k, e.g., by concatenating them after Gray-coding. Biometric features are subject to noise, which will lead to some errors in the reproduced secret $\hat{k}$; hence, a second stage of error correction is done with another HDS. The output of the second-stage `Gen` algorithm is discrete helper data r and a practically noiseless string c. The hash $h(c\|z)$ is stored in the enrollment database, along with the helper data $(w_i)_{i=1}^M$ and r. Here ,z is salt, a random string to prevent easy cross-matching.

In the authentication phase, a fresh biometric measurement $\underline{y}$ is obtained and split into components $(y_i)_{i=1}^M$. For each i independently, the estimator $\hat{s}_i$ is computed from y_i and w_i. The $\hat{s}_i$ are combined into an estimator $\hat{k}$, which is then input into the 2nd-stage HDS reconstruction together with r. The result is an estimator $\hat{c}$. Finally, $h(\hat{c}\|z)$ is compared with the stored hash $h(c\|z)$.

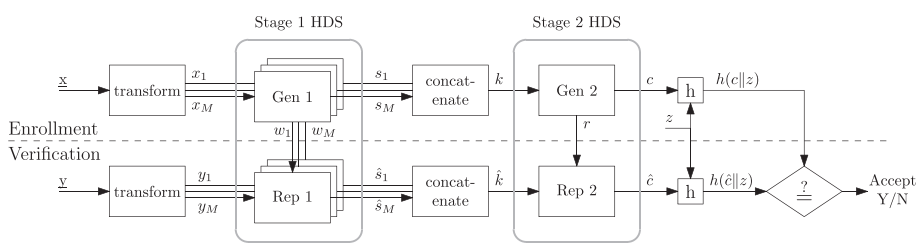

Fig. 1 Common steps in a privacy-preserving biometric verification scheme

1.4 Fuzzy extractors and secure sketches

Special algorithms have been developed for HDSs [4, 6, 8, 9]: Fuzzy extractors (FE) and secure sketches (SS). The FE and SS are special cases of the general HDS concept. They have different requirements,

- *Fuzzy extractor*
  The probability distribution of s given w has to be (nearly) uniform.
- *Secure sketch*
  s given w must have high entropy, but does not have to be uniform. Typically, s is equal to (a discretized version of) x.

The FE is typically used for the extraction of cryptographic keys from noisy sources such as physical unclonable functions (PUFs) [14–16]. Some fixed quantization schemes support the use of a fuzzy extractor, provided that the quantization intervals can be chosen such that each secret s is equiprobable, as in [17].

The SS is very well suited to the biometrics scenario described above.

1.5 Security and privacy

In the HDS context, the main privacy question is how much information, and *which* information, about the biometric $\underline{\mathrm{x}}$ is leaked by the helper data. Ideally, the helper data would contain just enough information to enable the error correction. Roughly speaking, this means that the vector $\underline{\mathrm{w}} = (w_i)_{i=1}^{M}$ consists of the noisy "least significant bits" of $\underline{\mathrm{x}}$, which typically do not reveal sensitive information since they are noisy anyway. In order to make this kind of intuitive statement more precise, one studies the information-theoretic properties of HDSs. In the system as sketched in Fig. 1, the mutual information[1] $I(C; \underline{\mathrm{W}}, R)$ is of particular interest: it measures the leakage about the string c caused by the fact that the attacker observes $\underline{\mathrm{w}}$ and r. By properly separating the "most significant digits" of $\underline{\mathrm{x}}$ from the "least significant digits", it is possible to achieve $I(C; \underline{\mathrm{W}}, R) = 0$. We call this zero secrecy leakage or, more compactly, zero leakage (ZL).[2] HDSs with the ZL property are very interesting for quantifying privacy guarantees: if a privacy-sensitive piece of a biometric is fully contained in c, and not in $(\underline{\mathrm{w}}, r)$, then a ZL HDS based database reveals *absolutely nothing* about that piece.[3]

We will focus in particular on schemes whose first stage has the ZL property for each feature separately: $I(S_i; W_i) = 0$. If the transform in Fig. 1 yields independent features, then automatically $I(S_j; W_i) = 0$ for all i, j, and the whole first stage has the ZL property.

1.6 Contributions and organization of this paper

In this paper, we zoom in on the first-stage HDS and focus on the ZL property in particular. Our aim is to minimize reconstruction errors in ZL HDSs that have scalar input $x \in \mathbb{R}$. We treat the helper data as being real-valued, $w \in \mathbb{R}$, though of course w is in practice stored as a finite-precision value.

- We show that the ZL constraint for continuous helper data necessitates the existence of "Sibling Points", points x that correspond to different s but give rise to the same helper data w.
- We prove that the ZL constraint for $x \in \mathbb{R}$ implies "quantile" helper data. This holds for uniformly distributed s as well as for non-uniform s. Thus, we identify a simple quantile construction as being the generic ZL scheme for all HDS types, including the FE and SS as special cases. It turns out that the continuum limit of a FE scheme of Verbitskiy et al. [7] precisely corresponds to our quantile HDS.
- We derive a reconstruction algorithm for the quantile ZL FE that minimizes the reconstruction errors. It amounts to using a set of optimized threshold values, and is very suitable for low-footprint implementation.
- We analyze, in an all-Gaussian example, the performance (in terms of reconstruction error rate) of our ZL FE combined with the optimal reconstruction algorithm. We compare this scheme to fixed quantization and a likelihood-based classifier. It turns out that our error rate is better than that of fixed quantization, and not much worse than that of the likelihood-based classifier.

The organization of this paper is as follows. Section 2 discusses quantization techniques. After some preliminaries (Section 3), the sibling points and the quantile helper data are treated in Section 4. Section 5 discusses the optimal reconstruction thresholds. The performance analysis in the Gaussian model is presented in Section 6.

2 Related work on biometric quantization

Many biometric parameters can be converted by a principal component analysis (PCA) into a vector of (near)independent components [18]. For this reason, most papers on helper data in the analog domain can restrict themselves to a one-dimensional quantization, e.g., [4, 6, 18]. Yet, the quantization strategies differ, as we will review below. Figure 2 shows the probability density function (PDF) of the measurement y in the verification phase and how the choice of quantization regions in the verification phase affects the probability of erroneous reconstruction (shaded area) in the various schemes.

2.1 Fixed quantization (FQ)

The simplest form of quantization applies a uniform, fixed quantization grid during both enrollment and verification. An example for $N = 4$ quantization regions is depicted

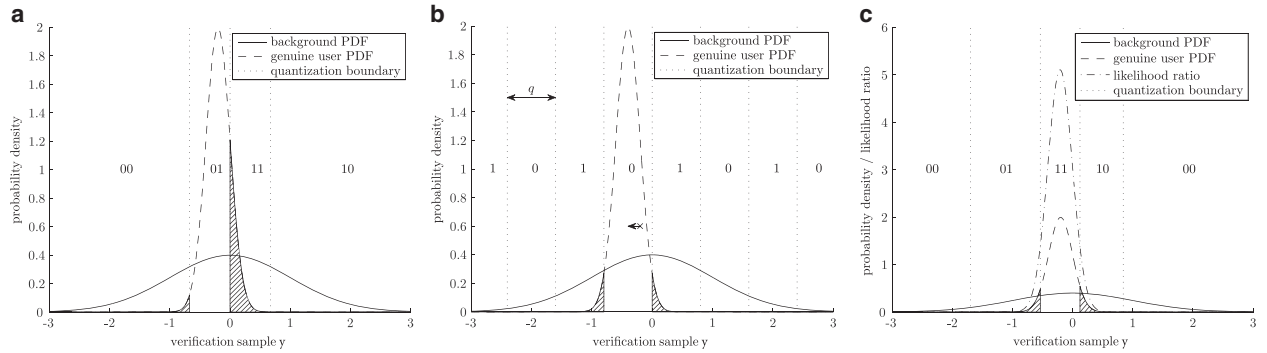

Fig. 2 Examples of adaptating the genuine user PDF in the verification phase. FQ does not translate the PDF; QIM centers the PDF on a quantization interval; LQ uses a likelihood ratio to adjust the quantization regions. **a** Fixed equiprobable quantization. **b** Quantization Index Modulation. **c** Multi-bits based on likelihood ratio [6]

in Fig. 2a. An unfavorably located genuine user pdf, near a quantization boundary, can cause a high reconstruction error.

The inherently large error probability can be mitigated by "reliable component" selection [17]. Only components x_i far away from a boundary are selected; the rest are discarded. The indices of the reliable components constitute the helper data. Such a scheme is very inefficient, as it wastes resources: features that are unfavorably located w.r.t. the quantization grid, but nonetheless carry information, are eliminated. Furthermore, the helper data leaks information about the biometric, since the intervals have unequal width and therefore unequal probabilities of producing reliable components [19].

2.2 Quantization index modulation (QIM)

QIM borrows principles from digital watermarking [20] and writing on dirty paper [21]. QIM has quantization intervals alternatingly labeled with 's'0" and "1" as the values for the secret s. The helper data w is constructed as the distance from x to the middle of a quantization interval; adding w to y then offsets the pdf so that the pdf is centered on the interval (Fig. 2b), yielding a significantly lower reconstruction error probability than FQ.

The freedom to choose quantization step sizes allows for a trade-off between reconstruction performance and leakage [4]. The alternating labeling was adopted to reduce leakage but sacrifices a large part of the source's entropy.

2.3 Likelihood-based quantization (LQ)

At enrollment, the LQ scheme [6] allocates N quantization regions as follows. The first two boundaries are chosen such that they yield the same probability of y given x, and at the same time enclose a probability mass $1/N$ on the background distribution (the whole population's distribution). Subsequent quantization intervals are chosen contiguous to the first and again enclose a $1/N$ probability mass. Finally, the probability mass in the tails of the background distribution is added up as a wrap-around interval, which also holds a probability mass of $1/N$. Since the quantization boundaries are at fixed probability mass intervals, it suffices to communicate a single boundary t as helper data to the verification phase.

In LQ, the secret s is not equiprobable. The error rates are low, but the revealed t leaks information about s.

2.4 Dynamic detection-rate-based bit allocation

In [22], Lim et al. proposed dynamic genuine interval search (DGIS) as an improvement of the bit allocation scheme of Chen et al. [23]. The resulting scheme has some similarity to our approach in that they both determine discretization intervals per user and store these intervals as helper data. However, their scheme is motivated solely by optimization of the detection rate, whereas in our scheme the optimization is subject to the zero leakage restriction. Applying the DGIS method introduces some additional leakage to the underlying bit allocation scheme. Furthermore, DGIS performs its search for the optimal discretization intervals using a sliding window algorithm, which in general will not succeed in finding the exact optimum. In contrast, in our scheme, we analytically derive the optimal solution from the background distribution.

3 Preliminaries

3.1 Notation

Random variables are denoted with capital letters and their realizations in lowercase. The notation $\mathbb{E}$ stands for expectation. Sets are written in calligraphic font. We zoom in on the one-dimensional first-stage HDS in Fig. 1. For brevity of notation the index $i \in \{1, \ldots, M\}$ on x_i, w_i, s_i, y_i and $\hat{s}_i$ will be omitted.

The probability density function (PDF) or probability mass function (PMF) of a random variable A is denoted

as f_A, and the cumulative distribution function (CDF) as F_A. We consider $X \in \mathbb{R}$. The helper data is considered continuous, $W \in \mathcal{W} \subset \mathbb{R}$. Without loss of generality we fix $\mathcal{W} = [0, 1)$. The secret S is an integer in the range $\mathcal{S} = \{0, \ldots, N-1\}$, where N is a system design choice, typically chosen according to the signal to noise ratio of the biometric feature. The helper data is computed from X using a function g, i.e., $W = g(X)$. Similarly, we define a quantization function Q such that $S = Q(X)$. The enrollment part of the HDS is given by the pair Q, g. We define quantization regions as follows,

$$A_s = \{x \in \mathbb{R} : Q(x) = s\}. \tag{1}$$

The quantization regions are non-overlapping and cover the complete feature space, hence form a partitioning:

$$A_s \cap A_t = \emptyset \quad \text{for } s \neq t\,; \qquad \bigcup_{s \in \mathcal{S}} A_s = \mathbb{R}. \tag{2}$$

We consider only quantization regions that are contiguous, i.e., for all s it holds that A_s is a simple interval. In Section 5.3, we will see that many other choices may work equally well, *but not better*; our preference for contiguous A_s regions is tantamount to choosing the simplest element Q out of a whole equivalence class of quantization functions that lead to the same HDS performance. We define quantization boundaries $q_s = \inf A_s$. Without loss of generality, we choose Q to be a monotonically increasing function. This gives $\sup A_s = q_{s+1}$. An overview of the quantization regions and boundaries is depicted in Fig. 3.

In a generic HDS, the probabilities $\mathbb{P}[S = s]$ can be different for each s. We will use shorthand notation

$$\mathbb{P}[S = s] = p_s > 0. \tag{3}$$

The quantization boundaries are given by

$$q_s = F_X^{-1}\left(\sum_{t=0}^{s-1} p_t\right), \tag{4}$$

where F_X^{-1} is the inverse CDF. For a Fuzzy extractor, one requires $p_s = 1/N$ for all s, in which case (4) simplifies to

$$q_s^{\mathrm{FE}} = F_X^{-1}\left(\frac{s}{N}\right). \tag{5}$$

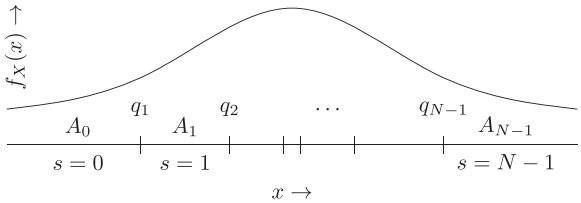

Fig. 3 Quantization regions A_s and boundaries q_s. The locations of the quantization boundaries are based on the distribution of x, such that secret s occurs with probability p_s

3.2 Zero leakage

We will work with a definition of the zeroleakage property that is a bit stricter than the usual formulation [7], which pertains to mutual information. This is necessary in order to avoid problems caused by the fact that W is a continuum variable (e.g., pathological cases where some property does not hold on measure-zero subsets of $\mathcal{W}$),

Definition 3.1. We call a helper data scheme ***Zero Leakage*** *if and only if*

$$\forall_{\mathcal{V} \subseteq \mathcal{W}} \quad \mathbb{P}[S = s | W \in \mathcal{V}] = \mathbb{P}[S = s]\,. \tag{6}$$

In words, we define the ZL property as independence between S and W. Knowledge about W has no effect on the adversary's uncertainty about S. ZL implies $I(S; W) = 0$ or, equivalently, $H(S|W) = H(S)$. Here H stands for Shannon entropy, and I for mutual information (see, e.g., ([24], Eq. (2.35)–(2.39)).

3.3 Noise model

It is common to assume a noise model in which the enrollment measurement x and verification measurement y are both derived from a hidden 'true' biometric value z, i.e., $X = Z + N_{\mathrm{e}}$ and $Y = Z + N_{\mathrm{v}}$, where N_{e} stands for the noise in the enrollment measurement and N_{v} for the noise in the verification measurement. It is assumed that N_{e} and N_{v} are mutually independent and independent of X and Y. The $N_{\mathrm{e}}, N_{\mathrm{v}}$ have zero mean and variance $\sigma_{\mathrm{e}}^2, \sigma_{\mathrm{v}}^2$ respectively. The variance of z is denoted as σ_Z^2. This is a very generic model. It allows for various special cases such as noiseless enrollment, equal noise at enrollment, and verification, etc.

It is readily seen that the variance of X and Y is given by $\sigma_X^2 = \sigma_Z^2 + \sigma_{\mathrm{e}}^2$ and $\sigma_Y^2 = \sigma_Z^2 + \sigma_{\mathrm{v}}^2$. The correlation coefficient ρ between X and Y is defined as $\rho = (\mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y])/(\sigma_X \sigma_Y)$ and it can be expressed as $\rho = \sigma_Z^2/(\sigma_X \sigma_Y)$.

For zero-mean X, Y, it is possible to write

$$Y = \lambda X + R, \quad \text{with } \lambda = \frac{\sigma_Y}{\sigma_X}\rho \quad \text{and } \mathrm{Var}(R) = \sigma^2 \stackrel{\mathrm{def}}{=} \sigma_Y^2(1 - \rho^2), \tag{7}$$

where R is zero-mean noise. This way of expressing Y is motivated by the first and second order statistics, i.e., the variance of Y is $\lambda^2 \sigma_X^2 + \sigma^2 = \sigma_Y^2$, and the correlation between X and Y is $\mathbb{E}[XY]/(\sigma_X \sigma_Y) = \lambda \sigma_X^2/(\sigma_X \sigma_Y) = \rho$.

In the case of Gaussian $Z, N_{\mathrm{e}}, N_{\mathrm{v}}$, the X and Y are Gaussian, and the noise R is Gaussian as well.

From (7), it follows that the PDF of y given x (i.e., the noise between enrollment and verification) is centered on λx and has variance σ^2. The parameter λ is called the attenuation parameter. In the "identical conditions" case $\sigma_{\mathrm{e}} = \sigma_{\mathrm{v}}$ it holds that $\lambda = \rho = \sigma_Z^2/\left(\sigma_Z^2 + \sigma_{\mathrm{v}}^2\right)$. In the

"noiseless enrollment" case $\sigma_e = 0$ we have $\lambda = 1$ and $\rho = \sigma_Z / \sqrt{\sigma_Z^2 + \sigma_v^2}$.

We will adopt expression (7) as our noise model.

In Section 5.1, we will be considering a class of noise distributions that we call *symmetric fading noise*.

Definition 3.2. Let X be the enrollment measurement and let Y be the verification measurement, where we adopt be model of Eq. (7). Let $f_{Y|X}$ denote the probability density function of Y given X. The noise is called symmetric fading noise *if for all x, y_1, y_2 it holds that*

$$|y_1 - \lambda x| < |y_2 - \lambda x| \implies f_{Y|X}(y_1|x) > f_{Y|X}(y_2|x). \quad (8)$$

Equation (8) reflects the property that small noise excursions are more likely than large ones, and that the sign of the noise is equally likely to be positive or negative. Gaussian noise is an example of symmetric fading noise.

4 Zero leakage: quantile helper data

In Section 4.1, we present a chain of arguments from which we conclude that, for ZL helper data, it is sufficient to consider only functions g with the following properties: (1) covering $\mathcal{W}$ on each quantization interval (surjective); (2) monotonically increasing on each quantization interval. This is then used in Section 4.2 to derive the main result, Theorem 4.8: Zero Leakage is equivalent to having helper data obeying a specific quantile rule. This rule makes it possible to construct a very simple ZL HDS which is entirely generic.

4.1 Why it is sufficient to consider monotonically increasing surjective functions *g*

The reasoning in this section is as follows. We define *sibling points* as points x in different quantization intervals but with equal w. We first show that for every w, there must be at least one sibling point in each interval (surjectivity); then, we demonstrate that having more than one is bad for the reconstruction error rate. This establishes that each interval must contain exactly one sibling point for each w. Then, we show that the ordering of sibling points must be the same in each interval, because otherwise the error rate increases. Finally, assuming g to be differentiable yields the monotonicity property.

The verifier has to reconstruct x based on y and w. In general this is done by first identifying which points $x \in \mathbb{R}$ are compatible with w, and then selecting which one is most likely, given y and w. For the first step, we introduce the concept of *sibling points*.

Definition 4.1. (Sibling points): *Two points $x, x' \in \mathbb{R}$, with $x \neq x'$, are called Sibling Points if $g(x) = g(x')$.*

The verifier determines a set $\mathcal{X}_w = \{x \in \mathbb{R} | g(x) = w\}$ of sibling points that correspond to helper data value w. We write $\mathcal{X}_w = \cup_{s \in \mathcal{S}} \mathcal{X}_{sw}$, with $\mathcal{X}_{sw} = \{x \in \mathbb{R} | Q(x) = s \wedge g(x) = w\}$. We derive a number of requirements on the sets $\mathcal{X}_{sw}$.

Lemma 4.2. ZL implies that

$$\forall_{w \in \mathcal{W}, s \in \mathcal{S}} \quad \mathcal{X}_{sw} \neq \emptyset. \quad (9)$$

Proof: see Appendix A1. Lemma 4.2 tells us that there is significant leakage if there is not at least one sibling point compatible with w in each interval A_s, for all $w \in \mathcal{W}$. Since we are interested in zero leakage, we will from this point onward consider only functions g such that $\mathcal{X}_{sw} \neq \emptyset$ for all s, w.

Next, we look at the requirement of low reconstruction error probability. We focus on the *minimum distance* between sibling points that belong to different quantization intervals.

Definition 4.3. The minimum distance between sibling points in different quantization intervals is defined as

$$D_{\min}(w) = \min_{s,t \in \mathcal{S}:\, s<t} |\min \mathcal{X}_{tw} - \max \mathcal{X}_{sw}|, \quad (10)$$

$$D_{\min} = \min_{w \in \mathcal{W}} D_{\min}(w). \quad (11)$$

We take the approach of maximizing $D_{\min}$. It is intuitively clear that such an approach yields low error rates given the noise model introduced in Section 3.3. The following lemma gives a constraint that improves the $D_{\min}$.

Lemma 4.4. Let $w \in \mathcal{W}$ and $\mathcal{X}_{sw} \neq \emptyset$ for all $s \in \mathcal{S}$. The $D_{min}(w)$ is maximized by setting $|\mathcal{X}_{sw}| = 1$ for all $s \in \mathcal{S}$.

Proof: see Appendix A.2. Lemma 4.4 states that each quantization interval A_s should contain exactly one point x compatible with w. From here onward we will only consider functions g with this property.

The set $\mathcal{X}_{sw}$ consists of a single point which we will denote as x_{sw}. Note that g is then an *invertible* function on each interval A_s. For given $w \in \mathcal{W}$, we now have a set $\mathcal{X}_w = \cup_{s \in \mathcal{S}} x_{sw}$ that consists of one sibling point per quantization interval. This vastly simplifies the analysis. Our next step is to put further constraints on g.

Lemma 4.5. Let $x_1, x_2 \in A_s$ and $x_3, x_4 \in A_t$, $s \neq t$, with $x_1 < x_2 < x_3 < x_4$ and $g(x_1) = w_1$, $g(x_2) = w_2$. Consider two cases,

1. $g(x_3) = w_1; \quad g(x_4) = w_2$
2. $g(x_4) = w_1; \quad g(x_3) = w_2.$

Then it holds that

$$\min_{w\in\{w_1,w_2\}} D_{\min}^{\text{case 2}}(w) \leq \min_{w\in\{w_1,w_2\}} D_{\min}^{\text{case 1}}(w). \tag{12}$$

Proof: see Appendix A.3. Lemma 4.5 tells us that *the ordering of sibling points should be the same in each quantization interval*, for otherwise the overall minimum distance $D_{\min}$ suffers. If, for some s, a point x with helper data w_2 is higher than a point with helper data w_1, then this order has to be the same for all intervals.

The combination of having a preserved order (Lemma 4.5) together with g being invertible on each interval (Lemma 4.4) points us in the direction of "smooth" functions. If g is piecewise differentiable, then we can formulate a simple constraint as follows.

Theorem 4.6. (sign of g' equal on each A_s) : *Let $x_s \in A_s, x_t \in A_t$ be sibling points as defined in Def. 4.1. Let g be differentiable in x_s and x_t. Then having sign $g'(x_s) =$ sign $g'(x_t)$ leads to a higher D_{min} than sign $g'(x_s) \neq$ sign $g'(x_t)$.*

Proof: see Appendix A.4.

If we consider a function g that is differentiable on each quantization interval, then (1) its piecewise invertibility implies that it has to be either monotonously increasing or monotonously decreasing on each interval, and (2) Theorem 4.6 then implies that it has to be either increasing on *all* intervals or decreasing on all intervals.

Of course there is no reason to assume that g is piecewise differentiable. For instance, take a piecewise differentiable g and apply a permutation to the w-axis. This procedure yields a function g_2 which, in terms of error probabilities, has exactly the same performance as g, but is not differentiable (nor even continuous). Thus, there exist huge equivalence classes of helper data generating functions that satisfy invertibility (Lemma 4.4) and proper ordering (Lemma 4.5). This brings us to the following conjecture, which allows us to concentrate on functions that are easy to analyze.

Conjecture 4.7. Without loss of generality we can choose the function g to be differentiable on each quantization interval A_s, $s \in \mathcal{S}$.

Based on Conjecture 4.7, we will consider only functions g that are monotonically *increasing* on each interval. This assumption is in line with all (first stage) HDSs [4, 6, 8] known to us.

4.2 Quantile helper data

We state our main result in the theorem below.

Theorem 4.8. (**ZL is equivalent to quantile relationship between sibling points**): *Let g be monotonously increasing on each interval A_s, with $g(A_0) = \cdots = g(A_{N-1}) = \mathcal{W}$. Let $s, t \in \mathcal{S}$. Let $x_s \in A_s, x_t \in A_t$ be sibling points as defined in Def. 4.1. In order to satisfy Zero Leakage we have the following necessary and sufficient condition on the sibling points,*

$$\frac{F_X(x_s) - F_X(q_s)}{p_s} = \frac{F_X(x_t) - F_X(q_t)}{p_t}. \tag{13}$$

Proof: see Appendix A.5.

Corollary 4.9. (ZL FE sibling point relation): *Let g be monotonously increasing on each interval A_s, with $g(A_0) = \cdots = g(A_{N-1}) = \mathcal{W}$. Let $s, t \in \mathcal{S}$. Let $x_s \in A_s, x_t \in A_t$ be sibling points as defined in Def. 4.1. Then for a Fuzzy Extractor we have the following necessary and sufficient condition on the sibling points in order to satisfy Zero Leakage,*

$$F_X(x_s) - \frac{s}{N} = F_X(x_t) - \frac{t}{N}. \tag{14}$$

Proof. Immediately follows by combining Eq. (13) with the fact that $p_s = 1/N \ \forall s \in \mathcal{S}$ in a FE scheme, and with the FE quantization boundaries given in Eq. (5). □

Theorem 4.8 allows us to define the enrollment steps in a ZL HDS in a very simple way,

$$\begin{aligned} s &= Q(x) \\ w &= g(x) = \frac{F_X(x) - F_X(q_s)}{p_s}. \end{aligned} \tag{15}$$

Note that $w \in [0, 1)$, and $F_X(q_s) = \sum_{t=0}^{s-1} p_t$. The helper data can be interpreted as a quantile distance between x and the quantization boundary q_s, normalized with respect to the probability mass p_s in the interval A_s. An example of such a function is depicted in Fig. 4. For a specific distribution, e.g., a standard Gaussian distribution, the helper data generation function is depicted in Fig. 5. In the FE case, Eq. (15) simplifies to

$$F_X(x) = \frac{s + w}{N}; \qquad w \in [0, 1) \tag{16}$$

and the helper data generation function becomes

$$w = g^{\text{FE}}(x) = N \cdot F_X(x) - s. \tag{17}$$

Equation (16) coincides with the continuum limit of the Fuzzy extractor construction by Verbitskiy et al. [7]. A similar equation was later independently proposed for uniform key generation from a noisy channel by Ye et al. [25]. Equation (15) is the *simplest* way to implement

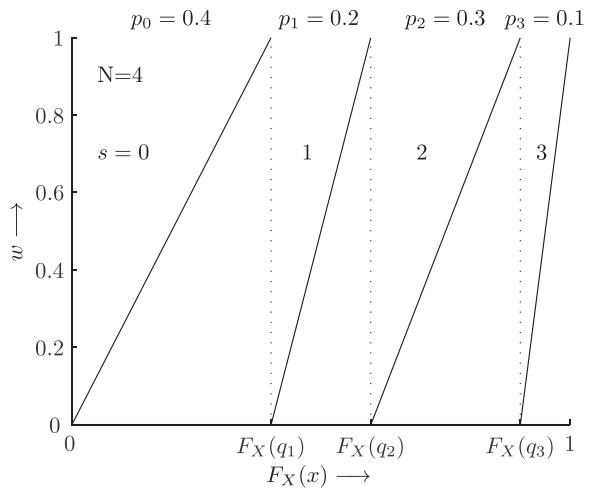

Fig. 4 Example of helper data generating function g for $N = 4$ on quantile x, i.e., $F_X(x)$. The probabilities of the secrets do not have to be equal; in this case, we have used $(p_0, \ldots, p_3) = (0.4, 0.2, 0.3, 0.1)$

an enrollment that satisfies the sibling point relation of Theorem 4.8. However, it is not the *only* way. For instance, by applying any invertible function to w, a new helper data scheme is obtained that also satisfies the sibling point relation (13) and hence is ZL. Another example is to store the whole set of sibling points $\{x_{tw}\}_{t\in\mathcal{S}}$; this contains exactly the same information as w. The transformed scheme can be seen as merely a different representation of the "basic" ZL HDS (15). Such a representation may have various advantages over (15), e.g., allowing for a faster reconstruction procedure, while being completely equivalent in terms of the ZL property. We will see such a case in Section 5.3.

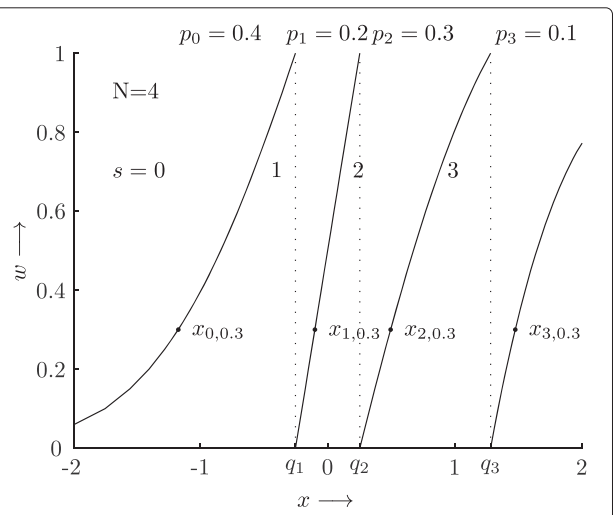

Fig. 5 Example of a helper data generating function g for a standard Gaussian distribution, i.e., $x \sim \mathcal{N}(0, 1)$, and $N = 4$. Sibling points x_{sw} are given for $s \in \{0, \ldots, 3\}$ and $w = 0.3$

5 Optimal reconstruction

5.1 Maximum likelihood and thresholds

The goal of the HDS reconstruction algorithm $\mathrm{Rep}(y, w)$ is to reliably reproduce the secret s. The best way to achieve this is to choose the most probable $\hat{s}$ given y and w, i.e., a maximum likelihood algorithm. We derive optimal decision intervals for the reconstruction phase in a Zero Leakage Fuzzy Extractor.

Lemma 5.1. Let $\mathrm{Rep}(y, w)$ *be the reproduction algorithm of a ZL FE system. Let* g_s^{-1} *be the inverse of the helper data generation function for a given secret s. Then optimal reconstruction is achieved by*

$$\mathrm{Rep}(y, w) = \arg\max_{s\in\mathcal{S}} f_{Y|X}\left(y|g_s^{-1}(w)\right). \tag{18}$$

Proof: see Appendix A.6. To simplify the verification phase we can identify thresholds τ_s that denote the lower boundary of a decision region. If $\tau_s \le y < \tau_{s+1}$, we reconstruct $\hat{s} = s$. The $\tau_0 = -\infty$ and $\tau_N = \infty$ are fixed, which implies we have to find optimal values only for the $N - 1$ variables $\tau_1, \ldots, \tau_{N-1}$ as a function of w.

Theorem 5.2. Let $f_{Y|X}$ *represent symmetric fading noise. Then optimal reconstruction in a FE scheme is obtained by the following choice of thresholds*

$$\tau_s = \lambda\frac{g_s^{-1}(w) + g_{s-1}^{-1}(w)}{2}. \tag{19}$$

Proof. In case of symmetric fading noise we know that

$$f_{Y|X}(y|x) = \varphi(|y - \lambda x|), \tag{20}$$

with φ some monotonic decreasing function. Combining this notion with that of Eq. (18) to find a point $y = \tau_s$ that gives equal probability for s and $s - 1$ yields

$$\varphi\left(|\tau_s - \lambda g_{s-1}^{-1}(w)|\right) = \varphi\left(|\tau_s - \lambda g_s^{-1}(w)|\right). \tag{21}$$

The left and right hand side of this equation can only be equal for equal arguments, and hence

$$\tau_s - \lambda g_{s-1}^{-1}(w) = \pm\left(\tau_s - \lambda g_s^{-1}(w)\right). \tag{22}$$

Since $g_s^{-1}(w) \neq g_{s-1}^{-1}(w)$ the only viable solution is Eq. (19). □

Instead of storing the ZL helper data w according to (15), one can also store the set of thresholds $\tau_1, \ldots, \tau_{N-1}$. This contains precisely the same information, and allows for quicker reconstruction of s: just a thresholding operation on y and the τ_s values, which can be implemented on computationally limited devices.

5.2 Special case: 1-bit secret

In the case of a one-bit secret s, i.e., $N = 2$, the above ZL FE scheme is reduced to storing a single threshold τ_1.

It is interesting and somewhat counterintuitive that this yields a threshold for verification that does not leak information about the secret. In case the average of X is zero, one might assume that a positive threshold value implies $s = 0$. However, both $s = 0$ and $s = 1$ allow positive as well as negative τ_1, dependent on the relative location of x in the quantization interval.

5.3 FE: equivalent choices for the quantization

Let us reconsider the quantization function $Q(x)$ in the case of a Fuzzy extractor. Let us fix N and take the $g(x)$ as specified in Eq. (16). Then, it is possible to find an infinite number of different functions Q that will conserve the ZL property and lead to exactly the same error rate as the original scheme. This is seen as follows. For any $w \in [0, 1)$, there is an N-tuplet of sibling points. Without any impact on the reconstruction performance, we can permute the s-values of these points; the error rate of the reconstruction procedure depends only on the x-values of the sibling points, not on the s-label they carry. It is allowed to do this permutation for every w independently, resulting in an infinite equivalence class of Q-functions. The choice we made in Section 3 yields the simplest function in an equivalence class.

6 Example: Gaussian features and BCH codes

To benchmark the reproduction performance of our scheme, we give an example based on Gaussian-distributed variables. In this example, we will assume all variables to be Gaussian distributed, though we remind the reader that our scheme specifies optimal reconstruction thresholds even for non-Gaussian distributions.

We compare the reproduction performance of our ZL quantization scheme with Likelihood-based reproduction (ZLQ-LR) to a scheme with (1) fixed quantization (FQ), see Section 2.1, and (2) likelihood classification (LC). The former is, to our knowledge, the only other scheme sharing the zero secrecy leakage property, since it does not use any helper data. An example with $N = 4$ intervals is depicted in Fig. 6a. LC is not an actual quantization scheme since it requires the enrollment sample to be stored in-the-clear. However, a likelihood based classifier provides an optimal trade-off between false acceptance and false rejection according to communication theory [24] and should therefore yield the lowest possible error rate. Instead of quantization boundaries, the classifier is characterized by decision boundaries as depicted in Fig. 6b.

A comparison with QIM cannot be made since there the probability for an impostor to guess the enrolled secret cannot be made equal to $1/N$. This would result in an unfair comparison since the other schemes are designed to possess this property. Moreover, the QIM scheme allows the reproduction error probability to be made arbitrary small by increasing the quantization width at the cost of leakage.

Also, the likelihood based classification can be tuned by setting the decision threshold. However, for this scheme, it is possible to choose a threshold such that an impostor will have a probability of $1/N$ to be accepted, which corresponds to the $1/N$ probability of guessing the enrolled secret in a FE scheme. Note that for a likelihood classifier, there is no enrolled secret since this is not a quantization scheme.

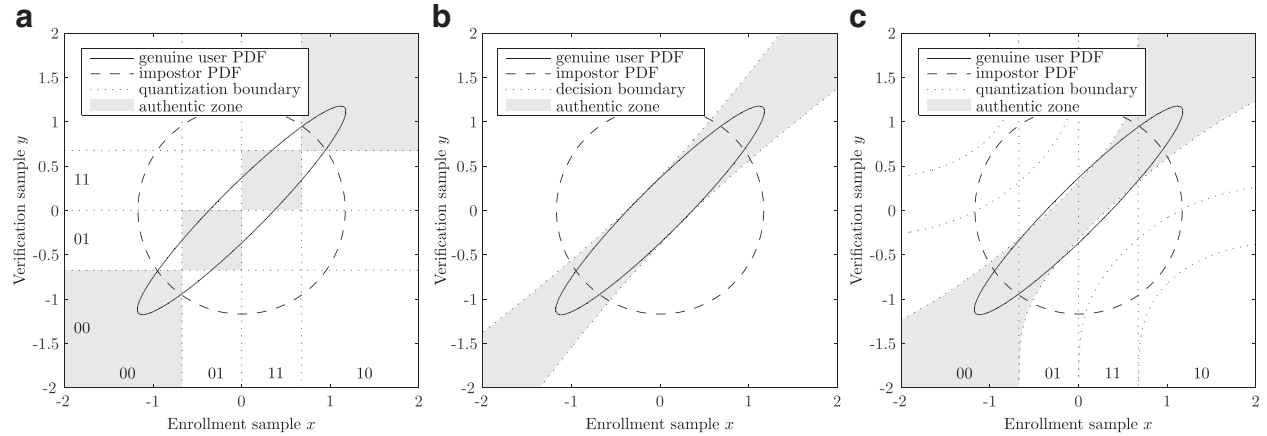

Fig. 6 Quantization and decision patterns based on the genuine user and impostor PDFs. Ideally the genuine user PDF should be contained in the authentic zone and the impostor PDF should have a large mass outside the authentic zone. Fifty percent probability mass is contained in the genuine user and impostor PDF ellipse and circle. The genuine user PDF is based on a 10 dB SNR. **a** Fixed equiprobable quantization (FQ). **b** Likelihood classification (LC). **c** Zero leakage quantization scheme with likelihood based reproduction (ZLQ-LR)

As can be seen from Fig. 7, the reproduction performance for a ZL scheme with likelihood based reproduction is always better than that of a fixed quantization scheme. However, it is outperformed by the likelihood classifier. Differences are especially apparent for features with a higher signal-to-noise ratio. In these regions, the fixed quantization struggles with a inherent high error probability, while the ZL scheme follows the LC.

In a good quantization scheme, the gap between $I(X;Y)$ and $I(S;\hat{S})$ must be small. For a Gaussian channel, standard expressions are known from ([24], Eq. (9.16)). Figure 8 shows that a fixed quantization requires a higher SNR on order to converge to the maximum number of bits, whereas the ZLQ-LR scheme directly reaches this value.

Finally, we consider the vector case of the two quantization schemes discussed above. We concluded that FQ has a larger error probability, but we now show how this relates to either false rejection or secret length when combined with a code offset method [3].

We assume i.i.d. features and therefore we can calculate false acceptance rate (FAR) and false rejection rate (FRR) based on a binomial distribution. In practice, features can be made (nearly) independent, but they will in general not be identically distributed. However, results will be similar. Furthermore we assume the error correcting code can be applied such that its error correcting properties can be fully exploited. This implies that we have to use a Gray code to label the extracted secrets before concatenation.

We used 64 i.i.d. features, each having a SNR of 17 dB, which is a typical average value for biometric features [8, 17]. From these features, we extract 2 bits per feature on which we apply BCH codes with a code length of 127. (We omit one bit). For analysis, we have also included the code $(127, 127, 0)$, which is not an actual code, but represents the case in which no error correction is applied.

Suppose we want to achieve a target FRR of $1 \cdot 10^{-3}$, the topmost dotted line in Fig. 9, then we require a BCH $(127, 92, 5)$ code for the ZLQ-LR scheme, while a BCH $(127, 15, 27)$ code is required for the FQ scheme. This implies that we would have a secret key size of 92 bits versus 15 bits. Clearly, the latter is not sufficient for any security application. At the same time, due to the small key size, FQ has an increased FAR.

7 Conclusions

In this paper, we have studied a generic helper data scheme (HDS) which comprises the Fuzzy extractor (FE) and the secure sketch (SS) as special cases. In particular, we have looked at the zero leakage (ZL) property of HDSs in the case of a one-dimensional continuous source X and continuous helper data W.

We make minimal assumptions, justified by Conjecture 4.7: we consider only monotonic $g(x)$. We have shown that the ZL property implies the existence of sibling points $\{x_{sw}\}_{s\in\mathcal{S}}$ for every w. These are values of x that have the same helper data w. Furthermore, the ZL requirement is equivalent to a quantile relationship (Theorem 4.8) between the sibling points. This directly leads to Eq. (15) for computing w from x. (Applying any reversible function to this w yields a completely equivalent helper data

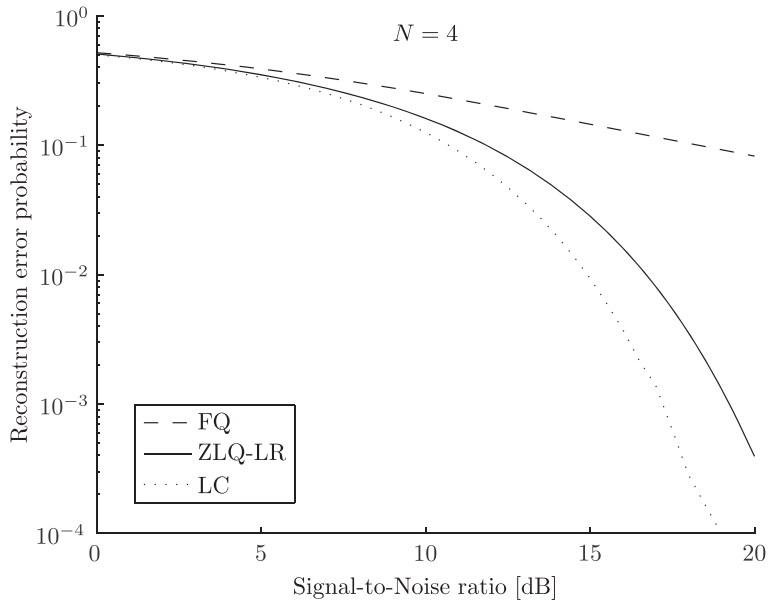

Fig. 7 Reconstruction error probability for Gaussian-distributed features and Gaussian noise for $N = 4$

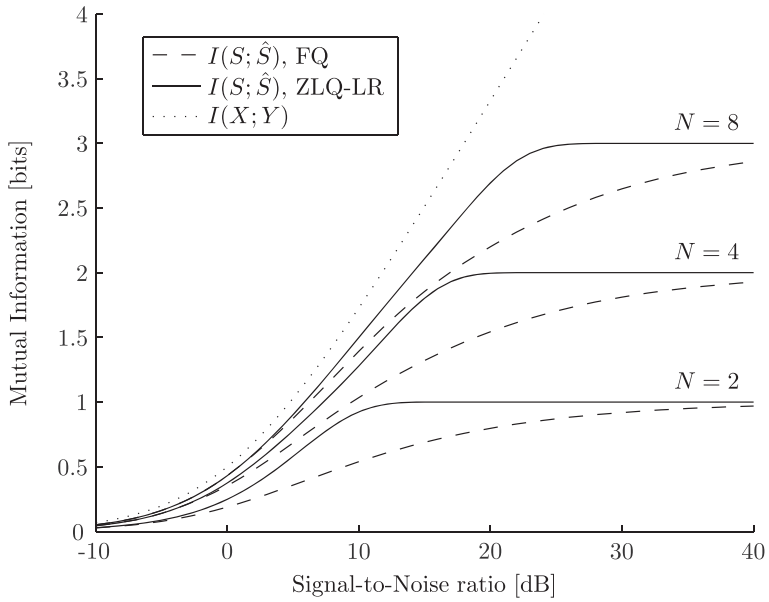

Fig. 8 Mutual information between S and $\hat{S}$ for Gaussian-distributed features and Gaussian noise

system.) The special case of a FE ($p_s = 1/N$) yields the $m \to \infty$ limit of the Verbitskiy et al. [7] construction.

We have derived reconstruction thresholds τ_s for a ZL FE that minimize the error rate in the reconstruction of s (Theorem 5.2). This result holds under very mild assumptions on the noise: symmetric and fading. Equation (19) contains the attenuation parameter λ, which follows from the noise model as specified in Section 3.3.

Finally, we have analyzed reproduction performance in an all-Gaussian example. Fixed quantization struggles with inherent high error probability, while the ZL FE with optimal reproduction follows the performance of the optimal classification algorithm. This results in a larger key size in the protected template compared to the fixed quantization scheme, since an ECC with a larger message length can be applied in the second stage HDS to achieve the same FRR.

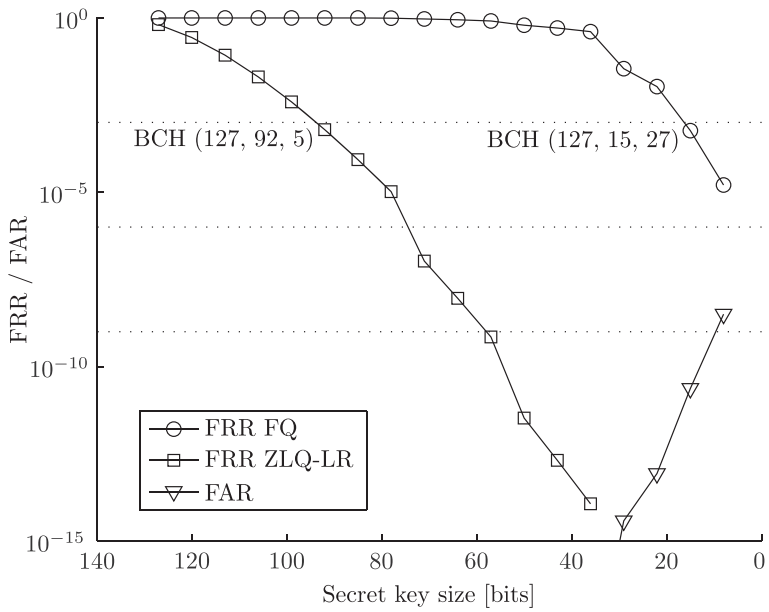

Fig. 9 System performance of ZLQ-LR and FQ

In this paper, we have focused on arbitrary but known probability densities. Experiments with real data are beyond the scope of this paper, but have been reported in [26, 27]. A key finding there was that modeling the distributions can be problematic, especially due to statistical outliers. Even so, improvements were obtained with respect to earlier ZL schemes. We see modeling refinements as a topic for future research.

Endnotes

[1] For information-theoretic concepts such as Shannon entropy and mutual information we refer to e.g. [24].

[2] This concept is not new. Achieving zero mutual information has always been a (sometimes achievable) desideratum in the literature on fuzzy vaults/fuzzy extractors/secure sketches for discrete and continuous sources.

[3] An overall Zero-Leakage scheme can be obtained in the final HDS stage even from a leaky HDS by applying privacy amplification as a post-processing step. However, this procedure discards substantial amounts of source entropy, while in many practical applications it is already a challenge to achieve reasonable security levels from biometrics without privacy protection.

Appendix

Appendix A: Proofs

A.1 Proof of Lemma 4.2

Pick any $w \in \mathcal{W}$. The statement $w \in \mathcal{W}$ means that there exists at least one $s' \in \mathcal{S}$ such that $\mathcal{X}_{s'w} \neq \emptyset$. Now suppose there exists some $s'' \in \mathcal{S}$ with $\mathcal{X}_{s''w} = \emptyset$. Then knowledge of w reveals information about S, i.e. $\mathbb{P}[S = s|W = w] \neq \mathbb{P}[S = s]$, which contradicts ZL.

A.2 Proof of Lemma 4.4

Let g be such that $|\mathcal{X}_{sw}| > 1$ for some s, w. Then choose a point $\bar{x} \in \mathcal{X}_{sw}$. Construct a function g_2 such that

$$g_2(x) \begin{cases} = g(x) & \text{if } x = \bar{x} \text{ or } x \notin \mathcal{X}_{sw} \\ \neq g(x) & \text{otherwise} \end{cases}. \tag{23}$$

The $D_{\min}(w)$ for g_2 cannot be smaller than $D_{\min}(w)$ for g.

A.3 Proof of Lemma 4.5

We tabulate the $D_{\min}(w)$ values for case 1 and 2,

	case 1	case 2
w_1	$x_3 - x_1$	$x_4 - x_1$
w_2	$x_4 - x_2$	$x_3 - x_2$

The smallest of these distances is $x_3 - x_2$.

A.4 Proof of Theorem 4.6

Let $0 < \varepsilon \ll 1$ and $0 < \delta \ll 1$. Without loss of generality we consider $s < t$. We invoke Lemma 4.5 with $x_1 = x_s$, $x_2 = x_s + \varepsilon, x_3 = x_t, x_4 = x_t + \delta$. According to Lemma 4.5 we have to take $g(x_2) = g(x_4)$ in order to obtain a large $D_{\min}$. Applying a first order Taylor expansion, this gives

$$g(x_s) + \varepsilon g'(x_s) + \mathcal{O}(\varepsilon^2) = g(x_t) + \delta g'(x_t) + \mathcal{O}(\delta^2). \tag{24}$$

We use the fact that $g(x_s) = g(x_t)$, and that ε and δ are positive. Taking the sign of both sides of (24) and neglecting second order contributions, we get sign $g'(x_s) =$ sign $g'(x_t)$.

A.5 Proof of Theorem 4.8

The ZL property is equivalent to $f_W = f_{W|S}$, which gives for all $s \in \mathcal{S}$

$$f_W(w) = f_{W|S}(w|s) = \frac{f_{W,S}(w,s)}{p_s}, \tag{25}$$

where $f_{W,S}$ is the joint distribution for W and S. We work under the assumption that $w = g(x)$ is a monotonous function on each interval A_s, fully spanning $\mathcal{W}$. Then for given s and w there exists exactly one point x_{sw} that satisfies $Q(x) = s$ and $g(x) = w$. Furthermore, conservation of probability then gives $f_{W,S}(w,s)\,\mathrm{d}w = f_X(x_{sw})\,\mathrm{d}x_{sw}$. Since the right hand side of (25) is independent of s, we can write $f_W(w)\mathrm{d}w = p_s^{-1} f_X(x_{sw})\mathrm{d}x_{sw}$ for *any* $s \in \mathcal{S}$. Hence for any $s, t \in \mathcal{S}$, $w \in \mathcal{W}$ it holds that

$$\frac{f_X(x_{sw})\mathrm{d}x_{sw}}{p_s} = \frac{f_X(x_{tw})\mathrm{d}x_{tw}}{p_t}, \tag{26}$$

which can be rewritten as

$$\frac{\mathrm{d}F_X(x_{sw})}{p_s} = \frac{\mathrm{d}F_X(x_{tw})}{p_t}. \tag{27}$$

The result (13) follows by integration, using the fact that A_s has lower boundary q_s.

A.6 Proof of Lemma 5.1

Optimal reconstruction can be done by selecting the most likely secret given y, w,

$$\mathrm{Rep}(y,w) = \arg\max_{s\in\mathcal{S}} f_{S|Y,W}(s|y,w) = \arg\max_{s\in\mathcal{S}} \frac{f_{Y,S,W}(y,s,w)}{f_{Y,W}(y,w)}. \tag{28}$$

The denominator does not depend on s, and can hence be omitted. This gives

$$\mathrm{Rep}(y,w) = \arg\max_{s\in\mathcal{S}} f_{S,Y,W}(s,y,w) \tag{29}$$

$$= \arg\max_{s\in\mathcal{S}} f_{Y|S,W}(y|s,w) f_{W|S}(w|s) p_s. \tag{30}$$

We constructed the scheme to be a FE with ZL, and therefore $p_s = 1/N$ and $f_{W|S}(w|s) = f_W(w)$. We see that both p_s and $f_{W|S}(w|s)$ do not depend on s, which implies they can be omitted from Eq. (30), yielding $\mathrm{Rep}(y,w) = \arg\max_{s\in\mathcal{S}} f_{Y|S,W}(y|s,w)$. Finally, knowing S and W is

equivalent to knowing X. Hence $f_{Y|S,W}(y|s,w)$ can be replaced by $f_{Y|X}(y|x)$ with x satisfying $Q(x) = s$ and $g(x) = w$. The unique x value that satisfies these constraints is $g_s^{-1}(w)$.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Signal Processing Systems group, Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB, Eindhoven, The Netherlands. [2]Security and Embedded Networked Systems group, Department of Mathematics and Computer Science, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands. [3]Discrete Mathematics group, Department of Mathematics and Computer Science, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands. [4]Genkey Solutions B.V., High Tech Campus 69, 5656 AG Eindhoven, The Netherlands.

Received: 22 October 2015 Accepted: 20 April 2016
Published online: 05 May 2016

References

1. T van der Putte, J Keuning, in *Proceedings of the Fourth Working Conference on Smart Card Research and Advanced Applications on Smart Card Research and Advanced Applications*. Biometrical fingerprint recognition: don't get your fingers burned (Kluwer Academic Publishers, Norwell, MA, USA, 2001), pp. 289–303
2. T Matsumoto, H Matsumoto, K Yamada, S Hoshino, Impact of artificial "gummy" fingers on fingerprint systems. Opt. Secur. Counterfeit Deterrence Tech. **4677**, 275–289 (2002)
3. A Juels, M Wattenberg, in *CCS '99: Proceedings of the 6th ACM Conf on Comp and Comm Security*. A fuzzy commitment scheme (ACM, New York, NY, USA, 1999), pp. 28–36. doi:10.1145/319709.319714. http://doi.acm.org/10.1145/319709.319714
4. J-P Linnartz, P Tuyls, in *New Shielding Functions to Enhance Privacy and Prevent Misuse of Biometric Templates*, ed. by J Kittler, Mark Nixon. Audio- and Video-Based Biometric Person Authentication: 4th International Conference, AVBPA 2003 Guildford, UK, June 9–11, 2003 Proceedings (Springer Berlin Heidelberg, Berlin, Heidelberg, 2003), pp. 393–402. doi:10.1007/3-540-44887-X_47. http://dx.doi.org/10.1007/3-540-44887-X_47
5. P Tuyls, B Škorić, T Kevenaar, *Security with Noisy Data: Private Biometrics, Secure Key Storage and Anti-Counterfeiting*. (Springer, Secaucus, NJ, USA, 2007)
6. C Chen, RNJ Veldhuis, TAM Kevenaar, AHM Akkermans, in *Proc. IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems*. Multi-bits biometric string generation based on the likelihood ratio (IEEE, Piscataway, 2007)
7. EA Verbitskiy, P Tuyls, C Obi, B Schoenmakers, B Škorić, Key extraction from general nondiscrete signals. Inform. Forensics Secur. IEEE Trans. **5**(2), 269–279 (2010)
8. JA de Groot, J-PMG Linnartz, in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process*. Zero leakage quantization scheme for biometric verification, (Piscataway, 2011)
9. Y Dodis, L Reyzin, A Smith, in *Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data*, ed. by C Cachin, JL Camenisch. Advances in Cryptology - EUROCRYPT 2004: International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken, Switzerland, May 2-6, 2004. Proceedings (Springer Berlin Heidelberg, Berlin, Heidelberg, 2004), pp. 523–540. doi:10.1007/978-3-540-24676-3_31 http://dx.doi.org/10.1007/978-3-540-24676-3_31
10. AV Herrewege, S Katzenbeisser, R Maes, R Peeters, A-R Sadeghi, I Verbauwhede, C Wachsmann, in *Reverse Fuzzy Extractors: Enabling Lightweight Mutual Authentication for PUF-Enabled RFIDs*, ed. by AD Keromytis. Financial Cryptography and Data Security: 16th International Conference, FC 2012, Kralendijk, Bonaire, Februray 27-March 2, 2012, Revised Selected Papers (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), pp. 374–389. doi:10.1007/978-3-642-32946-3_27. http://dx.doi.org/10.1007/978-3-642-32946-3_27
11. B Škorić, N de Vreede, The spammed code offset method. IEEE Trans. Inform. Forensics Secur. **9**(5), 875–884 (2014)
12. F MacWilliams, N Sloane, *The Theory of Error Correcting Codes*. (Elsevier, Amsterdam, 1978)
13. JL Wayman, AK Jain, D Maltoni, D Maio (eds.), *Biometric Systems: Technology, Design and Performance Evaluation*, 1st edn. (Spring Verlag, London, 2005)
14. B Škorić, P Tuyls, W Ophey, in *Robust Key Extraction from Physical Uncloneable Functions*, ed. by J Ioannidis, A Keromytis, and M Yung. Applied Cryptography and Network Security: Third International Conference, ACNS 2005, New York, NY, USA, June 7-10, 2005. Proceedings (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005), pp. 407–422. doi:10.1007/11496137_28. http://dx.doi.org/10.1007/11496137_28
15. GE Suh, S Devadas, in *Proceedings of the 44th Annual Design Automation Conference*. DAC '07. Physical unclonable functions for device authentication and secret key generation (ACM, New York, NY, USA, 2007), pp. 9–14
16. DE Holcomb, WP Burleson, K Fu, Power-Up SRAM state as an identifying fingerprint and source of true random numbers. Comput. IEEE Trans. **58**(9), 1198–1210 (2009)
17. P Tuyls, A Akkermans, T Kevenaar, G-J Schrijen, A Bazen, R Veldhuis, in *Practical Biometric Authentication with Template Protection*, ed. by T Kanade, A Jain, and N Ratha. Audio- and Video-Based Biometric Person Authentication: 5th International Conference, AVBPA 2005, Hilton Rye Town, NY, USA, July 20-22, 2005. Proceedings (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005), pp. 436–446. doi:10.1007/11527923_45. http://dx.doi.org/10.1007/11527923_45
18. EJC Kelkboom, GG Molina, J Breebaart, RNJ Veldhuis, TAM Kevenaar, W Jonker, Binary biometrics: an analytic framework to estimate the performance curves under gaussian assumption. Syst. Man Cybernetics, Part A: Syst. Hum. IEEE Trans. **40**(3), 555–571 (2010). doi:10.1109/TSMCA.2010.2041657
19. EJC Kelkboom, KTJ de Groot, C Chen, J Breebaart, RNJ Veldhuis, in *Biometrics: Theory, Applications, and Systems, 2009. BTAS '09. IEEE 3rd International Conference On*. Pitfall of the detection rate optimized bit allocation within template protection and a remedy, (2009), pp. 1–8. doi:10.1109/BTAS.2009.5339046
20. B Chen, GW Wornell, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding. Inform. Theory IEEE Trans. **47**(4), 1423–1443 (2001). doi:10.1109/18.923725
21. MHM Costa, Writing on dirty paper (corresp.) IEEE Trans. Inform. Theory. **29**(3), 439–441 (1983)
22. M-H Lim, ABJ Teoh, K-A Toh, Dynamic detection-rate-based bit allocation with genuine interval concealment for binary biometric representation. IEEE Trans. Cybernet., 843–857 (2013)
23. C Chen, RNJ Veldhuis, TAM Kevenaar, AHM Akkermans, Biometric quantization through detection rate optimized bit allocation. EURASIP J. Adv. Signal Process. **2009**, 29–12916 (2009). doi:10.1155/2009/784834
24. TM Cover, JA Thomas, *Elements of Information Theory*, 2nd edn. (John Wiley & Sons, Inc., 2005)
25. C Ye, S Mathur, A Reznik, Y Shah, W Trappe, NB Mandayam, Information-theoretically secret key generation for fading wireless channels. IEEE Trans. Inform. Forensics Secur. **5**(2), 240–254 (2010)
26. JA de Groot, J-PMG Linnartz, in *Proc. WIC Symposium on Information Theory in the Benelux*. Improved privacy protection in authentication by fingerprints (WIC, The Netherlands, 2011)
27. JA de Groot, B Škorić, N de Vreede, J-PMG Linnartz, in *Security and Cryptography (SECRYPT), 2013 International Conference On*. Diagnostic category leakage in helper data schemes for biometric authentication (IEEE, Piscataway, 2013), pp. 1–6