Security analysis of Quantum-Readout PUFs in the case of generic challenge-estimation attacks

Boris Škorić

Abstract

Quantum Readout PUFs (QR-PUFs) have been proposed as a technique for remote authentication of objects. The security is based on basic quantum information theoretic principles and the assumption that the adversary cannot efficiently implement arbitrary unitary transformations. We analyze the security of QR-PUF schemes in the case where each challenge consists of precisely n quanta and the dimension K of the Hilbert space is larger than n^2. We consider a class of attacks where the adversary first tries to learn as much as possible about the challenge and then bases his response on his estimate of the challenge. For this class of attacks we derive an upper bound on the adversary's success probability as a function of K and n.

1 Introduction

1.1 Physical Unclonable Functions

Authentication is usually based on either "something that you know" or "something that you possess". In the second case it is desirable to work with tokens that are difficult to clone, even for the manufacturer of the token. With the advent of Physical Unclonable Functions (PUFs), physical systems have been developed which satisfy strong uniqueness and unclonability properties, e.g. phenomena such as laser speckle based on multiple scattering. A PUF is a complex piece of material whose structure is difficult to reproduce accurately because its manufacture contains uncontrollable steps [1, 2, 3, 4, 5, 6, 7, 8, 9]. A stimulus can be applied to the PUF ('challenge'), leading to a 'response' that depends in a complex way on the challenge and the precise details of the PUF's structure. The combination of a challenge and the corresponding response is called a Challenge-Response Pair (CRP).

An example of a physical system satisfying the above requirements is the so-called Optical PUF: a three-dimensional diffusive structure containing randomly positioned optical scatterers. When an Optical PUF is illuminated by a laser, the transmitted and reflected light has a random-looking pattern of dark and bright spots known as speckle. The characteristics of the laser beam (e.g. wavelength, angle, focus) constitute the challenge; the speckle pattern is the response. The response depends strongly on the challenge and on the exact positions of the scatterers. Optical PUFs support a large number of independent CRPs [10, 11, 12].

1.2 Quantum readout of PUFs

A PUF-based authentication or anti-counterfeiting system typically has two phases: enrollment and verification. In the enrollment phase the Verifier applies a limited number of random challenges to a PUF and records the CRPs in a database. Later, in the verification phase, the Verifier has to decide whether a PUF is authentic. He looks up the CRPs listed for that given PUF, and by challenging the PUF anew verifies if it produces the listed responses. The procedure sketched above is extremely reliable when the Verifier has full physical control over the PUF. There are many cases, however, where the PUF owner is unwilling or unable to hand over his PUF. In such situations the Verifier must do verification without having full control. This is referred to as "hands-off" verification. Achieving a high level of security is far more difficult in this setting, since there is a serious danger of emulation ('spoofing').

In practical situations, the number of supported independent CRPs is 'small' in the sense that anyone holding the PUF can, in a feasible amount of time, extract enough information from the PUF to be able to compute (or look up) the response to any future PUF challenge without having to use the PUF any more. Thus we must assume that a PUF can be *emulated* once the adversary has had a chance to examine it. This also holds for Optical PUFs, though the emulation may require quite a large database of CRPs. In general, the stricter the robustness requirements (i.e. reproducibility of responses), the smaller the challenge space and hence the more serious the danger of emulation.

The usual way to retain control in the "hands-off" setting is to have a trusted measurement device in the field or extra sensors for detecting specific kinds of spoofing. This approach has a drawback: The extra anti-spoofing hardware adds cost, while it is difficult to ascertain how secure the system actually is. For instance, remote trusted devices need to be tamper-proofed, but hardware attacks improve over time. Similarly, new techniques are continuously developed to spoof sensors. Thus, it is an arms race situation.

An elegant way out of this expensive arms race was proposed in [13]: Quantum Readout (QR) of PUFs. It makes spoofing fundamentally difficult by making use of basic quantum information theoretic principles. The main idea is to have PUF challenges that are quantum states, so that the adversary cannot extract all information from them; if he does not know the challenge, he does not know what to emulate. This approach is fundamentally secure as long as the adversary does not have the means to efficiently[1] apply arbitrary unitary transformations to the quantum state.

More specifically, the scheme works as follows. The PUF interacts with the challenge state via unitary evolution and produces a response that is also a quantum state. The Verifier, who knows from the enrollment phase what the response state is supposed to be, is able to verify if the response is correct. All this can be done without a trusted remote device, because of the inherent tamper-resistant properties of single quanta. The No Cloning Theorem [16, 17] ensures that an unknown single quantum state cannot be copied onto another particle. One of the implications is that the state of an unknown quantum challenge cannot be fully determined. By repeatedly sending random challenges, the verifier ensures that the probability of successful spoofing is brought down exponentially. Nice properties of the QR-PUF technique are that the challenge space does not have to be large, and that the scheme is still secure if the list of responses is publicly known.

Quantum Readout of PUFs was first experimentally realized by Goorden et al. [15] in an Optical PUF system. The challenge was implemented as a weak coherent light pulse with average photon number n and a randomly chosen wavefront that has K degrees of freedom, with $K \gg n$. The scattering in the PUF scrambles the wavefront. The response is the scrambled light pulse. Verification was performed using a spatial light modulator and a photon counter. The security is based on the fact that performing measurements on n photons reveals too little information to characterize the K-mode challenge state.

1.3 Security of Quantum-Readout PUFs: previous work

The existing security analyses of QR-PUFs assume that the adversary does not have a way to perform arbitrary unitary operations. The analyses are restricted to so-called *challenge estimation* attacks, in which the adversary first does a measurement on the challenge, from the outcome calculates an estimate of the challenge and finally produces a response quantum state consistent with the estimated challenge. We will also work in this context.

In [14] it was proven for the case of single-quantum challenges that the per-round false accept probability cannot exceed $2/(K + 1)$, where K is the dimension of the Hilbert space. Hence, QR-OUFs can be secure even if the dimension of the Hilbert space is low, e.g. $K = 2$.

Ref. [18] analyzed the case of Optical QR-PUFs [15] with K modes and average photon number n. The adversary has (on average) n quanta to examine, which gives him more information than in

[1] By 'efficient' we mean without particle losses, fast, and at a reasonable cost. At the moment, and in the foreseeable future, there is no lossless way to apply arbitrary unitary operations in the optical PUF system of [15].

the $n = 1$ case. A specific type of measurement was considered known as *quadrature*, which is the most informative kind of measurement on electromagnetic fields. It was shown that a challenge estimation attack in this context cannot achieve a per-quantum false accept probability better than approximately $n/(K + n)$. Thus, security is achievable as long as n is not large compared to K.

The previous work does not provide a result for the case of more than one quantum in the challenge ($n > 1$) combined with *arbitrary* measurements by the adversary.

1.4 Contributions and outline

We analyze the security of QR-PUFs against challenge-estimation attacks. We consider the case where the challenge comprises exactly n quanta, without making assumptions about the measurements allowed to the adversary.

- Our main result is an upper bound on the adversary's per-quantum accept probability. For $n \ll \sqrt{K}$ the result is approximately $(6n - 4)/K$. A bound on his per-round accept probability and the overall false accept probability are straightforwardly obtained.
- For $n = 1$ our bound reduces to the known result $2/(K + 1)$, but the proof is much more elegant than the proof in [14].

In Section 2 we detail the attacker model and briefly review Mutually Unbiased Bases (MUBs). MUBs are important to us since they represent a set of most informative measurements for the adversary. We also discuss Gauss sums, which arise because of the use of MUBs, and generalized Beta functions, which pop up when one averages over the challenge space. In Section 3 we derive our bound. Section 4 contains a short discussion.

2 Preliminaries

2.1 Notation

Quantum states are represented as vectors in a Hilbert space. We adopt the usual Dirac 'bra' and 'ket' notation; the ket vector $|\psi\rangle$ stands for a quantum state labelled by some description ψ which summarizes all the knowable information about the state. The Hermitian conjugate is denoted as the bra vector $\langle\psi|$. The notation for the inner product between two states is $\langle\psi_1|\psi_2\rangle$. We will consider only normalized states, i.e. satisfying $\langle\psi|\psi\rangle = 1$. Real-valued observables are represented by Hermitian operators acting on the Hilbert space. The expectation value of an operator A, given state $|\psi\rangle$, is denoted as $\langle\psi|A|\psi\rangle$, or in shorthand notation $\langle A\rangle$ when it is clear from the context what the state is. The Hermitian conjugate of A is denoted as $A^\dagger$.

We will consider a K-dimensional Hilbert space, with $K \gg 1$. The set $\{0, \dots, K - 1\}$ will be abbreviated as $\mathcal{K}$; the set $\{1, \dots, n\}$ as $[n]$. We reserve the symbol δ for the Kronecker delta (as in δ_{ab}) and for the Dirac delta function. The standard basis states are written as $|z\rangle$, with $z \in \mathcal{K}$. We define $\omega = \exp(i2\pi/K)$.

Vectors that are not quantum states will be written in boldface. We define $\mathbf{1}_K$ as the K-component vector $(1, \cdots, 1)$.

Furthermore, we use multi-index notation: $\boldsymbol{x}^{\boldsymbol{u}}$ stands for the product $\prod_j x_j^{u_j}$.

The properties of the PUF are summarized as a unitary $K \times K$ transition matrix R. The PUF response to a challenge $|\psi\rangle$ is $R|\psi\rangle$.

2.2 Attacker model

We consider the following attacker model. The verifier prepares a challenge consisting of exactly n quanta (with $n < K$) that are all in the same state $|\psi\rangle$. He is allowed to choose any $|\psi\rangle$ in the Hilbert space. He sends the challenge to the PUF holder. There the challenge interacts with the

PUF, resulting in a response state. The challenge state can be written as $|\Psi\rangle = \otimes_{\alpha=1}^{n}|\psi\rangle_\alpha$, and the expected response state is $|\Omega\rangle = \otimes_{\alpha=1}^{n}|\omega\rangle_\alpha$ with $|\omega\rangle = R|\psi\rangle$. The response state is returned to the verifier.

For each quantum independently the verifier checks the validity of the response. He does this by projecting each response quantum onto $|\omega\rangle$ (with measurement outcome 1 in the case of a match and 0 otherwise). We assume that he has the technological means to measure the projection operator $|\omega\rangle\langle\omega|$ for arbitrary $|\omega\rangle$. Ideally, the correct response yields n matches. However, imperfections in the equipment may cause some noise. (Noise can occur at any stage: challenge preparation, state transport, interaction with the PUF, and measurement of $|\omega\rangle\langle\omega|$.) In order to accommodate for such noise, the verifier tolerates a fraction $\varepsilon_{\text{noise}}$ of all projection outcomes to be zero.

We investigate the following attack. The attacker fully knows R but does not possess the PUF. Furthermore, he does not possess a quantum computer or, equivalently, a device that can perform arbitrary unitary operations in a lossless way. The attacker performs a measurement on each of the n quanta separately, in order to estimate $|\psi\rangle$ as accurately as he can. He chooses n Hermitian operators $B^{(1)}, \cdots, B^{(n)}$. The set of operators is denoted as $\mathcal{B} = \{B^{(\alpha)}\}_{\alpha=1}^{n}$. Each operator $B^{(\alpha)}$ has its own orthonormal eigenbasis of eigenvectors $|b_{\alpha j}\rangle$ with $j \in \mathcal{K}$. Without loss of generality we scale the eigenvalues such that $B^{(\alpha)}|b_{\alpha j}\rangle = j|b_{\alpha j}\rangle$; this is allowed, since we are only interested in the eigenvectors. The attacker performs measurement $B^{(\alpha)}$ on the α'th quantum in the challenge. The outcome of measurement α is denoted as $k_\alpha \in \mathcal{K}$, and we define a vector $\boldsymbol{k} = (k_\alpha)_{\alpha\in[n]}$. The outcome state of measurement α is $|b_{\alpha k_\alpha}\rangle$. Based on $\mathcal{B}$ and $\boldsymbol{k}$, the attacker computes an estimate of $|\psi\rangle$. We denote this estimate as $|\hat{\psi}_{\boldsymbol{k}}\rangle$. The best estimate (i.e. the one with the highest probability of being correct, conditioned on the observed $\boldsymbol{k}$) is given by the average of the outcome vectors $|b_{\alpha k_\alpha}\rangle$, where each of these vectors is given equal weight. We define

$$|\hat{\psi}_{\boldsymbol{k}}\rangle \propto \sum_{\alpha\in[n]} |b_{\alpha k_\alpha}\rangle, \tag{1}$$

with $\langle\hat{\psi}_{\boldsymbol{k}}|\hat{\psi}_{\boldsymbol{k}}\rangle = 1$. The normalization constant for $|\hat{\psi}_{\boldsymbol{k}}\rangle$ is denoted as $\mathcal{N}_{\boldsymbol{k}}$.

$$|\hat{\psi}_{\boldsymbol{k}}\rangle = \frac{\sum_\alpha |b_{\alpha k_\alpha}\rangle}{\sqrt{\sum_{\alpha\beta}\langle b_{\alpha k_\alpha}|b_{\beta k_\beta}\rangle}} = \frac{1}{\sqrt{\mathcal{N}_{\boldsymbol{k}}}}\sum_\alpha |b_{\alpha k_\alpha}\rangle. \tag{2}$$

The attacker computes $|\hat{\omega}_{\boldsymbol{k}}\rangle = R|\hat{\psi}_{\boldsymbol{k}}\rangle$, prepares this state n times and sends $\otimes_{\alpha=1}^{n}|\hat{\omega}_{\boldsymbol{k}}\rangle$ back to the verifier.

We say that the attack has succeeded if, in a succession of challenge-response protocols, the success rate exceeds $1 - \varepsilon_{\text{noise}}$.

2.3 Mutually unbiased operators

In the above described setting with n quanta, it is known[19, 20] that a set of n *mutually unbiased* operators optimally extracts information from the state $|\psi\rangle$. We briefly review the main properties of mutually unbiased bases.

Definition 1 (Mutually unbiased) *Let $\{M_i\}_{i=1}^{s}$ be a set of Hermitian operators on a K-dimensional Hilbert space. Let the orthonormal basis associated with M_i be denoted as $|i, a\rangle$, with $a \in \mathcal{K}$. The set of operators is called mutually unbiased if*

$$\left|\langle i, a|j, b\rangle\right|^2 = \frac{1}{K} \quad \forall_{a,b\in\mathcal{K},\ i,j\in[s],\ i\neq j}. \tag{3}$$

In other words, the 'mutually unbiased' property means: if a system is in an eigenstate of M_i, and a measurement is done of some M_j with $j \neq i$, then there is no bias towards any of the possible outcomes.

For general K it is hard to determine how large the maximal set size s is. However, if K is a prime number, then a set of $K+1$ mutually unbiased operators exists. A construction is as follows. Let $|z\rangle$, with $z \in \mathcal{K}$, be the standard basis. In this basis define a diagonal operator Z and a 'rotation' operator X as

$$Z = \sum_{z\in\mathcal{K}} \omega^z |z\rangle\langle z| \qquad ; \qquad X = \sum_{z\in\mathcal{K}} |z+1\rangle\langle z|. \tag{4}$$

Here the numbers in the bra and ket brackets are taken modulo K. For $j \in \mathcal{K}$ we define operators $M_j = XZ^j$. Their eigensystems are given by

$$|j,a\rangle = \frac{1}{\sqrt{K}} \sum_{z\in\mathcal{K}} \omega^{-az} \omega^{jz(z-1)/2} |z\rangle \qquad ; \qquad M_j|j,a\rangle = \omega^a |j,a\rangle \tag{5}$$

where $a \in \mathcal{K}$. The set $\{M_j\}_{j=0}^{K-1}$ together with Z forms a set of $K+1$ mutually unbiased operators.

2.4 Gauss sums and Legendre symbols

The appearance in (5) of the z^2 in the exponent will lead to so-called Gauss sums.

Lemma 1 (Gauss sum) *Let p be a prime. Let s_p be defined such that $s_p = 1$ if $p = 1 \mod 4$ and $s_p = i$ if $p = 3 \mod 4$. Let $b \in \{1, \ldots, p-1\}$. Let L_p^b be the Legendre symbol (+1 if b is a square modulo p, -1 if it is not). Then*

$$\sum_{\ell=0}^{p-1} (\exp i\frac{2\pi}{p})^{b\ell^2} = s_p L_p^b \sqrt{p}. \tag{6}$$

Lemma 2 *For $s_p = 1$ it holds that $L_p^{-x} = L_p^x$, while for $s_p = i$ one has $L_p^{-x} = -L_p^x$.*

2.5 Averaging over the random challenge

In Section 3 we will need to compute expectation values over the randomly chosen challenge $|\psi\rangle$. This will be handled as follows. We pick an arbitrary orthonormal basis and expand $|\psi\rangle$ as

$$|\psi\rangle = \sum_{j\in\mathcal{K}} r_j e^{i\varphi_j} |j\rangle, \tag{7}$$

where the angles φ_j are uniformly drawn from $[-\pi, \pi)$, and the vector $\boldsymbol{r} = (r_j)_{j\in\mathcal{K}}$, with $r_j \geq 0$, is uniformly drawn from an orthant of the unit hypersphere $\sum_{j\in\mathcal{K}} r_j^2 = 1$. The r_j and the angles are all mutually independent.

Taking the expectation over $|\psi\rangle$ will be denoted as $\mathbb{E}_\psi$. The $\mathbb{E}_\psi$ can be split up into independent expectations, $\mathbb{E}_\psi = \mathbb{E}_{\boldsymbol{r}} \mathbb{E}_{\varphi_0} \cdots \mathbb{E}_{\varphi_{K-1}}$. We will use shorthand notation $\mathbb{E}_\varphi$ for the expectation over all the angles.

The following lemmas will help us to compute expectations.

Lemma 3 *Let $t \in \mathbb{Z}$ and $j \in \mathcal{K}$. Then $\mathbb{E}_\varphi[e^{it\varphi_j}] = \delta_{t0}$.*

Proof: For $t = 0$ we have $\mathbb{E}_\varphi[1] = 1$. For $t \neq 0$ we have $\mathbb{E}_\varphi[e^{it\varphi_j}] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}\varphi_j \; e^{it\varphi_j} = 0$. □

Lemma 4 (Dirichlet integral) *Let $\boldsymbol{v} = (v_j)_{j\in\mathcal{K}}$ be a vector, with $v_j > 0$ for all j. Then*

$$\int_0^1 \mathrm{d}^K p \; \delta(1 - \sum_{a\in\mathcal{K}} p_a) \boldsymbol{p}^{-1+\boldsymbol{v}} = B(\boldsymbol{v}) = \frac{\prod_{\alpha\in\mathcal{K}} \Gamma(v_a)}{\Gamma(\sum_{b\in\mathcal{K}} v_b)}. \tag{8}$$

The B is the generalized Beta function, and Γ is the Gamma function.

Lemma 5 *Let f be a function of $\boldsymbol{r}$. The expectation $\mathbb{E}_{\boldsymbol{r}}[f(\boldsymbol{r})]$ can be computed as*

$$\mathbb{E}_{\boldsymbol{r}}[f(\boldsymbol{r})] = \frac{1}{2^{-K}B(\mathbf{1}_K)} \int_0^1 \mathrm{d}^K r\ \delta(1-|\boldsymbol{r}|^2)\ \boldsymbol{r}^1 f(\boldsymbol{r}). \tag{9}$$

<u>*Proof:*</u> The f is integrated over the whole hypersphere orthant, with equal weight in every point, and with the correct area element $r\mathrm{d}r$ in each dimension. We can see as follows that the normalization is correct. We take $\mathbb{E}_{\boldsymbol{r}}[1]$, which leads to an integral $\int_0^1 \mathrm{d}^K r\ \delta(1-|\boldsymbol{r}|^2)\boldsymbol{r}^1$ proportional to the area of the hypersphere. We change integration variables to $p_a = r_a^2$, which gives $r_a \mathrm{d}r_a = \frac{1}{2}\mathrm{d}p_a$. Applying Lemma 4 we find the normalization $2^{-K}B(\mathbf{1}_K)$. □

Lemma 6 *Let $\boldsymbol{u} = (u_j)_{j\in\mathcal{K}}$ be a vector satisfying $u_j > -1$ for all j. Then*

$$\mathbb{E}_{\boldsymbol{r}}[\boldsymbol{r}^{2\boldsymbol{u}}] = \frac{B(\mathbf{1}_K + \boldsymbol{u})}{B(\mathbf{1}_K)}. \tag{10}$$

<u>*Proof:*</u> Follows directly from Lemmas 5 and 4. □

3 Security analysis

Now we determine the attacker's success probability in the model specified in Section 2.2. In Section 2.3 we established that it is optimal for the attacker if $\mathcal{B}$ consists of n mutually unbiased operators. In order to further create optimal attack conditions, we assume K to be prime; then, since we assume $n < K$, a mutually unbiased $\mathcal{B}$ can always be realized. We write $|b_{\alpha k_\alpha}\rangle = |g_\alpha, k_\alpha\rangle$, with $g_\alpha \in \mathcal{K}$ and $\alpha \neq \beta \implies g_\alpha \neq g_\beta$. The notation $|\cdot,\cdot\rangle$ was introduced in Section 2.3 for the mutually unbiased basis states.

3.1 The False Accept probability per quantum

For each of the attacker's response quanta independently there is a probability $P_{\psi\boldsymbol{k}}$ that the state will be projected to $|\omega\rangle$,

$$P_{\psi\boldsymbol{k}} = |\langle\omega|\hat{\omega}_{\boldsymbol{k}}\rangle|^2 = |\langle\psi|R^\dagger R|\hat{\psi}_{\boldsymbol{k}}\rangle|^2 = |\langle\psi|\hat{\psi}_{\boldsymbol{k}}\rangle|^2. \tag{11}$$

We have used the fact that R is unitary, i.e. $R^\dagger R = \mathbf{1}$. We define an averaged version of $\lambda_{\psi\boldsymbol{k}}$ as

$$P_{\text{av}} = \mathbb{E}_\psi \mathbb{E}_{\boldsymbol{k}|\psi} P_{\psi\boldsymbol{k}}. \tag{12}$$

Here $\mathbb{E}_\psi$ denotes the expectation value over the random challenge, and $\mathbb{E}_{\boldsymbol{k}|\psi}$ the expectation over the outcome $\boldsymbol{k}$ for given $|\psi\rangle$. The expected number of passing attacker quanta is nP_{av}.
In order to pass one round of the verification protocol, the adversary must have at least $\lceil(1-\varepsilon_{\text{noise}})n\rceil$ of his quanta project onto $|\omega\rangle$. The probability P_{pass1} of this happening is

$$P_{\text{pass1}} = \mathbb{E}_\psi \mathbb{E}_{\boldsymbol{k}|\psi} \sum_{u=\lceil(1-\varepsilon_{\text{noise}})n\rceil}^{n} \binom{n}{u} P_{\psi\boldsymbol{k}}^u (1-P_{\psi\boldsymbol{k}})^{n-u} \tag{13}$$

since the projections of the quanta are independent events, giving rise to a binomial distribution of the number of passed quanta. The adversary's overall probability of passing the whole protocol depends on the number of rounds: $P_{\text{pass1}}^{\#\text{rounds}}$. Hence, as long as P_{pass1} is sufficiently below 1, the overall False Accept probability can be made exponentially small. Using Jensen's inequality, we can bound (13) from above as

$$P_{\text{pass1}} \leq \sum_{u=\lceil(1-\varepsilon_{\text{noise}})n\rceil}^{n} \binom{n}{u} P_{\text{av}}^u (1-P_{\text{av}})^{n-u}. \tag{14}$$

Remark: When $P_{\rm av}$ becomes higher than $1-\varepsilon_{\rm noise}$, the attacker's overall success probability becomes non-negligible.

For proof-technical reasons we will concentrate on the quantity $P_{\rm av}$ instead of $P_{\rm pass1}$. Once we have an upper bound on $P_{\rm av}$, the inequality (14) allows us to obtain an upper bound on $P_{\rm pass1}$. Expanding $\mathbb{E}_{\boldsymbol{k}|\psi}$ in (12) and then $|\hat{\psi}_{\boldsymbol{k}}\rangle$, we write

$$\begin{aligned} P_{\rm av} &= \mathbb{E}_\psi \sum_{\boldsymbol{k}\in\mathcal{K}^n} |\langle\psi|\hat{\psi}_{\boldsymbol{k}}\rangle|^2 \prod_{\alpha\in[n]} |\langle b_{\alpha k_\alpha}|\psi\rangle|^2 \\ &= \mathbb{E}_\psi \sum_{\boldsymbol{k}\in\mathcal{K}^n} \frac{1}{\mathcal{N}_{\boldsymbol{k}}} \sum_{\beta,\gamma\in[n]} \langle\psi|b_{\beta k_\beta}\rangle\langle b_{\gamma k_\gamma}|\psi\rangle \prod_{\alpha\in[n]} |\langle b_{\alpha k_\alpha}|\psi\rangle|^2. \end{aligned} \tag{15}$$

We will upper bound (15) as follows. First we derive an upper bound $1/\mathcal{N}_{\boldsymbol{k}}$ which depends on K and n, but not on $\boldsymbol{k}$. The thus obtained upper bound on (15) can be drastically simplified, allowing for a final bounding using Cauchy-Schwartz.

3.2 Bound on the norm $\mathcal{N}_{\boldsymbol{k}}$

Theorem 1 *Let $n<\sqrt{K}$. The normalization constant $\mathcal{N}_{\boldsymbol{k}}$ as defined in (2) can be bounded as*

$$\frac{1}{\mathcal{N}_{\boldsymbol{k}}} \leq \frac{1}{n}\cdot\frac{1}{1-(n-1)/\sqrt{K}}. \tag{16}$$

Proof: We start from the definition $\mathcal{N}_{\boldsymbol{k}} = \sum_{\alpha\beta}\langle b_{\alpha k_\alpha}|b_{\beta k_\beta}\rangle$ and substitute $|b_{\alpha k_\alpha}\rangle = |g_\alpha, k_\alpha\rangle$, with the mutually unbiased basis states as defined in (5). We introduce shorthand notation $\triangle k_{\alpha\beta} = k_\alpha - k_\beta$ and $\triangle g_{\alpha\beta} = g_\alpha - g_\beta$.

$$\begin{aligned} \mathcal{N}_{\boldsymbol{k}} &= \frac{1}{K}\sum_{\alpha,\beta\in[n]}\sum_{\ell\in\mathcal{K}} \omega^{-\ell\triangle k_{\alpha\beta}}\omega^{\triangle g_{\alpha\beta}\ell(\ell-1)/2} \\ &= n+\frac{1}{K}\sum_{\substack{\alpha,\beta\in[n]\\ \alpha\neq\beta}} \omega^{-\frac{\triangle g_{\alpha\beta}}{2}(\frac{1}{2}+\frac{\triangle k_{\alpha\beta}}{\triangle g_{\alpha\beta}})^2} \sum_{\ell\in\mathcal{K}} \omega^{\frac{\triangle g_{\alpha\beta}}{2}[\ell-(\frac{1}{2}+\frac{\triangle k_{\alpha\beta}}{\triangle g_{\alpha\beta}})]^2}. \end{aligned} \tag{17}$$

We perform the ℓ-summation in (17) using Lemma 1 (Gauss sum). This yields

$$\mathcal{N}_{\boldsymbol{k}} = n+\frac{s_K}{\sqrt{K}}\sum_{\substack{\alpha,\beta\in[n]\\ \alpha\neq\beta}} L_K^{\triangle g_{\alpha\beta}/2}\omega^{-\frac{\triangle g_{\alpha\beta}}{2}(\frac{1}{2}+\frac{\triangle k_{\alpha\beta}}{\triangle g_{\alpha\beta}})^2}. \tag{18}$$

The notation s_K and L are explained in Lemma 1. In the Legendre symbol, the expression $\triangle g_{\alpha\beta}/2$ for odd $\triangle g_{\alpha\beta}$ should be read as $\triangle g_{\alpha\beta}\cdot 2^{-1} \mod K$. Using Lemma 2, we rewrite (18) as

$$\mathcal{N}_{\boldsymbol{k}} = n+\frac{2}{\sqrt{K}}\sum_{\substack{\alpha,\beta\in[n]\\ \alpha<\beta}} \mathrm{Re}\left[s_K L_K^{\triangle g_{\alpha\beta}/2}\omega^{-\frac{\triangle g_{\alpha\beta}}{2}(\frac{1}{2}+\frac{\triangle k_{\alpha\beta}}{\triangle g_{\alpha\beta}})^2}\right]. \tag{19}$$

At worst, every term in the summation is equal to -1. There are $n(n-1)/2$ terms. We conclude that $\mathcal{N}_{\boldsymbol{k}} \geq n - n(n-1)/\sqrt{K}$. □

Remark: The equal sign in Theorem 1 is attained when $n=1$.

3.3 Main result

Our main result is the theorem below.

Theorem 2 *Let $n<\sqrt{K}$. Then the quantity $P_{\rm av}$ is bounded as*

$$P_{\rm av} \leq \left[\frac{6(n-1)K}{(K+1)(K+2)}+\frac{2}{K+1}\right]\frac{1}{1-\frac{n-1}{\sqrt{K}}}. \tag{20}$$

Proof: We introduce the abbreviation $C_{Kn} = n^{-1}[1 - \frac{n-1}{\sqrt{K}}]^{-1}$, so that Theorem 1 is compactly written as $1/\mathcal{N}_{\boldsymbol{k}} \le C_{Kn}$. Since every term multiplying $1/\mathcal{N}_{\boldsymbol{k}}$ in (15) is nonnegative, we can write

$$\begin{aligned}
P_{\rm av} &\le C_{Kn}\mathbb{E}_\psi \sum_{\boldsymbol{k}\in\mathcal{K}^n} \sum_{\beta,\gamma\in[n]} \langle\psi|b_{\beta k_\beta}\rangle\langle b_{\gamma k_\gamma}|\psi\rangle \prod_{\alpha\in[n]} |\langle b_{\alpha k_\alpha}|\psi\rangle|^2 \\
&= C_{Kn} \sum_{\beta\in[n]} \sum_{k_\beta\in\mathcal{K}} \mathbb{E}_\psi |\langle\psi|b_{\beta k_\beta}\rangle|^4 \sum_{\boldsymbol{k}\setminus k_\beta} \prod_{\alpha\in[n]\setminus\{\beta\}} |\langle b_{\alpha k_\alpha}|\psi\rangle|^2 \\
&\quad + C_{Kn} \sum_{\substack{\beta,\gamma\in[n]\\ \beta\neq\gamma}} \sum_{k_\beta,k_\gamma} \mathbb{E}_\psi \langle\psi|b_{\beta k_\beta}\rangle\langle b_{\gamma k_\gamma}|\psi\rangle |\langle\psi|b_{\beta k_\beta}\rangle|^2 |\langle b_{\gamma k_\gamma}|\psi\rangle|^2 \cdot \\
&\quad \sum_{\boldsymbol{k}\setminus k_\beta,k_\gamma} \prod_{\alpha\in[n]\setminus\{\beta,\gamma\}} |\langle b_{\alpha k_\alpha}|\psi\rangle|^2 \\
&= C_{Kn} \sum_{\beta\in[n]} \sum_{k\in\mathcal{K}} \mathbb{E}_\psi |\langle\psi|b_{\beta k}\rangle|^4 \\
&\quad + C_{Kn} \sum_{\substack{\beta,\gamma\in[n]\\ \beta\neq\gamma}} \sum_{k,\ell\in\mathcal{K}} \mathbb{E}_\psi \langle\psi|b_{\beta k}\rangle\, |\langle\psi|b_{\beta k}\rangle|^2\, \langle b_{\gamma\ell}|\psi\rangle\, |\langle b_{\gamma\ell}|\psi\rangle|^2. \qquad (21)
\end{aligned}$$

In the last step we repeatedly used that $\sum_{k\in\mathcal{K}} |k\rangle\langle k| = 1$, followed by $\langle\psi|\psi\rangle = 1$. The first term in (21) is evaluated using Lemma 6, $\mathbb{E}_\psi[r^4_{\beta k}] = \frac{2}{K(K+1)}$. Thus we obtain

$$P_{\rm av} \le \frac{2nC_{Kn}}{K+1} + C_{Kn} \sum_{\substack{\beta,\gamma\in[n]\\ \beta\neq\gamma}} \sum_{k,\ell\in\mathcal{K}} \mathbb{E}_\psi \langle\psi|b_{\beta k}\rangle\, |\langle\psi|b_{\beta k}\rangle|^2\, \langle b_{\gamma\ell}|\psi\rangle\, |\langle b_{\gamma\ell}|\psi\rangle|^2. \qquad (22)$$

We consider the $\mathbb{E}_\psi$ in (22) as an inner product $\langle X, Y\rangle$ with $X = \langle b_{\beta k}|\psi\rangle\, |\langle\psi|b_{\beta k}\rangle|^2$ and $Y = \langle b_{\gamma\ell}|\psi\rangle\, |\langle b_{\gamma\ell}|\psi\rangle|^2$. Applying Cauchy-Schwartz, we have $\langle X, Y\rangle \le \sqrt{\langle X, X\rangle\langle Y, Y\rangle} = \sqrt{(\mathbb{E}_r r_k^6)(\mathbb{E}_r r_\ell^6)} = \frac{6}{K(K+1)(K+2)}$. The sum over k and ℓ gives rise to a factor K^2, and the sum over $\beta \neq \gamma$ gives a factor $n^2 - n$. □

We note the following:

- For $n = 1$ Theorem 2 reproduces the known result $2/(K + 1)$.
- Theorem 2 allows us to draw some conclusions about the speckle-based QR-PUF system of [15]. There the number of quanta is not fixed but Poisson-distributed, and the average number of photons $n_{\rm av}$ is known. For a powerful class of challenge estimation attacks (but not general challenge estimation attacks) the $P_{\rm av}$ was computed in [18], $P_{\rm av} \approx n_{\rm av}/(K+n_{\rm av})$ for $n_{\rm av} \ll K$. Theorem 2 with the substitution $n \to n_{\rm av}$ is consistent with that result, namely the adversary potentially has a higher probability of success when he is allowed more general attacks.
- The bound in (20) can probably be improved upon. First, we think that the value $\sqrt{K}$ in the condition $n < \sqrt{K}$ has no special meaning. It seems purely proof-technical. Given the result $\approx n/(K + n)$ for the quadrature based attack [18], we expect that a bound of order n/K can be derived for general challenge estimation attacks *without* the condition $n < \sqrt{K}$. Second, the factor '6' can probably be reduced. We suspect that our application of the Cauchy-Schwartz inequality does not give a tight bound.

4 Discussion

For a QR-PUF scheme with n quanta and a K-dimensional Hilbert space, we have derived an upper bound (Theorem 2) on the per-quantum success probability $P_{\rm av}$ of an adversary who does a generic challenge-estimation attack. From this result a bound is straightforwardly obtained on the adversary's per-round and overall success probability. Our bound is conservative (i.e. high) since we have assumed K to be a prime number, which guarantees the availability of a set of mutually unbiased measurements. For non-prime K the adversary may not have such an optimum set of measurements.

Our result reduces to the known bound $2/(K+1)$ for $n = 1$. The proof is more elegant than the proof in [14].

We expect that our bound can be made significantly tighter, and that the condition $n < \sqrt{K}$ is in fact not required for getting a bound of order n/K.

We realize that there is a whole class of attacks that we have not considered here, namely 'quantum' attacks involving additional particles that can be entangled with the challenge state, or involving 'partial' measurements. This is a topic for future work. Furthermore, another class of attacks that remains to be analyzed is the use of 'imperfect' unitary operations by the adversary (e.g. lossy and/or inaccurate). A first step in this direction was already made in [13] for $n = 1$, but this approach needs to be improved.

What we would like to have is a more complete security proof along the lines of the security proofs for Quantum Key Distribution, but under the condition that the adversary is not allowed to implement arbitrary unitaries in a lossless way. However, because of the differences between QR-PUFs and Quantum Key Distribution, most notably the fact that the adversary must *immediately* create a response state which immediately gets measured by the verifier, we suspect that such a full security proof will point to the challenge-estimation attack as the strongest possible attack.

Acknowledgments

We thank Berry Schoenmakers, Pepijn Pinkse and Allard Mosk for useful discussions.

References

[1] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld. Physical One-Way Functions. *Science*, 297:2026–2030, 2002.

[2] B. Gassend, D.E. Clarke, M. van Dijk, and S. Devadas. Silicon physical unknown functions. In *ACM Conference on Computer and Communications Security (CCS) 2002*, pages 148–160. ACM, 2002.

[3] J.D.R. Buchanan, R.P. Cowburn, A. Jausovec, D. Petit, P. Seem, G. Xiong, D. Atkinson, K. Fenton, D.A. Allwood, and M.T. Bryan. Forgery: 'fingerprinting' documents and packaging. *Nature, Brief Communications*, 436:475, 2005.

[4] P. Tuyls, G.J. Schrijen, B. Škorić, J. van Geloven, R. Verhaegh, and R. Wolters. Read-proof hardware from protective coatings. In *Cryptographic Hardware and Embedded Systems (CHES) 2006*, volume 4249 of *LNCS*, pages 369–383. Springer-Verlag, 2006.

[5] J. Guajardo, S.S. Kumar, G.J. Schrijen, and P. Tuyls. FPGA intrinsic PUFs and their use for IP protection. In *Cryptographic Hardware and Embedded Systems (CHES) 2007*, volume 4727 of *LNCS*, pages 63–80. Springer, 2007.

[6] G. DeJean and D. Kirovski. Radio frequency certificates of authenticity. In *IEEE Antenna and Propagation Symposium – URSI*, 2006.

[7] B. Škorić, T. Bel, A.H.M. Blom, B.R. de Jong, H. Kretschman, and A.J.M. Nellissen. Randomized resonators as uniquely identifiable anti-counterfeiting tags. *Secure Component and System Identification Workshop*, Berlin, March 2008.

[8] P. Tuyls, B. Škorić, and T. Kevenaar (Eds.). *Security with Noisy Data: Private Biometrics, Secure Key Storage and Anti-Counterfeiting.* Springer, London, 2007.

[9] A.-R. Sadeghi and D. Naccache (Eds.). *Towards Hardware-Intrinsic Security.* Springer, 2010.

[10] P. Tuyls, B. Škorić, S. Stallinga, A.H.M. Akkermans, and W. Ophey. Information-theoretic security analysis of physical uncloneable functions. In *9th Conf. on Financial Cryptography and Data Security*, volume 3570 of *LNCS*, pages 141–155. Springer, 2005.

[11] T. Ignatenko, G.-J. Schrijen, B. Škorić, P. Tuyls, and F.M.J. Willems. Estimating the Secrecy Rate of Physical Uncloneable Functions with the Context-Tree Weighting Method. In *Proc. IEEE International Symposium on Information Theory (ISIT) 2006*, pages 499–503, 2006.

[12] B. Škorić. On the entropy of keys derived from laser speckle; statistical properties of Gabor-transformed speckle. *Journal of Optics A: Pure and Applied Optics*, 10(5):055304–055316, 2008.

[13] B. Škorić. Quantum Readout of Physical Unclonable Functions. *International Journal of Quantum Information*, 10(1):1250001–1 – 125001–31, 2012.

[14] B. Škorić. Quantum Readout of Physical Unclonable Functions. `http://eprint.iacr.org/2009/369`.

[15] S.A. Goorden, M. Horstmann, A.P. Mosk, B. Škorić, and P.W.H. Pinkse. Quantum-Secure Authentication with a Classical Key. `http://arxiv.org/abs/1303.0142`, 2013.

[16] W.K. Wootters and W.H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.

[17] D. Dieks. *Phys. Lett. A*, 92:271, 1982.

[18] B. Škorić, A.P. Mosk, and P.W.H. Pinkse. Security of Quantum-Readout PUFs against quadrature-based challenge-estimation attacks. 2013. Accepted for publication in International Journal of Quantum Information. `http://eprint.iacr.org/2013/084`.

[19] I.D. Ivanović. Geometrical description of quantal state determination. *J. Phys. A*, 14(12):3241, 1981.

[20] W.K. Wootters and B.D. Fields. Optimal state-determination by mutually unbiased measurements. *Ann. Phys.*, 191:363–381, 1989.