Security Analysis of Identification Protocols based on quantum Physical Unclonable Functions

Frederick Hetherton

Master of Science
Cyber Security, Privacy and Trust
School of Informatics
University of Edinburgh
2020

Abstract

With the emergence of quantum technologies and ever-more sophisticated adversaries posing a growing threat to the security of existing cryptosystems, many schemes are being re-examined and revised for the quantum era. This work revisits the concept of identification (ID) protocols built upon Physical Unclonable Functions (PUFs). PUFs are hardware tokens that exploit arbitrary variations in microstructure caused by environmental and material fluctuations during manufacture, in order to display unique, highly unpredictable but repeatable response characteristics when challenged with a range of inputs. Such properties are considered prohibitively improbable to replicate, and thus the devices are considered to have distinct behavioural fingerprints, and regarded as 'unclonable' – defining their ability to serve as unique identifiers.

Recent technological advances and their future prospects pose a threat to this unclonability, and whilst quantum-enhanced PUF solutions with overlying protocols have been proposed in response, the process of proving security of a cryptographic protocol based upon quantum hardware properties is a vital component in developing a practical realisation. Previous work in the field has sought to analyse the security of PUF-based ID protocols against a variety of both practical and theoretical attacks; however, it has done so inconsistently, without a unified framework or set of definitions, with non-standard proof techniques, and with overly-specific threat models.

This work employs a quantum game-based security model to conduct analyses of PUF-based ID protocols in the presence of a general quantum adversary, reducing the notion of protocol security to a PUF's primary security property: unforgeability. We take a theoretical approach to investigating the security and efficiency of ID protocols based on a variety of PUF types; proving the superiority of 'quantum enhanced' protocols by both measures, as compared to their strictly classical analogue.

Keywords: *Physical unclonable functions, identification protocols, authentication, quantum cryptography, quantum cryptanalysis.*

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Frederick Hetherton*)

Acknowledgements

First and foremost, I would like to express sincere gratitude to my parents for their investment in my education, and the continued support from family and friends for my decisions that have led to this Master's degree. I recognise that this experience is a privilege not afforded to most, and for that I am thankful.

To my supervisor, Dr Delavar – thank you for your time, guidance and (most importantly) patience throughout this project. Your unfaltering enthusiasm to help, and respond to my incessant questioning during our weekly virtual meetings, is an attitude I could not have completed this project without.

Finally, to my flatmates, who had the misfortune to be confined with me for 4 months during the Covid-19 lockdown. You had a tough gig with me around, but you played the part well – thank you for putting up with me, and seeing me through to project submission with a decent measure of remaining sanity.

Table of Contents

Chapter 1

Introduction

1.1 Background

Physical Unclonable Functions (PUFs) are queryable devices that exploit unpredictable variations in microstructure to establish a distinctive behavioural fingerprint. Typically, PUFs are prohibitively hard to clone and have a unique challenge-response behaviour due to the random physical disorders that arise during manufacture. The core concept is that when presented with a challenge the PUF outputs a response, such that for a range of challenges the outputs appear random. The result is a hardware token able to serve as a unique identifier, characterised by a distinct, unpredictable set of challenge-response pairs (CRPs). An example is an optical PUF using the angle of incidence of a laser onto a glass token as a challenge and the resulting scatter pattern as the response. Aspects of the glass manufacturing process (such as the precise temperature, air pressure and mineral composition) are deemed uncontrollable with respect to producing the exact glass crystal formation – which uniquely determines the scatter pattern.

Identification (ID) protocols are methods of authenticating a device or individual for access to a privileged system. In general, ID protocols rest upon the notion of a *shared secret* between the trusted verifier and parties with a legitimate right to access. In the physical world, this might correspond to something you own (a key), something you know (a password) or something you are (a biometric). A PUF's ability to serve as a unique identifier makes it ideal for the basis of an ID protocol – the shared secret being the response to a randomly chosen challenge. Legitimate access is then defined by whether you can prove you are the PUF holder. In practice, the PUF is first *enrolled* with the verifier, allowing them to sample CRPs and later choose a random entry to challenge a party for the correct response when they request authentication.

1.2 Motivation

For efficiency's sake, modern cryptosystems tend not to provide information-theoretic security but computational security against realistic adversaries – requiring them to be based on hardness assumptions. However, quantum technologies are projected to emerge that will be capable of cracking almost all existing widespread cryptosystems by providing the ability to run quantum algorithms which break certain hardness assumptions with efficiency. Algorithms such as Shor's, Grover's and Simon's pose up to exponential speed-up from classical efforts in cases, and have been demonstrated to dismantle both public and private key cryptosystems with relative ease [2, 10, 13, 22, 30]. As a result, research has placed great emphasis on 'quantum-proofing' cryptosystems and their hardness assumptions for the quantum era [27].

In this project, the hardness assumption upon which PUFs are built is known as *unforgeability*: the notion that it should be infeasible to present a valid response for a chosen challenge without current PUF access or prior learning of the respective CRP. Several PUF types have been shown experimentally to be vulnerable to a variety of man-in-the-middle, impersonation and machine-learning attacks [7, 28] that break PUF unforgeability. In response, quantum-enhanced PUF solutions have been proposed to allow PUFs to be challenged and read out using quantum states rather than classical bit-strings [3, 24] – these devices claim to provide enhanced security by protecting against many of the attacks that classical PUFs are vulnerable to. Nonetheless, evidence of advantage to security or efficiency is paramount in motivating their manufacture and integration into protocols, especially since attacks feasible for a quantum adversary have been projected to yield success on stronger PUF types. Such attacks exploit laws of quantum mechanics to leverage PUF systems and leak more information than an adversary otherwise could [8, 15, 23, 29]. It is essential that we prove the security of quantum-enhanced PUF protocols against general attacks, and investigate additional advantages. Previous works have contributed a great deal to this effort [5, 18, 20, 24], however, they are inconsistent in approach, adopting non-standard proof techniques and overly-specific threat models [4]. To our knowledge, the literature lacks a unified approach with solid definitions to analyse security of protocols based on a variety of PUF types against a general quantum adversary.

We adopt a game-based security framework to analyse security of ID protocols based on three PUF types: classical (cPUF), quantum-readout (QR-PUF), and quantum (qPUF). Security is evaluated against a general quantum adversary to prove verifiable advantages in adopting quantum-enhanced PUF solutions. Protocol efficiency will also be investigated to provide further material for comparative evaluation.

1.3 Thesis Overview

Modelling PUFs and Unforgeability

We begin by laying theoretical foundations through modelling PUF manufacture and interaction in Chapter 2, defining the cPUF, QR-PUF and qPUF. Our model introduces unforgeability, what we deem to be the primary security property of PUFs, in Chapter 3. We define such terms as *quantum unconditional*, *existential* and *selective* unforgeability, each describing a different threat model. Chapter 3 concludes with unforgeability results: that quantum unconditional and existential unforgeability are too strong for our PUFs to fulfil, whilst all three are quantum selectively unforgeable. With these results we explore security of the ID protocols defined in the following Chapter.

PUF-based Identification Protocols

Chapter 4 sets out the ID protocols of interest for this work, each with an increasing degree of quantum enhancement. The first is a cPUF-based ID protocol, for which CRPs are composed of classical bit strings, with the protocol assuming classical honest behaviour. The second protocol is adapted from [24], based on a QR-PUF: it assumes underlying classical hardware, whilst honest parties are capable of querying the PUF with quantum-encoded challenges, and receiving similar responses. The final protocol is based on a qPUF, for which challenges are not restricted to encodings of classical bit strings. In this case, the verifier stores CRPs in a classical / quantum hybrid database.

Protocol Analysis

We begin analysis in Chapter 5 by exploring security of each protocol with reference to the PUF unforgeability property, through a game-based security framework with a varying threat model. Our primary analyses conclude that, via reduction of protocol security to selective unforgeability of the underlying PUF, we can prove security of all three protocols against a general quantum adversary. To demonstrate a scenario in which qPUFs present an advantage, we discuss how qPUF ID protocol security can be maintained if the verifier reuses CRPs for multiple protocol rounds whilst QR-PUF and cPUF protocol security is broken. We continue analyses to evaluate protocol efficiency in terms of the storage, communication and computation resources required – helping to differentiate where one protocol may be beneficial over another in the case that they provide similar security. We find that of our original protocols cPUF yields the greatest efficiency, but by reusing CRPs we can reduce qPUF enrollment costs to significantly lower than possible with a cPUF.

Discussion and Conclusion

We begin our discussion in Chapter 6 by bringing together results for comparative protocol assessment; declaring qPUF security and efficiency advantages by revising the ID protocol to reuse CRPs. We continue with a critical evaluation of our model, highlighting its limitations and posing how it might be improved upon to broaden applicability. The work concludes with an outline of project contributions, its significance in a wider context, and possible future research directions.

1.4 Related Work

Critical sources from which we develop theory are three papers [3, 4, 11] from our own institution, which inspired this project. From them we adopt a quantum game-based security modelling approach, the notion of PUF unforgeability, and the concept of a qPUF coupled with a state equality testing algorithm on which to base an enhanced ID protocol. We build on wide a variety of other crucial research, as outlined below.

The original cPUF ID protocol was proposed in [17] and revisited in [9], adapted in this work as the first of three protocols. As the first of its kind, there are a number of papers focused on analysing its security: a first paper of interest, [20], models attacks against PUFs; [21] presents attack models and security evaluations; [28] discusses hybrid side-channel/machine-learning attacks; and [23] discusses general cryptographic protocols under quantum attacks. To prove protocol security, we first prove cPUF unforgeability by adapting the definition of blinded unforgeability introduced in [1]. Our QR-PUF protocol is inspired by those proposed in [24], security of which was later evaluated in reviews by the original author [25, 26]. In addition, [8, 29] explore investigate how PUF-based quantum authentication systems can be leveraged by superposition and cloning attacks respectively. The qPUF protocol has been proposed in [11] with variations differing with regards to available resources.

An alternative framework with which to investigate PUF ID protocol security takes a quantitative approach to characterising PUFs, through defining robustness and unclonability properties, as a basis for protocol security [12]. However, due to the scope of this project, it will not be explored – we leave this for future work. Further discussions concerning a unified framework for analyses have been published in [5, 18].

1.5 Preliminaries

This work is aimed at readers with a general familiarity with the field of quantum information and an interest in cryptographic applications, though to improve accessibility we have endeavoured to introduce relevant non-trivial concepts below. For readers with a lesser familiarity, preliminary understanding at a more fundamental level may be required. For reference, we advise that [14] provides a comprehensive coverage.

1.5.1 Quantum Information

State Representation

We adopt the usual Dirac notation to represent quantum states as vectors in a Hilbert space $\mathcal{H}^D$: $|\psi\rangle$ represents a vector, with $\langle\psi|$ its dual – for which coefficients are complex conjugates of those of $|\psi\rangle$. E.g. for $|\psi\rangle \in \mathcal{H}^2$ and $|0\rangle, |1\rangle$ computational basis states of $\mathcal{H}^2$: $|\psi\rangle = a|0\rangle + b|1\rangle \implies \langle\psi| = a^* \langle 0| + b^* \langle 1|$. In matrix notation:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \qquad |\psi_1\rangle = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \quad \langle\psi_2| = \begin{bmatrix} a_2^* & b_2^* \end{bmatrix}$$

Operators

Action of a system on a state can be described by an operator, corresponding to a linear map, represented by a complex valued matrix; with the adjoint (†) of an operator defined transpose and conjugate element-wise. An operator U is *unitary* if $UU^\dagger = \mathbb{I}$.

State Equality

Definition 1 (Fidelity)**.** *The* fidelity *of two pure quantum states* $|\psi\rangle, |\phi\rangle$ *is a measure of their "closeness", defined to be* $F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$.

Definition 2 (Distinguishability)**.** *For* $0 \leq \mu, \nu \leq 1$, *two quantum states* $|\psi\rangle$ *and* $|\omega\rangle$ *are said to be* μ*-distinguishable if* $0 \leq F(|\psi\rangle, |\omega\rangle) \leq 1 - \mu$, *and* ν*-indistinguishable* *if* $\nu \leq F(|\psi\rangle, |\omega\rangle) \leq 1$*. Two states* $|\psi\rangle$ *and* $|\omega\rangle$ *are completely distinguishable or 1-distinguishable* $(\mu = 1)$*, if* $F(|\psi\rangle, |\omega\rangle) = 0$*, and completely indistinguishable or 1-indistinguishable* $(\nu = 1)$ *if* $F(|\psi\rangle, |\omega\rangle) = 1$.

The following is an abstraction of specific tests for state equality, adapted from [3, 4, 11]:

Definition 3 (Quantum Testing Algorithm)**.** *Let* $|\psi\rangle^{\otimes\kappa_1}$ *and* $|\phi\rangle^{\otimes\kappa_2}$ *be* κ_1 *and* κ_2 *copies of two pure quantum states* $|\psi\rangle$ *and* $|\phi\rangle$*, respectively. A quantum testing algorithm* $\mathcal{T}$ *is a quantum algorithm that takes as input the tuple* $(|\psi\rangle^{\otimes\kappa_1}, |\phi\rangle^{\otimes\kappa_2})$ *and some ancilla states and accepts* $|\psi\rangle$ *and* $|\phi\rangle$ *as equal (outputs 1) with the following probability:*

$$\Pr\left[1 \leftarrow \mathcal{T}(|\psi\rangle^{\otimes\kappa_1}, |\phi\rangle^{\otimes\kappa_2})\right] = 1 - \Pr\left[0 \leftarrow \mathcal{T}(|\psi\rangle^{\otimes\kappa_1}, |\phi\rangle^{\otimes\kappa_1})\right] = f(\kappa_1, \kappa_2, F(|\psi\rangle, |\phi\rangle))$$

where $f(\kappa_1, \kappa_2, F(|\psi\rangle, |\phi\rangle))$ *satisfies the following limits:*

$$\begin{cases} \lim_{F(|\psi\rangle,|\phi\rangle)\to 1} f(\kappa_1, \kappa_2, F(|\psi\rangle, |\phi\rangle)) = 1 & \forall(\kappa_1, \kappa_2) \\ \lim_{\kappa_1=1,\kappa_2\to\infty} f(\kappa_1, \kappa_2, F(|\psi\rangle, |\phi\rangle)) = F^2(|\psi\rangle, |\phi\rangle) & \\ \lim_{\kappa_1\to\infty,\kappa_2=1} f(\kappa_1, \kappa_2, F(|\psi\rangle, |\phi\rangle)) = F^2(|\psi\rangle, |\phi\rangle) & \\ \lim_{F(|\psi\rangle,|\phi\rangle)\to 0} f(\kappa_1, \kappa_2, F(|\psi\rangle, |\phi\rangle)) = \varepsilon(\kappa_1, \kappa_2) & \end{cases}$$

for $\varepsilon(\kappa_1, \kappa_2)$ *the statistical error of the test algorithm.*

1.5.2 Complexity

O notation is a way of describing mathematically the asymptotic growth in complexity of an algorithm or expression in terms of the input variables, described as follows:

Definition 4. *For a given function* $g(n)$*, we denote by* $O(g(n))$ *the set of functions*

$$O(g(n)) = \{f(n) : \text{there exist positive constants } c \text{ and } n_0 \text{ such that}$$
$$0 \leq f(n) \leq cg(n) \text{ for all } n \geq n_0\}.$$

$f(n) = O(g(n))$ says that f grows *with the order of* $g(n)$; e.g. $f(n) = O(1) \implies f(n)$ grows constantly with n, $f(n) = O(n^2) \implies f(n)$ has quadratic growth with n.

Definition 5. *A function* $g(n) : \mathbb{R} \to \mathbb{R}$ *is* polynomial *iff:* $g(n) = \sum_{m=0}^{k} a_m n^m : a_m \in \mathbb{R}$.

We say an algorithm or expression $f(n)$ has complexity polynomial in n, $f(n) = \text{poly}(n)$, if $f(n) = O(g(n))$ for $g(n)$ a polynomial function.

Definition 6. *A function* $g(n) : \mathbb{N} \to \mathbb{R}_{>0}$ *is* negligible $(g(n) = negl(n))$ *iff:*

$$\forall c > 0, \exists N_c \in \mathbb{N} \text{ such that } \forall n > N_c \text{ it holds that } g(n) < \frac{1}{n^c}.$$

I.e. $g(n)$ is negligible if it decreases with n faster than any inverse polynomial [23].

Chapter 2

Modelling PUFs

In this chapter we model the three PUF types upon which the protocols in Chapter 4 will be based: classical, quantum-readout and quantum. In general, each will have their manufacture and interaction behaviour modelled by algorithms 'Gen' and 'Eval' respectively. The PUF security parameter λ determines the 'size' of the CRP sample space, and thus the security of the PUF in a rough sense.

2.1 Classical PUF (cPUF)

cPUFs are PUFs for which CRPs consist of classical bits of information, and we restrict honest parties to interacting with cPUFs in this manner. However, for modelling in the context of security, it's insufficient to assume that cPUFs prohibit quantum interaction entirely. We suppose that an adversary in possession of a classical PUF is capable of exploiting its physical properties by exposing it to conditions in which it behaves quantumly – this represents a stronger, potentially more realistic threat model in the quantum era. To model this, we introduce two classical PUF evaluation algorithms: type I, which cannot be queried with quantum states; and type II, which can only be queried with quantum states. Type I cPUF evaluation will be employed in the PUF-based ID protocol, for which we assume all parties are honest and classical; type II will be used to model attack scenarios with a quantum adversary.

Definition 7 (cPUF Generation)**.** *The generation process of a* cPUF *is formalised by an algorithm* cGen *that takes a security parameter* λ *as input and generates a* cPUF *with a unique identifier i:*

$$\text{cPUF}_i \leftarrow \text{cGen}(\lambda).$$

Definition 8 (Type I cPUF Evaluation)**.** *The type I evaluation process of a* cPUF *with identifier i is formalised by an algorithm* $\mathrm{cEval_I} : \{0,1\}^{\lambda} \rightarrow \{0,1\}^{\lambda}$ *that maps a challenge input* $\mathbf{c}_j$ *to the corresponding response* $\mathbf{r}_{ij}$*:*

$$\mathbf{r}_{ij} \leftarrow \mathrm{cEval_I}(\mathrm{cPUF}_i, \mathbf{c}_j).$$

Requirements 1 (Classical Robustness, Uniqueness & Collision-resistance)**.** *The following are vital to prove correctness of any cryptographic scheme built upon PUFs with classical evaluation:*

i) ***Robustness*** *– To place a bound on noise, if* $\mathrm{cEval_I}(\mathrm{cPUF}_i, \mathbf{c}_j)$ *is run a number of times, the maximum distance between responses must at most be* δ_r*.*

ii) ***Uniqueness*** *– To be able to distinguish individual cPUFs from a family of* cPUF_i *created by the same cGen algorithm, and ran through* $\mathrm{cEval_I}$ *with the same challenge* $\mathbf{c}_j$*, the minimum distance between responses must be at least* δ_u*.*

iii) ***Collision-resistance*** *– Whenever* $\mathrm{cEval_I}$ *is run on a* cPUF_i *with multiple distinct challenges, the minimum distance between responses must be at least* δ_c*.*

The parameters $\delta_r, \delta_u, \delta_c$ *are determined by the security parameter* λ *and must satisfy the conditions* $\delta_r \leq \delta_u, \delta_r \leq \delta_c$ *to allow for distinguishing distinct challenges and PUFs.*

Definition 9 (Type II cPUF Evaluation)**.** *The type II evaluation process of a* cPUF *with identifier i is formalised by an algorithm* $\mathrm{cEval_{II}} : \mathcal{H}^D \rightarrow \mathcal{H}^D$*, where* $D = 2^{\lambda}$*, that maps a general superposition of quantum challenge states* $|c_j\rangle = \Sigma_n \alpha_n |c_n\rangle$ *to the corresponding superposition of responses* $|r_{ij}\rangle = \Sigma_n \alpha_n |r_{in}\rangle$*, for* $|c_n\rangle, |r_{in}\rangle \in \mathcal{H}^D$*:*

$$\Sigma_n \alpha_n |r_{in}\rangle \leftarrow \mathrm{cEval_{II}}(\mathrm{cPUF}_i, \Sigma_n \alpha_n |c_n\rangle).$$

Finally, after laying out these definitions and requirements, we can define a cPUF as follows:

Definition 10 (Classical Physical Unclonable Function)**.** *For* λ *the security parameter, and* $\delta_r, \delta_u, \delta_c \in [0,1]$ *the robustness, uniqueness and collision-resistance thresholds, a* $(\lambda, \delta_r, \delta_u, \delta_c)$-cPUF *is an instance of the tuple of algorithms*

$$\mathrm{cPUF} = (\mathrm{cGen}, \mathrm{cEval_I}, \mathrm{cEval_{II}})$$

defined according to Definitions 7, 8 and 9, satisfying Requirements 1.

If a particular cPUF is known to prohibit quantum queries in any physical environment, then it may be modelled as cPUF $= (\text{cGen}, \text{cEval}_{\text{I}}, \mathsf{null})$. However, throughout our analyses, we assume a quantum adversary can always leverage a cPUF to respond to quantum queries, and an honest party will only interact with it classically.

2.2 Quantum Readout PUF (QR-PUF)

The QR-PUF primitive is essentially a classical PUF that is read out using quantum states; the PUF is challenged from a space of classical bit strings encoded as pure quantum states, and responds similarly. We consider the more general case and model QR-PUFs as PUFs that can be challenged with any state in $\mathcal{H}^D$. Allowing an honest party this method of interaction with a PUF provides enhanced security through protection against eavesdroppers: third parties are unable to decode challenge states to their classical description without a super-polynomial number of state copies, and obtaining even partial descriptions disturbs honest transmission through state measurement – limiting their ability to characterise the PUF through learning CRPs. We assume, however, that classical descriptions of all response states from QR-PUFs are public, as in [24] – modelled by assuming that bases of the PUF Hilbert space are public.

Definition 11 (QR-PUF Generation)**.** *The generation process of a* QR-PUF *is formalised by an algorithm* qrGen *that takes a security parameter λ as input and generates a* QR-PUF *with a unique identifier i, along with bases B_i of the* QR-PUF$_i$ *Hilbert space $\mathcal{H}^D$:*

$$\{\text{QR-PUF}_i, B_i\} \leftarrow \text{qrGen}(\lambda).$$

Definition 12 (QR-PUF Evaluation)**.** *The evaluation process of a* QR-PUF *with identifier i is formalised by an algorithm* qrEval $: \mathcal{H}^D \to \mathcal{H}^D$*, where $D = 2^\lambda$, that maps a challenge state $|\psi_j\rangle$ to the corresponding response $|\omega_{ij}\rangle \in \mathcal{H}^D$:*

$$|\omega_{ij}\rangle \leftarrow \text{qrEval}(\text{QR-PUF}_i,\, |\psi_j\rangle).$$

Requirements 2 (Quantum Robustness, Uniqueness & Collision-resistance)**.** *The following are vital to prove correctness of any cryptographic scheme built upon PUFs with quantum evaluation:*

i) ***Robustness*** *– For any* PUF$_i$ *with quantum evaluation and any two input states $|c\rangle$ and $|\hat{c}\rangle$ that are δ_r-indistinguishable, the corresponding output quantum states*

$|r\rangle$ *and* $|\hat{r}\rangle$ *are also* δ_r*-indistinguishable with overwhelming probability,*

$$\Pr[\delta_r \leq F(|r\rangle, |\hat{r}\rangle) \leq 1] = 1 - negl(\lambda).$$

***ii) Uniqueness** – For any two PUFs with quantum evaluation generated by a* Gen *algorithm, i.e.* PUF_i *and* PUF_j*, the corresponding CPT map models, i.e.* Λ_i *and* Λ_j *are* δ_u*-distinguishable with overwhelming probability,*

$$\Pr\left[||(\Lambda_i - \Lambda_j)_{i \neq j}||_\diamond \geq \delta_u\right] = 1 - negl(\lambda).$$

***iii) Collision-resistance** – For any* PUF_i *with quantum evaluation and any two input states* $|c\rangle$ *and* $|\hat{c}\rangle$ *that are* δ_c*-distinguishable, the corresponding output states* $|r\rangle$ *and* $|\hat{r}\rangle$ *are also* δ_c*-distinguishable with overwhelming probability,*

$$\Pr[0 \leq F(|r\rangle, |\hat{r}\rangle) \leq 1 - \delta_c] = 1 - negl(\lambda)$$

We can then formally define a QR-PUF as follows:

Definition 13 (Quantum Readout Physical Unclonable Function)**.** *For* λ *the security parameter, and* $\delta_r, \delta_u, \delta_c \in [0,1]$ *the robustness, uniqueness and collision resistance thresholds: A* $(\lambda, \delta_r, \delta_u, \delta_c)$-QR-PUF *is an instance of the pair of algorithms*

$$\text{QR-PUF} = (\text{qrGen}, \text{qrEval})$$

defined according to Definitions 11 and 12, and satisfying Requirements 2.

Whilst we omit the preliminary discussion from [24], we assume that QR-PUF interaction can be modelled by a unitary operator $\mathrm{U}_{\text{QR-PUF}}$. It is proven in [4] that a unitary is sufficient to satisfy Requirements 2 for any δ_r, δ_u and δ_c, due to the distance-preserving property of unitary transformations – so we drop $(\delta_r, \delta_u, \delta_c)$ from notation.

2.3 Quantum PUF (qPUF)

The QR-PUF can be thought of as a special case of a more general variety of PUF, the qPUF. Generally, qPUFs do not restrict the CRP space to quantumly-encoded classical bit strings. Instead, a CRP can be any pair of pure quantum states in the corresponding Hilbert space – so long as they satisfy a number of minimum distance and distinguishability requirements. In contrast to QR-PUF, classical descriptions of qPUF response states are assumed unknown to all parties – just as with challenge states.

We now adapt the formal manufacture and evaluation models to the case of qPUFs:

Definition 14 (qPUF Generation)**.** *The generation process of a* qPUF *is formalised by an algorithm* qGen *that takes a security parameter* λ *as input and generates a* qPUF *with a unique identifier i:*

$$\mathrm{qPUF}_i \leftarrow \mathrm{qGen}(\lambda).$$

Definition 15 (qPUF Evaluation)**.** *The evaluation process of a* qPUF *with identifier i is formalised by an algorithm* qEval : $\mathcal{H}_{in} \rightarrow \mathcal{H}_{out}$*, that maps challenge quantum states* $|\psi_j\rangle$ *to response states states* $|\omega_{ij}\rangle$*:*

$$|\omega_{ij}\rangle \leftarrow \mathrm{qEval}(\mathrm{qPUF}_i, |\psi_j\rangle).$$

As well as adapting existing classical algorithms, we need an algorithm to efficiently test equality between two unknown quantum states. This is a necessary addition with respect to cPUF and QR-PUF primitives – CRPs now take on a more general quantum form, with unknown classical descriptions. We employ a testing algorithm $\mathcal{T}$ as in [4] (Definition 3). We then formally define a qPUF as follows:

Definition 16 (Quantum Physical Unclonable Function)**.** *For* λ *the security parameter, and* $\delta_r, \delta_u, \delta_c \in [0,1]$ *the robustness, uniqueness and collision resistance thresholds: A* $(\lambda, \delta_r, \delta_u, \delta_c)$-qPUF *is an instance of the tuple of algorithms*

$$\mathrm{qPUF} = (\mathrm{qGen}, \mathrm{qEval}, \mathcal{T})$$

defined according to definitions 14, 15, 3 respectively, and satisfying Requirements 2.

For the remainder of this work, we narrow our focus and assume qEval to be a unitary transformation, such that the respective qPUF is unitary. This is sufficient for an initial study, since qPUFs modelled by a unitary transformation satisfy Requirements 2 for any δ_r, δ_u and δ_c [3]. We drop $(\delta_r, \delta_c, \delta_u)$ from notation, and $\mathcal{H}_{in}^d$ and $\mathcal{H}_{out}^d$ become $\mathcal{H}^D$, where $D = 2^\lambda$. Finally, we formally define unitary qPUFs:

Definition 17 (Unitary qPUF)**.** *A unitary* qPUF *is an instance of the tuple of algorithms*

$$\mathrm{qPUF} = (\mathrm{qGen}, \mathrm{qEval}, \mathcal{T})$$

defined according to Definitions 14, 15 and 3 respectively, such that qEval *can be modelled by a unitary transformation* $\mathrm{U}_{\mathrm{qPUF}}$ *over a D-dimensional Hilbert space,* $\mathcal{H}^D$*, operating on pure quantum states* $|\psi\rangle \in \mathcal{H}^D$ *and returning pure outputs* $|\omega\rangle \in \mathcal{H}^D$*,*

$$|\omega_{ij}\rangle = \mathrm{qEval}(\mathrm{qPUF}_i, |\psi_j\rangle) = \mathrm{U}_{\mathrm{qPUF}_i} |\psi_j\rangle.$$

Chapter 3

Unforgeability of PUFs

In the following we define the unforgeability property for each of the PUF types defined in Chapter 2, presenting some relevant results. Generally speaking, a PUF is *unforgeable* if it is infeasible to produce $n+1$ legitimate CRPs from n prior PUF queries. Our definitions of unforgeability assume threat from a general quantum adversary, both polynomially-bounded and unbounded, with superposition access to the PUF before being challenged. We take unforgeability to be the primary security property of PUFs, and later use it to form the basis of PUF-based ID protocol security.

3.1 Game-based Framework

We model unforgeability of a PUF through means of a security game $\mathcal{G}_{c,\mu}^{\mathrm{PUF}}(\mathcal{A},\lambda)$ between a challenger $\mathcal{C}$ and a quantum adversary $\mathcal{A}$, where $c \in \{\mathsf{qEx}, \mathsf{qSel}\}$ denotes the type of challenge phase (described below) and PUF $\in$ {cPUF, QR-PUF, qPUF}. Our framework captures three threat models, each illustrating a different notion of unforgeability: *quantum unconditional unforgeability* defines the scenario in which $\mathcal{A}$ is an unbounded quantum adversary, and chooses the challenge to which it must respond (a 'qEx' challenge); *quantum existential unforgeability* is similar in the sense that $\mathcal{A}$ still chooses the challenge, except they are polynomially-bounded in their computational power; and *quantum selective unforgeability* also concerns a polynomially-bounded $\mathcal{A}$, whilst the challenge is chosen by $\mathcal{C}$ (a 'qSel' challenge).

In the following, $|x\rangle \in \mathcal{H}^D$ is the quantum encoding of a classical challenge / response $\mathbf{x} \in \{0,1\}^{\lambda}$, where λ is the PUF security parameter and $D = 2^{\lambda}$. For a set of states S, $|\psi\rangle \notin_{\mu} S$ denotes a state $|\psi\rangle$ that is at least μ-distinguishable from all states $|\omega\rangle \in S$.

cPUF Unforgeability Game

Let us assume the existence of a cPUF $= (\mathrm{cGen}, \mathrm{cEval}_{\mathrm{I}}, \mathrm{cEval}_{\mathrm{II}})$ according to Definition 10. Recall $\mathcal{C}$ interacts with cPUF only through cEval$_{\mathrm{I}}$ as an honest classical party, whilst $\mathcal{A}$ queries cPUF in superposition through cEval$_{\mathrm{II}}$ in the learning phase – for which CRPs are assumed to be encoded in the computational basis.

$$\mathcal{G}_c^{\mathrm{cPUF}}(\mathcal{A}, \lambda)$$

Setup

- The challenger $\mathcal{C}$ selects λ and runs cGen(λ) to build an instance of the cPUF family, cPUF$_i$. Then, $\mathcal{C}$ reveals λ, the Hilbert space $\mathcal{H}^D$, and the identifier of cPUF$_i$, i to the adversary, $\mathcal{A}$. The challenger creates two databases, S_{in} and S_{out} which are initially empty and shares them with the adversary $\mathcal{A}$.

Learning – For $j = 1 : k$, where $k \in O(\mathrm{poly}(\lambda))$:

- $\mathcal{A}$ prepares a superposition of challenge states $|\psi_j\rangle = \Sigma_n \alpha_n |c_n\rangle \in \mathcal{H}^D$, appends $|\psi_j\rangle$ to S_{in}, and sends $|\psi_j\rangle$ to $\mathcal{C}$;
- $\mathcal{C}$ runs cEval$_{\mathrm{II}}$(cPUF$_i$, $|\psi_j\rangle$) to obtain $|\omega_{ij}\rangle = \Sigma_n \alpha_n |r_{in}\rangle \in \mathcal{H}^D$, and sends $|\omega_{ij}\rangle$ to $\mathcal{A}$;
- $\mathcal{A}$ appends $|\omega_{ij}\rangle$ to S_{out}.

Challenge – Let c show the type of the challenge phase.

- If $c = \mathsf{qEx}$: $\mathcal{A}$ chooses a classical bit string $\mathbf{c}^* \in \{0,1\}^\lambda$, such that $|c^*\rangle \notin_\mu S_{in}$, and sends it to $\mathcal{C}$;
- If $c = \mathsf{qSel}$: $\mathcal{C}$ chooses a classical bit string $\mathbf{c}^* \in \{0,1\}^\lambda$ uniformly at random, and sends it to $\mathcal{A}$.

Guess

- $\mathcal{A}$ sends their guess $\hat{\mathbf{r}}$ on the output of cEval$_{\mathrm{I}}$(cPUF$_i, \mathbf{c}^*$) to $\mathcal{C}$;
- $\mathcal{C}$ computes $\mathbf{r}^* = \mathrm{cEval}_{\mathrm{I}}(\mathrm{cPUF}_i, \mathbf{c}^*)$;
- $\mathcal{C}$ compares $\hat{\mathbf{r}}$ and $\mathbf{r}^*$, outputting 0 if they are different and 1 if they are the same. $\mathcal{A}$ wins the game iff $b = 1$.

QR-PUF Unforgeability Game

Let QR-PUF = (qrGen, qrEval) be defined as in Definition 10.

$$\mathcal{G}^{\text{QR-PUF}}_{c,\mu}(\mathcal{A},\lambda)$$

Setup

- The challenger $\mathcal{C}$ selects λ and runs qrGen(λ) to build an instance of the QR-PUF family, QR-PUF$_i$. Then, $\mathcal{C}$ reveals λ, the identifier i of QR-PUF$_i$, and the Hilbert space $\mathcal{H}^D$ along with bases B_i to the adversary, $\mathcal{A}$. The challenger creates two databases, S_{in} and S_{out} which are initially empty and shares them with the adversary $\mathcal{A}$.

Learning – For $j = 1 : k$, where $k \in O(\text{poly}(\lambda))$:

- $\mathcal{A}$ prepares a quantum state $|\psi_j\rangle \in \mathcal{H}^D$, appends the classical description of $|\psi_j\rangle$ to S_{in} and sends the state to $\mathcal{C}$;
- $\mathcal{C}$ runs qrEval(QR-PUF$_i$, $|\psi_j\rangle$) $= |\omega_{ij}\rangle$ and sends the pair $|\omega_{ij}\rangle$ to $\mathcal{A}$.
- $\mathcal{A}$ uses bases B_i to obtain the classical description $\boldsymbol{\omega}_{ij}$ of $|\omega_{ij}\rangle$ and appends $\boldsymbol{\omega}_{ij}$ to S_{out}.

Challenge – Let c show the type of the challenge phase.

- If $c =$ qEx: $\mathcal{A}$ prepares a quantum state $|\psi^*\rangle \in \mathcal{H}^D$, such that $|\psi^*\rangle \notin_\mu S_{in}$. $\mathcal{A}$ keeps the classical description of $|\psi^*\rangle$ and sends the state to $\mathcal{C}$;
- If $c =$ qSel: $\mathcal{C}$ prepares a quantum state $|\psi^*\rangle \in \mathcal{H}^D$ uniformly at random. $\mathcal{C}$ keeps a copy of $|\psi^*\rangle$ and sends an extra copy to $\mathcal{A}$.

Guess

- $\mathcal{A}$ sends their guess $|\hat{\omega}\rangle$ on the output of qrEval(QR-PUF$_i$, $|\psi^*\rangle$) to $\mathcal{C}$;
- $\mathcal{C}$ computes $|\omega^*\rangle =$ qrEval(QR-PUF$_i$, $|\psi^*\rangle$);
- $\mathcal{C}$ performs a measurement $\mathcal{M}$ of the operator $|\omega^*\rangle\langle\omega^*|$ on $|\hat{\omega}\rangle$ to obtain an output $b \in [0,1]$. $\mathcal{C}$ outputs b. $\mathcal{A}$ wins the game iff $b = 1$.

qPUF Unforgeability Game

Let $\mathcal{T}$ and $\text{qPUF} = (\text{qGen}, \text{qEval}, \mathcal{T})$ be defined as in Definitions 3 and 16, respectively.

$$\mathcal{G}^{\text{qPUF}}_{c,\mu}(\mathcal{A}, \lambda)$$

Setup

- The challenger $\mathcal{C}$ selects λ and runs $\text{qGen}(\lambda)$ to build an instance of the qPUF family, qPUF_i. Then, $\mathcal{C}$ reveals λ, the domain and range Hilbert space $\mathcal{H}^D$ of the qPUF_i, and the identifier of qPUF_i, i to the adversary, $\mathcal{A}$. The challenger creates two databases, S_{in} and S_{out} which are initially empty and shares them with the adversary $\mathcal{A}$.

Learning – For $j = 1 : k$, where $k \in O(\text{poly}(\lambda))$:

- $\mathcal{A}$ prepares a quantum state $|\psi_j\rangle \in \mathcal{H}^D$, appends the classical description of $|\psi_j\rangle$ to S_{in} and sends the state to $\mathcal{C}$;
- $\mathcal{C}$ runs $\text{qEval}(\text{qPUF}_i, |\psi_j\rangle)$ and sends $|\omega_{ij}\rangle$ to $\mathcal{A}$;
- $\mathcal{A}$ appends $|\omega_{ij}\rangle$ to S_{out}.

Challenge – Let c show the type of the challenge phase.

- If $c = \mathsf{qEx}$: $\mathcal{A}$ prepares a quantum state $|\psi^*\rangle \in \mathcal{H}^D$, such that $|\psi^*\rangle \notin_\mu S_{in}$. $\mathcal{A}$ keeps the classical description of $|\psi^*\rangle$ and sends the state to $\mathcal{C}$;
- If $c = \mathsf{qSel}$: $\mathcal{C}$ prepares a quantum state $|\psi^*\rangle \in \mathcal{H}^D$ uniformly at random. The challenger keeps κ_1 copies of $|\psi^*\rangle$ and sends κ_2 copies of $|\psi^*\rangle$ to $\mathcal{A}$.

Guess

- $\mathcal{A}$ sends κ_2 copies of their guess $|\hat{\omega}\rangle$ on the output of $\text{qEval}(\text{qPUF}_i, |\psi^*\rangle)$ to $\mathcal{C}$;
- $\mathcal{C}$ runs $\text{qEval}(\text{qPUF}_i, |\psi^*\rangle^{\otimes\kappa_1})$, and gets $|\omega^*\rangle^{\otimes\kappa_1}$;
- $\mathcal{C}$ runs the test algorithm $b \leftarrow \mathcal{T}(|\omega^*\rangle^{\otimes\kappa_1}, |\hat{\omega}\rangle^{\otimes\kappa_2})$ where $b \in \{0, 1\}$ and outputs b. $\mathcal{A}$ wins the game iff $b = 1$.

3.2 Quantum Unconditional, Existential & Selective Unforgeability

Based on the above games we define the security notions quantum unconditional, existential and selective unforgeability for PUFs. The first implies the unforgeability of PUFs against an unbounded adversary with unlimited access to the PUF in the learning phase; the second is the most common and strongest type of the unforgeability against Quantum Polynomial-Time (QPT) adversaries; finally, the third is a weaker notion of unforgeability, but one we will prove to be sufficient for PUF-based ID protocols.

Now, for $\mathrm{PUF} \in \{\mathrm{cPUF}, \mathrm{QR\text{-}PUF}, \mathrm{qPUF}\}$:

Definition 18 (Quantum Unconditional Unforgeability)**.** *A* PUF *provides quantum unconditional unforgeability if the success probability of any* unbounded *adversary $\mathcal{A}$ in winning the game $\mathcal{G}^{\mathrm{PUF}}_{\mathsf{qEx},\mu}(\mathcal{A},\lambda)$ is negligible in λ:*

$$\Pr[1 \leftarrow \mathcal{G}^{\mathrm{PUF}}_{\mathsf{qEx},\mu}(\mathcal{A},\lambda)] = negl(\lambda)$$

Definition 19 (μ-Quantum Existential Unforgeability)**.** *A* PUF *provides μ-quantum existential unforgeability if the success probability of any QPT adversary $\mathcal{A}$ in winning the game $\mathcal{G}^{\mathrm{PUF}}_{\mathsf{qEx},\mu}(\mathcal{A},\lambda)$ is negligible in λ:*

$$\Pr[1 \leftarrow \mathcal{G}^{\mathrm{PUF}}_{\mathsf{qEx},\mu}(\mathcal{A},\lambda)] = negl(\lambda)$$

Note that for the case of cPUF in the above two definitions, $\mu = \mathsf{null}$.

Definition 20 (Quantum Selective Unforgeability)**.** *A* PUF *provides quantum selective unforgeability if the success probability of any QPT $\mathcal{A}$ in winning the game $\mathcal{G}^{\mathrm{PUF}}_{\mathsf{qSel}}(\mathcal{A},\lambda)$ is negligible in λ:*

$$\Pr[1 \leftarrow \mathcal{G}^{\mathrm{PUF}}_{\mathsf{qSel}}(\mathcal{A},\lambda)] = negl(\lambda)$$

3.2.1 PUF Unforgeability Results

The following are unforgeability results derived from the literature, concluding that quantum unconditional and existential unforgeability are too strong for any of our defined PUF types to provide whilst quantum selective unforgeability is provided under certain conditions.

Quantum Unconditional and Existential Unforgeability

We first present an impossibility result concerning existential unforgeability of unitary PUFs, borrowed from [4]:

Theorem 1 (No Unitary PUF provides Quantum Existential Unforgeability)**.** *No* PUF *with quantum evaluation algorithm modelled by a unitary* $\mathrm{U}_{\mathrm{PUF}}$ *of dimension* $D = 2^{\lambda}$ *can satisfy* μ*-quantum existential unforgeability for any* $0 < \mu \leq 1 - \textit{non-negl}(\lambda)$*.*

Given that we have confined our scope of investigation to PUFs modelled by unitary transformations, Theorem 1 applies to all PUF types in this work. Note the requirement for $0 < \mu \leq 1 - \textit{non-negl}(\lambda)$: we argue this to be sufficient to conclude quantum existential unforgeability is too strong a property for our PUFs, leaving the case $\mu \geq 1 - \textit{negl}(\lambda)$ for future work. We can thus also conclude that quantum unconditional unforgeability is too strong a property, since it concerns an *unbounded adversary* with existential challenge phase rather than a QPT adversary. It has also been shown in [3] that quantum unconditional unforgeability is too strong even for a selective challenge phase.

Theorem 2 (No Unitary PUF provides Quantum Unconditional Unforgeability)**.** *No* PUF *with quantum evaluation algorithm modelled by a unitary* $\mathrm{U}_{\mathrm{PUF}}$ *of dimension* $D = 2^{\lambda}$ *can satisfy quantum unconditional unforgeability for either existential or selective challenge phase.*

Quantum Selective Unforgeability

Before presenting any results regarding quantum selective unforgeability, we must introduce the notion of an *unknown unitary* – the assumption that $\mathrm{U}_{\mathrm{PUF}}$ is unknown to a prover prior to protocol execution. The formal definition of an unknown unitary is borrowed from [4]:

Definition 21 (Unknown Unitary Transformation)**.** *A set of unitary transformations* $\mathcal{U}$*, over* $\mathcal{H}^D$*, is termed a set of Unknown Unitaries if, for all QPT adversaries* $\mathcal{A}$*, and* $\mathrm{U} \in \mathcal{U}$ *sampled uniformly at random, the following holds:*

$$\Pr\left[\forall\,|\psi\rangle \in \mathcal{H}^D : F(\mathcal{A}(|\psi\rangle), \mathrm{U}\,|\psi\rangle) \geq \textit{non-negl}(\log(D))\right] = \textit{negl}(\log(D)).$$

Theorem 3 (Quantum Selective Unforgeability of Unknown Unitary PUFs)**.** *Any PUF of type* $\{\text{cPUF}, \text{QR-PUF}, \text{qPUF}\}$ *with quantum evaluation algorithm modelled by an* *unknown unitary* $\mathrm{U}_{\mathrm{PUF}}$*, according to Definition 21, satisfies quantum selective unforgeability so long as the challenge is chosen uniformly at random from the respective PUF Hilbert space* $\mathcal{H}^D$*.*

Proof. **Omitted.** For proof see Appendix E.5 of [4]; this Theorem is a direct corollary. □

An important point to note from Theorem 3 is that quantum selective unforgeability requires the PUF unitary to be unknown. If an adversary were to obtain a complete description of the unitary, say through performing the PUF characterisation method outlined in [24], we know from [4] that they would be able to perform a *quantum emulation attack* (introduced in [4], based upon the quantum emulation algorithm from [15]) to emulate the action of the PUF and break unforgeability. We establish the following results:

Corollary 1. *Any* QR-PUF *as in Definition 13 provides quantum selective unforgeability if its unitary transformation is unknown and the challenge is chosen uniformly at random from the* QR-PUF *Hilbert space* $\mathcal{H}^D$*.*

Corollary 2. *Any unitary* qPUF *as in Definition 17 provides quantum selective unforgeability if its unitary transformation is unknown and the challenge is chosen uniformly at random from the* qPUF *Hilbert space* $\mathcal{H}^D$*.*

Unfortunately, it's not immediately obvious from Theorem 3 whether our defined cPUF provides quantum selective unforgeability. The cPUF unforgeability game instructs that the challenge be chosen from $\{0,1\}^\lambda$, which clearly does not represent a uniformly random choice from $\mathcal{H}^D$. Intuitively, $\mathcal{H}^D$ contains an uncountably large number of possible quantum challenge states whilst $\{0,1\}^\lambda$ contains a finite 2^λ classical bit strings – so implication of the property is not clear. In light of this, we will take an alternative approach to proving quantum selective unforgeability of cPUFs.

3.3 Alternative Route to cPUF Selective Unforgeability

Our proof of cPUF quantum selective unforgeability is not as direct as the case of QR-PUF and qPUF. Instead, we introduce the concept of *blinded unforgeability*, repurposing the definition from [1] for the case of cPUF, showing that it implies quantum selective unforgeability. We then show that blinded unforgeability is fulfilled by cPUFs, if they are assumed to be *quantum-secure pseudorandom permutations*.

3.3.1 Blinded Unforgeability

Let $\mathcal{A}$ be a QPT adversary and $\varepsilon : \mathbb{N} \to [0,1]$ an efficiently computable function. We suppose the existence of an algorithm χ_{B_ε} (defined below) that acts to ensure $\mathcal{A}$ does not receive response information about queries in the blinded subset $B_\varepsilon \subseteq \{0,1\}^\lambda$.

$\mathrm{BlindForge}_{\mathrm{cPUF}}(\mathcal{A})$

Setup

- The challenger $\mathcal{C}$ selects the security parameter λ and runs $\mathrm{cGen}(\lambda)$ to build an instance of the cPUF family, cPUF_i. $\mathcal{C}$ reveals λ, the dimension D of the Hilbert space $\mathcal{H}^D$, and the identifier of cPUF_i, i, to the adversary $\mathcal{A}$. $\mathcal{A}$ selects the parameter ε. $\mathcal{C}$ generates the blinding set $B_\varepsilon \subseteq \{0,1\}^\lambda$ by placing each $\mathbf{x} \in \{0,1\}^\lambda$ into B_ε independently with probability $\varepsilon(\lambda)$.

Learning – For $j = 1 : k$, where $k \in O(\mathrm{poly}(\lambda))$:

- $\mathcal{A}$ prepares a superposition of challenge states $|\psi_j\rangle = \Sigma_n \alpha_n |c_n\rangle \in \mathcal{H}^D$, and sends $|\psi_j\rangle$ to $\mathcal{C}$.
- $\mathcal{C}$ runs an algorithm χ_{B_ε} that acts on $|\psi_j\rangle$ as follows:

$$\chi_{B_\varepsilon}(|\psi_j\rangle) = |\omega_{ij}\rangle = \Sigma_n \alpha_n \begin{cases} |r_{in}\rangle = \mathrm{cEval}_{\mathrm{II}}(\mathrm{cPUF}_i, |c_n\rangle), & \text{if } \mathbf{c}_n \notin B_\varepsilon \\ `\perp', & \text{if } \mathbf{c}_n \in B_\varepsilon \end{cases}$$

and sends $|\omega_{ij}\rangle$ to $\mathcal{A}$.

Forgery

- $\mathcal{A}$ selects $\mathbf{c}^* \in \{0,1\}^\lambda$ and sends $(\mathbf{c}^*, \hat{\mathbf{r}})$ to $\mathcal{C}$, where $\hat{\mathbf{r}}$ is $\mathcal{A}$'s guess on the output of $\mathrm{cEval}_{\mathrm{I}}(\mathrm{cPUF}_i, \mathbf{c}^*)$.

Outcome

- $\mathcal{C}$ runs $\mathrm{cEval}_{\mathrm{I}}(\mathrm{cPUF}_i, \mathbf{c}^*)$ to obtain $\mathbf{r}^*$.
- If $\hat{\mathbf{r}} \in B_\varepsilon$ and $\hat{\mathbf{r}} = \mathbf{r}^*$, $\mathcal{C}$ outputs 1 ($\mathcal{A}$ wins); else, $\mathcal{C}$ outputs 0 ($\mathcal{A}$ loses).

Definition 22 (Blinded Unforgeability)**.** *A* cPUF *provides blinded unforgeability if the success probability of any Quantum Polynomial-Time (QPT) adversary $\mathcal{A}$ in winning the game* $\mathrm{BlindForge}_{\mathrm{cPUF}}(\mathcal{A})$ *is negligible in* λ

$$\Pr[1 \leftarrow \mathrm{BlindForge}_{\mathrm{cPUF}}(\mathcal{A})] = negl(\lambda)$$

3.3.2 Selective Unforgeability from Blinded Unforgeability

Theorem 4 (Quantum Selective Unforgeability from Blinded Unforgeability)**.** *Let* Π *be a cryptographic primitive that is blinded unforgeable for any QPT adversary* $\mathcal{A}$*. Then* Π *is quantum selectively unforgeable.*

Proof. Let Π be a cryptographic primitive that is blinded unforgeable [1] for any QPT adversary $\mathcal{A}$. Suppose we require that the challenge be chosen by the challenger rather than $\mathcal{A}$. This describes a strictly weaker notion of security, that we will call *selective blinded unforgeability*, and is thus implied by blinded unforgeability. Suppose now we reduce the probability $\varepsilon(\lambda)$ that any one $\mathbf{x} \in \{0,1\}^\lambda$ will be included in B_ε, such that B_ε contains 1 element, say $\mathbf{c}$. Then we have two cases, either:

- $\mathcal{A}$ queries $\mathbf{c}$ during the learning phase, and receives '$\perp$'. Then $\mathcal{A}$ has learned what the challenge will be, and can be considered an adaptive adversary. This occurs with probability $negl(\lambda)$.

- $\mathcal{A}$ does not query $\mathbf{c}$ during the learning phase, then the scenario is indistinguishable from one in which there is no blinding. This occurs with probability $1 - negl(\lambda)$.

Therefore, for a blinding set of 1, the scenario is indistinguishable to any QPT adversary $\mathcal{A}$ from one in which there is no blinding set and the challenge is chosen by the challenger – thus, security in the former implies security in the latter and Π is quantum selectively unforgeable.

□

3.3.3 Unforgeability of cPUF

We recall a result from [1] implying that any quantum-secure pseudorandom function (qPRF) $F : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}^t$ (that is, F such that no efficient quantum algorithm $\mathcal{D}$ can reliably distinguish between F and a truly random function) is blinded unforgeable for $m, t \in \text{poly}(n)$. We use this to show that if a cPUF is a qPRF, then it is quantum selectively unforgeable.

Informally, a function is pseudorandom (PRF) if it behaves indistinguishably from a truly random function, and a PRF is a pseudorandom permutation (PRP) if it forms a bijection mapping between an equal domain and range. We define a PRP as follows [19]:

Definition 23 (Pseudorandom Permutation)**.** *Let $F : \{0,1\}^* \times \{0,1\}^\lambda \to \{0,1\}^\lambda$ be a deterministic and efficiently computable function $\forall k \in \{0,1\}^*$, and let R be a truly random permutation function on $\{0,1\}^\lambda$. F is a pseudorandom permutation if no efficient algorithm $\mathcal{D}$ can distinguish, with non-negligibly favourable probability, between F and R for any key $k \in \{0,1\}^*$. We say that F is* quantum secure *(qPRP) if the statement holds for $\mathcal{D}$ an efficient quantum algorithm.*

We now argue that cPUFs are indistinguishable from PRPs.

Theorem 5 (Indistinguishability of cPUFs from PRPs)**.** *Let a* cPUF *be as in Definition 10 for security parameter λ. Then no efficient distinguisher $\mathcal{D}$ can distinguish, with non-negligibly favourable probability, between the cPUF and a PRP F s.t. $F : \{0,1\}^* \times \{0,1\}^\lambda \to \{0,1\}^\lambda$ for fixed, unknown key $k \in \{0,1\}^*$.*

Proof. We have defined cPUFs in such a way that their action can be modelled by a unitary transformation $\mathrm{U}_{\mathrm{cPUF}}$, which by definition has an inverse $\mathrm{U}^{-1}_{\mathrm{cPUF}}$ and defines a permutation on the set of classical challenges $\{0,1\}^\lambda$. We can consider cPUFs to be pseudorandom permutations by modelling their unique manufacturing conditions, which determine their unique challenge-response behaviour, by a key $k \in \{0,1\}^*$. Thus, generation of a cPUF can be deemed indistinguishable from the process of selecting a k uniformly at random to produce a pseudorandom permutation for which k is entirely unknown to any party and infeasible to recover.

□

We argue, therefore, that cPUFs exhibit the same properties as PRPs with regards to unforgeability – and we claim the following:

Corollary 3. *Any cPUF as in Definition 10 provides quantum selective unforgeability so long as it is quantum secure – that is, if no efficient quantum algorithm $\mathcal{D}$ can distinguish between cPUF and a truly random permutation with non-negligibly favourable probability.*

Proof. According to Theorem 5 cPUFs are indistinguishable from PRPs which, assuming cPUFs are quantum secure, implies they are indistinguishable from qPRPs and therefore qPRFs. Using the result from [1] we conclude that such cPUFs provide blinded unforgeability, and therefore are quantum selectively unforgeable. □

Chapter 4

PUF-based Identification Protocols

Recall that an identification (ID) protocol is a method of authenticating a given device or individual for access to a privileged system or action. In this chapter we define PUF-based ID protocols, or such means of authenticating a given device or individual through use of a PUF. A protocol is referred to as PUF-based if at least one of its components can construct PUF CRPs to be used in the rest of the protocol.

Our aim is to compare a cPUF-based ID protocol with their quantum-enhanced PUF-based counterparts, and determine whether the quantum element adds an advantage. We focus on drawing comparisons across three protocol varieties: cPUF-based, as in [17] and [9]; QR-PUF-based, adapted from [24]; qPUF-based, inspired by [3].

General Protocol

We call the entity requesting authentication the prover, $\mathcal{P}$, and the trusted authority to which it must prove its legitimate status the verifier, $\mathcal{V}$. Our protocols are constructed with three phases: *Enrollment*, *Identification* and *Verification*. Enrollment sees $\mathcal{V}$ query the PUF in order to store a number of CRPs; the Identification and Verification phases are summarised as follows:

- $\mathcal{V}$ selects a challenge x from the table of stored CRPs and queries $\mathcal{P}$;
- $\mathcal{P}$, if they are an honest party with legitimate access to the PUF, queries the PUF with x and receives the output y_x, returning y_x to $\mathcal{V}$;
- If the y_x returned by $\mathcal{P}$ matches the corresponding y according to $\mathcal{V}$'s stored CRP (which it will, provided that $\mathcal{P}$ has access to the PUF) then $\mathcal{P}$ is authenticated.

4.1 cPUF ID Protocol

As this is the classical case, we begin by assuming the existence of a classical PUF according to Definition 10, with security parameter λ and unique identifier i. For this protocol, CRPs are of an entirely classical nature.

Enrollment

- Possession or control of cPUF_i is given to $\mathcal{V}$, such that $\mathcal{V}$ is free to interact with and query cPUF_i.
- $\mathcal{V}$ queries cPUF_i n times for n unique challenges $\mathbf{c}_j \in \{0,1\}^\lambda$ chosen uniformly at random, where $n \in O(\text{poly}(\lambda))$, obtaining

  $$\mathbf{r}_{ij} = \text{cEval}_\text{I}(\text{cPUF}_i, \mathbf{c}_j) \quad \text{for } j = 1...n,$$

  and stores the $\text{CRP}_{ij} = \{\mathbf{c}_j, \mathbf{r}_{ij}\}$ in a secure, private database CRT_i.
- cPUF_i is returned to its legitimate owner over a public channel.

Identification

- $\mathcal{V}$ selects a CRP_{ij} at random from CRT_i, and queries $\mathcal{P}$ with $\mathbf{c}_j$.
- $\mathcal{P}$ obtains a response $\mathbf{r}$ by querying cPUF_i with c_j, if they have access to cPUF_i as claimed, and returns $\mathbf{r}$, or nothing (‘$\perp$’), to $\mathcal{V}$.
- If ‘$\perp$’, the protocol aborts; otherwise, $\mathcal{V}$ proceeds to verification phase.

Verification

- $\mathcal{V}$ compares the stored response $\mathbf{r}_{ij}$ from CRP_{ij} with $\mathbf{r}$.
- If $\mathbf{r} = \mathbf{r}_{ij}$, $\mathcal{P}$ is authenticated as having access to cPUF_i, else $\mathcal{P}$ is denied authentication.
- $\mathcal{V}$ deletes CRP_{ij} from CRT_i so it is not reused in another protocol instance.

4.2 QR-PUF ID Protocol

This protocol is based on a QR-PUF as defined in Section 2.2, with security parameter λ, for which the classical descriptions of response states are public knowledge.

Enrollment

- Possession or control of QR-PUF$_i$ is given to $\mathcal{V}$, such that $\mathcal{V}$ is free to interact with and query QR-PUF$_i$.
- $\mathcal{V}$ queries QR-PUF$_i$ n times for n unique challenges $|\psi_j\rangle \in \mathcal{H}^D$ chosen uniformly at random, where $n \in O(\text{poly}(\lambda))$, obtaining

$$|\omega_{ij}\rangle = \text{qrEval}(\text{QR-PUF}_i, |\psi_j\rangle$$

  and stores the CRP$_{ij}$ = $\{\boldsymbol{\psi}_j, \boldsymbol{\omega}_{ij}\}$ in a secure, private database CRT$_i$.
- QR-PUF$_i$ is returned to its legitimate owner over a public channel.

Identification

- $\mathcal{V}$ selects a CRP$_{ij}$ at random from CRT$_i$, and queries $\mathcal{P}$ with $|\psi_j\rangle$.
- $\mathcal{P}$ obtains a response $|\omega\rangle$ by querying QR-PUF$_i$ with $|\psi_j\rangle$, if they have access to QR-PUF$_i$ as claimed, and returns $|\omega\rangle$, or nothing ('$\perp$'), to $\mathcal{V}$
- If '$\perp$', the protocol aborts; otherwise $\mathcal{V}$ proceeds to verification phase.

Verification

- $\mathcal{V}$ uses the stored response $\boldsymbol{\omega}_{ij}$ from CRP$_{ij}$ to prepare the state $|\omega_{ij}\rangle$.
- $\mathcal{V}$ performs a measurement $\mathcal{M}$ of the operator $|\omega_{ij}\rangle\langle\omega_{ij}|$ on $|\omega\rangle$ and obtains an output of 1 if $|\omega\rangle = |\omega_{ij}\rangle$ and 0 otherwise.
- If $|\omega\rangle = |\omega_{ij}\rangle$, $\mathcal{P}$ is authenticated as having access to QR-PUF$_i$, else $\mathcal{P}$ is denied authentication.
- $\mathcal{V}$ deletes CRP$_{ij}$ from CRT$_i$ so it is not reused in another protocol instance.

4.3 qPUF ID Protocol

This protocol is based on a unitary qPUF as defined in Section 2.3, with security parameter λ. $\mathcal{V}$ stores classical encodings of challenges but quantum responses, and no state's classical description is public. The quantities $\kappa_{1,2}$ are variable to suit the security needs of the system – tuning the accuracy of quantum testing algorithm $\mathcal{T}$.

Enrollment

- Possession or control of qPUF_i is given to $\mathcal{V}$, such that $\mathcal{V}$ is free to interact with and query qPUF_i.
- $\mathcal{V}$ queries qPUF_i κ_1 times for each of n unique challenges $|\psi_j\rangle \in \mathcal{H}^D$ chosen uniformly at random, where $n \in O(\text{poly}(\lambda))$, obtaining κ_1 copies of

  $$|\omega_{ij}\rangle = \text{qEval}(\text{qPUF}_i, |\psi_j\rangle) \quad \text{for } j = 1...n,$$

  and stores the $\text{CRP}_{ij} = \{\boldsymbol{\psi}_j, |\omega_{ij}\rangle^{\otimes\kappa_1}\}$ in a secure, private database CRT_i.
- qPUF_i is returned to its legitimate owner over a public channel.

Identification

- $\mathcal{V}$ selects a CRP_{ij} at random from CRT_i, and queries $\mathcal{P}$ κ_2 times with $|\psi_j\rangle$.
- $\mathcal{P}$ obtains κ_2 copies of a response $|\omega\rangle$ by querying qPUF_i with $|\psi_j\rangle$ κ_2 times, if they have access to qPUF_i as claimed, and returns κ_2 copies of $|\omega\rangle$, or nothing ('$\perp$'), to $\mathcal{V}$.
- If $\mathcal{V}$ receives '$\perp$', the protocol aborts; otherwise, $\mathcal{V}$ proceeds to verification.

Verification

- $\mathcal{V}$ compares $|\omega_{ij}\rangle$ from CRP_{ij} with $|\omega\rangle$, by performing a quantum test for equality on the κ_1 stored copies of $|\omega_{ij}\rangle$ and κ_2 copies of $|\omega\rangle$:

  $$\alpha \leftarrow \mathcal{T}(|\omega_{ij}\rangle^{\otimes\kappa_1}, |\omega\rangle^{\otimes\kappa_2}) \qquad \alpha \in \{0,1\}.$$

- If $\alpha = 1$, $\mathcal{P}$ is authenticated as having access to qPUF_i; else $\mathcal{P}$ is denied.
- $\mathcal{V}$ deletes CRP_{ij} from CRT_i so it is not reused in another protocol instance.

Chapter 5

Analysis of ID Protocols

5.1 Security Analysis of Chapter 4 Protocols

The principal attack any ID protocol must protect against is impersonation – an adversary *pretending* to have legitimate access to the privileged system when they do not, so as to be falsely authenticated. We characterise a *general quantum adversary* as a realistic computationally bound quantum adversary that seeks to achieve successful impersonation, without reference to their particular attack method. More specifically, we assume $\mathcal{A}$ to be a QPT adversary with both local quantum computational abilities and quantum access to the PUF primitives.

We have constructed our concept of a general quantum adversary with reference to a variety of theoretical QR-PUF ID protocol attacks that have been explored in the literature. To support our model, descriptions of these attacks have been outlined in Appendix A for the reader's interest. To the best of our knowledge, quantum attacks against cPUF-based ID protocols have not yet been introduced.

5.1.1 Security Against a General Quantum Adversary

We model security of a PUF-based ID protocol through means of a security game between a challenger $\mathcal{C}$ and a QPT adversary $\mathcal{A}$. In the following, λ is the PUF security parameter, $D = 2^\lambda$ is the dimension of the PUF Hilbert space $\mathcal{H}^D$, and $|x\rangle \in \mathcal{H}^D$ is the quantum encoding of a classical challenge / response $\mathbf{x} \in \{0,1\}^\lambda$.

cPUF

Let the cPUF-based ID protocol be defined as in Section 4.1. Security of the protocol against a general quantum adversary is defined upon the following security game:

$$\mathcal{G}^{\text{cPUFIDSec}}(\mathcal{A},\lambda)$$

Setup

- $\mathcal{C}$ runs cGen(λ) to create an instance of the cPUF family, cPUF$_i$.
- $\mathcal{C}$ creates a CRP table in a local classical database CRT$_i^{\mathcal{C}}$.
- For $j = 1 : n$, where $n \in O(\text{poly}(\lambda))$
    - * $\mathcal{C}$ selects $\mathbf{c}_j \in \{0,1\}^\lambda$ uniformly at random;
    - * $\mathcal{C}$ runs cEval$_\text{I}$(cPUF$_i$, $\mathbf{c}_j$) $= \mathbf{r}_{ij}$;
    - * $\mathcal{C}$ appends CRP$_{ij}^{\mathcal{C}} = \{\mathbf{c}_j, \mathbf{r}_{ij}\}$ to CRT$_i^{\mathcal{C}}$.

Learning

- $\mathcal{A}$ creates a CRP table in a local quantum database CRT$_i^{\mathcal{A}}$.
- For $j = 1 : k$, where $k \in O(\text{poly}(\lambda))$:
    - * $\mathcal{A}$ prepares a superposition of challenge states $|\psi_j\rangle = \Sigma_n \alpha_n |\psi_n\rangle \in \mathcal{H}^D$, and sends $|\psi_j\rangle$ to $\mathcal{C}$;
    - * $\mathcal{C}$ runs cEval$_\text{II}$(cPUF$_i$, $|\psi_j\rangle$) $= \Sigma_n \alpha_n |\omega_{in}\rangle = |\omega_{ij}\rangle$, and sends $|\omega_{ij}\rangle$ to $\mathcal{A}$;
    - * $\mathcal{A}$ appends CRP$_{ij}^{\mathcal{A}} = \{\boldsymbol{\psi}_j, |\omega_{ij}\rangle\}$ to CRT$_i^{\mathcal{A}}$.

Challenge

- $\mathcal{C}$ selects CRP$^* = \{\mathbf{c}^*, \mathbf{r}^*\}$ uniformly at random from CRT$_i^{\mathcal{C}}$ and sends $\mathbf{c}^*$ to $\mathcal{A}$.

Guess

- $\mathcal{A}$ returns $\hat{\mathbf{r}}$ to $\mathcal{C}$, where $\hat{\mathbf{r}}$ is $\mathcal{A}$'s guess for $\mathbf{r}^*$.
- $\mathcal{C}$ compares $\hat{\mathbf{r}}$ with $\mathbf{r}^*$ from CRP*, and outputs 1 if they are the same, or 0 otherwise.

QR-PUF

Let the QR-PUF-based ID protocol be defined as in Section 4.2. Security of the protocol against a general quantum adversary is defined upon the following security game:

$\mathcal{G}^{\text{QR-PUFIDSec}}(\mathcal{A}, \lambda)$

Setup

- $\mathcal{C}$ runs qrGen(λ) to create an instance of the QR-PUF family, QR-PUF$_i$, revealing bases B_i.
- $\mathcal{C}$ creates a CRP table in a local classical database CRT$_i^{\mathcal{C}}$.
- For $j = 1 : n$, where $n \in O(\text{poly}(\lambda))$
  - * $\mathcal{C}$ prepares a quantum state $|\psi_j\rangle \in \mathcal{H}^D$ uniformly at random;
  - * $\mathcal{C}$ runs qrEval(QR-PUF$_i$, $|\psi_j\rangle) = |\omega_{ij}\rangle$;
  - * $\mathcal{C}$ uses bases B_i to obtain $\boldsymbol{\omega}_{ij}$ and appends CRP$_{ij}^{\mathcal{C}} = \{\boldsymbol{\psi}_j, \boldsymbol{\omega}_{ij}\}$ to CRT$_i^{\mathcal{C}}$.

Learning

- $\mathcal{A}$ creates a CRP table in a local classical database CRT$_i^{\mathcal{A}}$.
- For $j = 1 : k$, where $k \in O(\text{poly}(\lambda))$:
  - * $\mathcal{A}$ prepares a quantum state $|\psi'_j\rangle \in \mathcal{H}^D$, and sends $|\psi'_j\rangle$ to $\mathcal{C}$;
  - * $\mathcal{C}$ runs qrEval(QR-PUF$_i$, $|\psi'_j\rangle) = |\omega'_{ij}\rangle$ and sends $|\omega'_{ij}\rangle$ to $\mathcal{A}$;
  - * $\mathcal{A}$ uses bases B_i to obtain $\boldsymbol{\omega}'_{ij}$, and appends CRP$_{ij}^{\mathcal{A}} = \{\boldsymbol{\psi}'_j, \boldsymbol{\omega}'_{ij}\}$ to CRT$_i^{\mathcal{A}}$.

Challenge

- $\mathcal{C}$ selects CRP$^* = \{\boldsymbol{\psi}^*, \boldsymbol{\omega}^*\}$ uniformly at random from CRT$_i^{\mathcal{C}}$ and sends $|\psi^*\rangle$ to $\mathcal{A}$.

Guess

- $\mathcal{A}$ returns $|\hat{\omega}\rangle$ to $\mathcal{C}$, where $|\hat{\omega}\rangle$ is $\mathcal{A}$'s guess for $|\omega^*\rangle$.
- $\mathcal{C}$ prepares $|\omega^*\rangle$ and performs a measurement $\mathcal{M}$ of $|\omega^*\rangle\langle\omega^*|$ on $|\hat{\omega}\rangle$, outputting 1 if the measurement yields 1 ($|\hat{\omega}\rangle = |\omega^*\rangle$), and 0 otherwise.

qPUF

Let the qPUF-based ID protocol be defined as in Section 4.3. Security of the protocol against a general quantum adversary is defined upon the following security game:

$\mathcal{G}^{\text{qPUFIDSec}}(\mathcal{A}, \lambda)$

Setup

- $\mathcal{C}$ runs qGen(λ) to create an instance of the qPUF family, qPUF$_i$.
- $\mathcal{C}$ creates a CRP table in a local quantum database CRT$_i^{\mathcal{C}}$.
- For $j = 1 : n$, where $n \in \mathcal{O}(\text{poly}(\lambda))$
  - * $\mathcal{C}$ prepares κ_1 copies of a quantum state $|\psi_j\rangle \in \mathcal{H}^D$ uniformly at random;
  - * $\mathcal{C}$ runs qEval(qPUF$_i$, $|\psi_j\rangle^{\otimes\kappa_1}$) $= |\omega_{ij}\rangle^{\otimes\kappa_1}$;
  - * $\mathcal{C}$ appends CRP$_{ij}^{\mathcal{C}} = \{\boldsymbol{\psi}_j, |\omega_{ij}\rangle^{\otimes\kappa_1}\}$ to CRT$_i^{\mathcal{C}}$.

Learning

- $\mathcal{A}$ creates a CRP table in a local quantum database CRT$_i^{\mathcal{A}}$.
- For $j = 1 : k$, where $k \in \mathcal{O}(\text{poly}(\lambda))$
  - * $\mathcal{A}$ prepares a quantum state $|\psi'_j\rangle \in \mathcal{H}^D$, and sends $|\psi'_j\rangle$ to $\mathcal{C}$;
  - * $\mathcal{C}$ runs qEval(qPUF$_i$, $|\psi'_j\rangle$) $= |\omega'_{ij}\rangle$, and sends $|\omega'_{ij}\rangle$ to $\mathcal{A}$;
  - * $\mathcal{A}$ appends CRP$_{ij}^{\mathcal{A}} = \{\boldsymbol{\psi}'_j, |\omega'_{ij}\rangle\}$ to CRT$_i^{\mathcal{A}}$.

Challenge

- $\mathcal{C}$ selects CRP$^* = \{\boldsymbol{\psi}^*, |\omega^*\rangle^{\otimes\kappa_1}\}$ uniformly at random from CRT$_i^{\mathcal{C}}$ and sends κ_2 copies of $|\psi^*\rangle$ to $\mathcal{C}$.

Guess

- $\mathcal{A}$ returns κ_2 copies of $|\hat{\omega}\rangle$ to $\mathcal{C}$, where $|\hat{\omega}\rangle$ is $\mathcal{A}$'s guess for $|\omega^*\rangle$.
- $\mathcal{C}$ compares $|\omega^*\rangle$ with $|\hat{\omega}\rangle$, by performing $\mathcal{T}(|\omega^*\rangle^{\otimes\kappa_1}, |\hat{\omega}\rangle^{\otimes\kappa_2})$, outputting 1 if the test yields 1 (i.e. $|\omega^*\rangle = |\hat{\omega}\rangle$) and 0 otherwise.

Definition 24 (Security of PUF ID Protocol)**.** *For* PUF $\in \{$cPUF, QR-PUF, qPUF$\}$*, we say a* PUF*-based ID protocol is secure against a general quantum adversary if the success probability of any QPT adversary $\mathcal{A}$ in the game $\mathcal{G}^{\text{PUFIDSec}}(\mathcal{A},\lambda)$ is negligible in the security parameter λ:*

$$\Pr\big[1 \leftarrow \mathcal{G}^{\text{PUFIDSec}}(\mathcal{A},\lambda)\big] = negl(\lambda).$$

5.1.2 Reduction of Security to Selective Unforgeability

In this section we reduce the definition of PUF ID protocol security to selective unforgeability of the underlying PUF, proving that if a PUF is quantum selectively unforgeable, then the respective ID protocol is secure against a general quantum adversary. In fact we prove the contrapositive, showing that if there is an adversary $\mathcal{A}$ who can win $\mathcal{G}^{\text{PUFIDSec}}(\mathcal{A},\lambda)$ with non-negligible probability, thereby breaking the security of a PUF ID protocol, then an adversary $\mathcal{B}$ can use $\mathcal{A}$'s algorithm to break selective unforgeability of the PUF. First, recall Definition 20 of quantum selective unforgeability for a general PUF of type {cPUF, QR-PUF, qPUF}:

Definition (Quantum Selective Unforgeability of PUF)**.** *A* PUF *is said to provide quantum selective unforgeability if the success probability of any Quantum Polynomial-Time (QPT) $\mathcal{A}$ in winning the game $\mathcal{G}^{\text{PUF}}_{\text{qSel}}(\mathcal{A},\lambda)$ is negligible in λ:*

$$\Pr[1 \leftarrow \mathcal{G}^{\text{PUF}}_{\text{qSel}}(\mathcal{A},\lambda)] = negl(\lambda).$$

Our statement is as follows:

Theorem 6 (ID Protocol Security from PUF Unforgeability)**.** *A* PUF*-based ID protocol is* secure against a general quantum adversary *if the underlying* PUF *of type* {cPUF, QR-PUF, qPUF} *is* quantum selectively unforgeable*:*

$$\Pr[1 \leftarrow \mathcal{G}^{\text{PUF}}_{\text{qSel}}(\mathcal{A},\lambda)] = negl(\lambda) \implies \Pr\big[1 \leftarrow \mathcal{G}^{\text{PUFIDSec}}(\mathcal{A},\lambda)\big] = negl(\lambda).$$

Proof. We proceed by proving the contrapositive of the above statement, i.e.

$$\Pr\big[1 \leftarrow \mathcal{G}^{\text{PUFIDSec}}(\mathcal{A},\lambda)\big] = \textit{non-negl}(\lambda) \implies \Pr\Big[1 \leftarrow \mathcal{G}^{\text{PUF}}_{\text{qSel}}(\mathcal{A},\lambda)\Big] = \textit{non-negl}(\lambda).$$

Let $\mathcal{A}$ be an adversarial algorithm that breaks security of a PUF-based ID protocol, i.e. $\Pr\big[1 \leftarrow \mathcal{G}^{\text{PUFIDSec}}(\mathcal{A},\lambda)\big] = \textit{non-negl}(\lambda)$. The game $\mathcal{G}^{\text{PUFIDSec}}(\mathcal{A},\lambda)$ is precisely the scenario captured in $\mathcal{G}_c^{\mathrm{PUF}}(\mathcal{A},\lambda)$ for a qSel challenge phase; both require that $\mathcal{A}$ respond to a $\mathcal{C}$-chosen challenge, which has been selected uniformly at random from the PUF challenge space. Whilst $\mathcal{G}^{\mathrm{PUFIDSec}}(\mathcal{A},\lambda)$ instructs the intermediary step of creating a challenger CRT, the scenarios are indistinguishable from an adversary's perspective – the resulting challenge can still be considered to be chosen uniformly at random from the PUF challenge space, since the CRT was constructed from challenges chosen this way. Therefore, any adversary $\mathcal{B}$ performing $\mathcal{A}$ as a subroutine can win the game $\mathcal{G}_{\mathsf{qSel}}^{\mathrm{PUF}}(\mathcal{B},\lambda)$ with non-negligible probability, $\Pr\left[1 \leftarrow \mathcal{G}_{\mathsf{qSel}}^{\mathrm{PUF}}(\mathcal{A},\lambda)\right] = \mathit{non\text{-}negl}(\lambda)$.

□

5.1.3 ID Protocol Security Results

Recalling our unforgeability results from Sections 3.2.1 and 3.3.3, we may use Theorem 6 to conclude the following:

Theorem 7. *Any unitary* PUF*-based ID protocol as defined in Chapter 4 provides security against a general QPT adversary $\mathcal{A}$, so long as the PUF's unitary transformation is unknown to $\mathcal{A}$ and the challenges are chosen uniformly at random.*

Proof. Follows directly from Corollaries 1, 2 & 3 and Theorem 6, assuming that cPUFs are qPRPs. □

We can also conclude insecurity in the case that the PUF's unitary transformation is *known*, using results from [4]:

Corollary 4. *No unitary* PUF*-based ID protocol as defined in Chapter 4 provides security against a general QPT adversary $\mathcal{A}$ if the PUF's unitary transformation is known to $\mathcal{A}$.*

Proof. **Omitted.** Result derives from the conclusion in [4] that if $\mathcal{A}$ has full knowledge of the PUF unitary then they can perform a quantum emulation attack to emulate the PUF's action on an unknown challenge, thereby forging a CRP without access to the PUF and breaking selective unforgeability. Insecurity follows. □

Despite our initial hypothesis that quantum-enhanced PUFs provide advantages, we have shown all three PUF types to be quantum selectively unforgeable – proving security of Chapter 4 ID protocols against a general quantum adversary. cPUF quantum selective unforgeability requires that cPUFs be qPRPs, which may be hard to achieve, however, other PUF types have no such requirement. In any case, we present scenarios which *do* highlight quantum advantage by revising ID protocols to reuse CRPs.

5.2 Revising Protocols to Reuse CRPs

The following discusses PUF ID protocol security in the instance that $\mathcal{V}$ reuses CRPs.

5.2.1 cPUF

If the verifier $\mathcal{V}$ reuses the same CRP $= (\mathbf{c}, \mathbf{r})$ for multiple protocol rounds, security of the protocol against a general QPT adversary $\mathcal{A}$ no longer holds. We support our claim by observing that, since $\mathbf{c}, \mathbf{r}$ are classical, $\mathcal{A}$ could eavesdrop on an honest protocol to learn them both. This enables a trivial forgery by allowing $\mathcal{A}$ to initiate a protocol round and return $\mathbf{r}$ to $\mathcal{V}$ when challenged with $\mathbf{c}$. Initially $\mathcal{A}$ may not be aware that $\mathcal{V}$ reuses $(\mathbf{c}, \mathbf{r})$, but since all challenges and responses are classical this can quickly be discovered over multiple protocol rounds. Thus, by altering the cPUF ID protocol such that $\mathcal{V}$ reuses CRPs, protocol security can be broken even by a classical adversary.

5.2.2 QR-PUF

Recall that classical descriptions of all QR-PUF responses are assumed to be public knowledge. This allows $\mathcal{A}$ to deduce that $\mathcal{V}$ reuses the same CRP $= (|\psi\rangle, |\omega\rangle)$ by eavesdropping on multiple honest protocol rounds as before. $\mathcal{A}$ can now store the description of $|\omega\rangle$ and submit $|\omega\rangle$ to $\mathcal{V}$ as the response during a new protocol round without needing any information about $|\psi\rangle$. Therefore, as with cPUF, the QR-PUF ID protocol security can be broken by a general quantum adversary if CRPs are reused.

5.2.3 qPUF

In the case of qPUF, exchange of both challenge $|\psi\rangle$ and response $|\omega\rangle$ between $\mathcal{V}$ and $\mathcal{P}$ is quantum-encoded, with classical descriptions unknown to $\mathcal{A}$. The quantum nature of CRPs prohibits $\mathcal{A}$ from learning the full classical description of $(|\psi\rangle, |\omega\rangle)$ through eavesdropping. $\mathcal{A}$ would need to perform super-polynomial number of measurements on the same number of state copies in order to acquire complete knowledge of either $(|\psi\rangle, |\omega\rangle)$, which is infeasible with QPT bounding. Even attempting to leak information would alert honest parties to $\mathcal{A}$'s presence: $\mathcal{A}$'s measurements on either state would necessarily collapse them into a basis of measurement and disturb transmission between $\mathcal{V}$ and $\mathcal{P}$. Thus, $\mathcal{A}$ can not determine the reuse of CRPs in the case of qPUF.

However, let us assume protocol details to be public, making $\mathcal{A}$ aware that the qPUF ID protocol reuses CRP $= (|\psi\rangle, |\omega\rangle)$. Since the classical description of $|\omega\rangle$ is unknown to $\mathcal{A}$, and the no-cloning theorem therefore prohibits its copying, in order to submit $|\omega\rangle$ as the response in a new protocol round $\mathcal{A}$ must capture the original $|\omega\rangle$ sent by $\mathcal{P}$ and store it in quantum memory. This requirement allows us to engineer security of the qPUF ID protocol by considering minor adjustments, as outlined below.

Time-Bounding Round Separation

Suppose we impose that the time between two rounds of the ID protocol is greater than the coherence time of a state in quantum memory that is possible through current (or near future) technologies. This guarantees that even if $\mathcal{A}$ successfully captures $|\omega\rangle$ in quantum memory, the state will be insufficiently coherent for use in a future ID protocol instance. However, the coherence time bound on $\mathcal{A}$'s quantum memory necessarily affects $\mathcal{V}$ too. To account for this we may assume that there exists a method of quantum error correction sufficient to maintain $|\omega\rangle$ in quantum memory using κ_1 copies of it, for large enough κ_1. If such a method is feasible, the ID protocol can remain secure with repeated use of a single CRP since $\mathcal{V}$ has κ_1 copies of $|\omega\rangle$. However, if infeasible, $\mathcal{V}$ must resample a new CRP between each ID protocol round – which is less efficient even than our original protocol, factoring in qPUF transfer.

Adaptive Change in CRPs

Alternatively, suppose we change which qPUF CRP is used whenever the ID protocol fails – which necessarily happens whenever $\mathcal{A}$ intercepts and captures $|\omega\rangle$ (as then $\mathcal{P}$ does not successfully deliver $|\omega\rangle$ to $\mathcal{V}$). We then guarantee that $\mathcal{A}$ can never hold the correct response in quantum memory – maintaining the security of the protocol whilst minimising change in CRP to only when necessary.

5.3 Efficiency Analysis

We investigate efficiency of ID protocols assuming n rounds of Identification and Verification phases for each protocol, for which $m \leq n$ rounds fail, where $n, m \in \text{poly}(\lambda)$. Protocols are assumed noiseless, i.e. honest provers always return precisely the correct response, QR-PUF Verification phase measurements are perfectly accurate in deducing equivalence of states, as is the qPUF testing algorithm $\mathcal{T}$ for sufficient $\kappa_{1,2}$. Protocol rounds are therefore assumed to fail only when subverted by an adversary – either through eavesdropping, or by acting as the prover and returning an invalid response.

Protocol efficiency is evaluated based on dependence of the required storage, computation and communication costs on the PUF security parameter λ, with each category further divided into classical and quantum costs. To represent dependence on λ, we recall the notion of computational complexity. In the context of protocol efficiency: if, for example, the classical storage cost for a protocol grows linearly with λ by a factor of k we may say it grows as $O(k\lambda)$, which is equivalent to $O(\lambda)$.

5.3.1 Original Protocols

Our original ID protocols assume change in CRP with each round, and therefore $n \in \text{poly}(\lambda)$ CRPs for n rounds. PUF Enrollment costs are a straightforward increase from cPUF through QR-PUF to qPUF – whilst QR-PUF costs have a similar dependence to cPUF they require quantum abilities, which we assume to carry inherent greater costs. qPUF is significantly more taxing since it requires κ_1 copies of each quantum response state in storage and communication, for κ_1 unbounded. The comparison in Identification and Verification phases is just as clear-cut, with costs at their lowest for cPUF and highest for qPUF. We note that verifier-borne quantum computation costs are dependent on the cost of the specific quantum testing algorithm $\mathcal{T}$ employed.

5.3.2 Revised qPUF Protocols

The main body of the qPUF-based ID protocol (Identification and Verification phases) is left unchanged by revised sampling of qPUF CRPs. However, cost of qPUF Enrollment is significantly reduced from being potentially unbounded polynomial in λ to simply linear in λ when we time-bound protocol round separation, assuming a method for $\mathcal{V}$ to retain the quality of their stored response states. We support this claim by recalling that the number of stored CRPs n is originally polynomial in λ but is reduced to 1 in the time-bounded protocol – eliminating the polynomial dependence on λ. The revision to adaptively change CRP with protocol failure presents an intermediate efficiency with costs dependent on the number m of protocols that $\mathcal{A}$ attempts to subvert, which represents a potential range from linear in λ up to polynomial.

5.3.3 Efficiency Tables

Below are efficiency results summarised into two tables displaying protocol costs as dependent on PUF security parameter λ. Table 5.1 concerns the PUF Enrollment phase, for which the verifier bears all costs, and shows both the original ID protocols and qPUF revisions. Table 5.2 concerns protocol Identification and Verification phases, displayed combined, for each ID protocol (including qPUF revisions, since costs in these phases are not affected). Costs are broken down and explained in Appendix B.

PUF	Storage		Computation		Communication	
	Quantum	Classical	Quantum	Classical	Quantum	Classical
Original Protocol						
cPUF	–	$O(2n\lambda)$	–	–	–	$O(n\lambda)$
QR-PUF	–	$O(2n\lambda)$	$n\mathcal{D}+n\mathcal{Q}$	–	$O(n\lambda)$	–
qPUF	$O(\kappa_1 n\lambda)$	$O(n\lambda)$	$n\mathcal{Q}$	–	$O(\kappa_1 n\lambda)$	–
Time-Bounded Protocol						
qPUF	$O(\kappa_1\lambda)$	$O(\lambda)$	$\mathcal{Q}$	–	$O(\kappa_1\lambda)$	–
Adaptive CRP Protocol						
qPUF	$O(\kappa_1 m\lambda)$	$O(m\lambda)$	$m\mathcal{Q}$	–	$O(\kappa_1 m\lambda)$	–

Table 5.1: Classical / quantum storage, computation and communication costs for the Enrollment stage of Chapter 4 protocols and the qPUF reused CRP revisions. All costs are borne by the verifier. $\mathcal{Q}$, $\mathcal{D}$ are costs of preparing and decoding a state respectively.

PUF	Storage		Computation		Communication	
	Quantum	Classical	Quantum	Classical	Quantum	Classical
Verifier						
cPUF	–	–	–	$O(1)$	–	$O(\lambda)$
QR-PUF	–	–	$2\mathcal{Q}+\mathcal{M}$	–	$O(\lambda)$	–
qPUF	–	–	$\kappa_2\mathcal{Q}+\mathcal{T}$	–	$O(\kappa_2\lambda)$	–
Prover						
cPUF	–	–	–	–	–	$O(2\lambda)$
QR-PUF	–	–	–	–	$O(2\lambda)$	–
qPUF	–	–	–	–	$O(2\kappa_2\lambda)$	–

Table 5.2: Classical / quantum storage, computation and communication costs for the Identification and Verification stages of ID protocols. Verifier and prover costs are displayed separately. $\mathcal{Q}$ is cost of preparing a state, $\mathcal{M}$ is cost of projective measurement in QR-PUF verification, $\mathcal{T}(\kappa_1, \kappa_2)$ the cost of the quantum testing algorithm.

Chapter 6

Discussion & Conclusion

6.1 Protocol Assessment & Comparison

We deduced with Theorem 7 that, so long as PUF unitaries are unknown and challenges are chosen uniformly at random, the ID protocols set out in Chapter 4 provide security against a general quantum adversary. Considering efficiency of these protocols and the degree of quantum capability they require (Section 5.3) we can conclude that the cPUF ID protocol is the most efficient, and requires the least quantum ability, though it does require that the underlying cPUF be a qPRP. However, by revising protocols such that they reuse CRPs, we devised two schemes that each show definitive qPUF advantage.

The first scheme instructs that time between successive qPUF ID protocol rounds be lower-bounded by an upper-bound in coherence time for a single state in quantum memory, given current or near-future technologies. This guarantees that no adversary can hold a sufficiently coherent response state in memory to successfully authenticate, and the verifier can continue to reuse the same CRP – reducing enrollment costs from polynomial to linear in λ (assuming constant $\kappa_{1,2}$). On the other hand, it requires there to exist a method of retaining coherence of κ_1 copies of a state over the same time period. This is a confident assumption, but the pay-off in efficiency is great.

The second scheme instructs that CRPs be changed only in the event that the protocol fails, which, assuming a noiseless protocol, is equivalent to when an adversary $\mathcal{A}$ attempts to subvert it. This method does not guarantee to present as stark an improvement in efficiency as the first, however, in the *best case* it is as efficient as the time-bounded revision (which is assuming that $\mathcal{A}$ does not interfere with any protocol round); and in the *worst case* it is as efficient as the original qPUF protocol (which is assuming that $\mathcal{A}$ interferes with the protocol in every round). It therefore provides a promising alternative for if there does not exist a method of retaining κ_1 copies of a state in quantum memory for longer than the required time-bound. It is also important to consider the implications of time-bounding successive protocol rounds in the presence of advanced quantum memory capabilities, since longer wait times between protocol rounds will result in the *availability* of the system being significantly compromised. In any case we have found the original Chapter 4 qPUF ID protocol to be sub-optimal, and reused CRP revisions to be a demonstration of quantum advantage.

6.2 Critical Model Evaluation

Given the natural restraints on time and resources for an MSc project, we made a number of model assumptions in order to simplify our analyses. We argue these assumptions to be sufficient for an initial study of this kind, however, we describe the corresponding model limitations below and how they affect applicability of our results.

PUF Modelling

In modelling QR-PUFs, we captured the assumption from [24] that the classical descriptions of all QR-PUF responses are public knowledge by having bases be output at PUF generation such that any party can may decode response states. We suggest that, in fact, this may not be a plausible feature, and the QR-PUF may require alternative modelling to reflect the assumption. Furthermore, whilst Corollary 1 concerning QR-PUF quantum selective unforgeability requires challenges be chosen uniformly at random from the PUF Hilbert space, such a method of sampling the challenge is not compatible in the case that CRPs are restricted to quantumly encoded classical bit strings. In such a case, quantum selective unforgeability of the QR-PUF may need more investigation – though this does not affect our conclusion of qPUF advantage.

Importantly, we assumed that each of the PUFs in this project could be modelled by a unitary transformation. Whilst this restriction significantly streamlined our analysis by allowing us to avoid dealing with mixed states, there are potentially important results concerning unforgeability to be deduced from non-unitary PUF analyses.

Existential Unforgeability Requirements

Recall that in the existential unforgeability game the adversary can query any state they wish during the learning phase and then select the challenge to which they must respond. But, if they choose one of the learnt CRPs then they have not actually fulfilled the requirement for forgery, since the requirement is generating a *new* CRP. The assumption in [3] is that the adversary will not query the exact challenge they want to forge in the learning phase, though they can query a superposition of the challenge with some others. However, since the learnt queries are quantum, it is not specified how the challenger can check whether the adversary has not queried the challenge in the learning phase. The quantum random oracle model proposed in [6] provides a way for recording queried challenges, which may provide a solution for this.

Adaptive Adversaries

Our analyses did not consider *adaptive* adversaries, as in [4], with PUF access after being challenged (without the ability to query the challenge itself). For ID protocols we believe ours to be a realistic threat model, given the assumptions that adversaries only have PUF access during the transfer period after Enrollment, and that Identification and Verification phases do not commence until transfer of the PUF is complete. However, the opportune scenario for an adaptive adversary may well arise in considering other protocols, so it should not be ignored if our model is to take on broader applications.

Alternative Attack Objectives

Our model of a general quantum adversary is such that we investigated security against impersonation, yielding a scenario in which qPUFs gave a distinct advantage. However, by considering the possibility that an adversary does not seek authentication for itself but instead seeks to deprive access to others, we see that the quantum-enhanced protocols suffer a disadvantage: quantum challenge states are fragile and susceptible to distortion by an adversary imposing a measurement or operator on them, forcing a protocol round to fail. We suggest that future models and ID protocol designs should explore mitigation strategies for denial-of-service attacks and other attack objectives.

Protocol Variations

Our variation in ID protocols was fairly limited, focusing only on a standard protocol based on three different PUF types and a consideration of qPUF protocol adjustments. Adaptations such as the multiple-round QR-PUF protocols from [24] and the resource-conscious qPUF protocols from [11] provide examples of how the protocols can be manipulated beyond our basic set-up for differing security and efficiency requirements.

Practical Implementation

Our theoretical analyses do not address the issue of practical implementation, potentially missing out on crucial vulnerabilities that may arise through composition of PUF-based protocols with adjoining systems. For example, we assumed noiseless protocols – whilst in practice, a degree of protocol noise raises the probability of a forgery being accepted, and may mask the potential for attacks based on statistical analysis.

6.3 Contributions & Significance

We have adapted definitions from [3, 4, 11] to introduce a unified PUF framework through which we define a range of PUF types and their unforgeability properties. This allows for a more straightforward and fair comparison of PUF properties, to determine their suitability to certain applications. Our PUF modelling and unforgeability definitions are independent of later chapters concerning ID protocols, and so can be used as a basis for more general PUF-based cryptographic protocols.

We have conducted a broadened analysis of PUF-based ID protocols, abstracting from overly-specific threat models in previous works, to define protocol security against a general quantum adversary. This strengthens the applicability of our model by taking the focus away from a specific attack method – allowing one to be confident in defining secure PUF-based ID protocols given unforgeability of the underlying PUF. Our protocol analyses also provide an example of how one might build and analyse a more general cryptographic protocol.

Crucially we have concluded, as we set out to, verifiable security and efficiency advantages in adopting quantum PUFs as a basis for ID protocols. This result alone is promising for the development of secure authentication systems in the quantum era, but also holds significance for the wider topic of PUFs in cryptosystems as it gives cause to investigating quantum advantages in other PUF applications.

6.3.1 Future Directions

We stress that although we have developed proof of verifiable quantum advantages in PUFs for certain settings, there is a great deal more research required, both theoretical and experimental, to make the prospect of quantum-enhanced PUF solutions a viable widespread security development. We recognise that quantum technologies are not yet sufficient to support some of our assumptions, such as classical / quantum hybrid memory, and memory capable of maintaining coherent quantum states for sufficient time (currently approx. 10^{-3}s [16]). We outlined in Section 6.2 a number of points for improvement in our theoretical model, that we believe ought to be addressed – particularly a considered approach to the highlighted issues of PUF modelling, including an investigation into non-unitary PUFs, and protocol design to mitigate vulnerability to denial-of-service attacks.

6.4 Project Reflection & Process Review

As is often the case with theoretical theses, a lack of experimental methodology and planning discussions mean that only the 'final product' of developed theory is presented. I would like here to take the opportunity to reflect on the MSc project process.

Overall, workflow was heavily disrupted by COVID-19 safety measures, such as closure of university facilities and support hubs. Improper workspace, limited resources and inadequate computer access made certain periods especially taxing. I am thankful that this project did not require location-bound experimental work or access to specialist equipment, which may have faced suspension. My supervisor and I kept to a schedule of weekly meetings, supported by frequent email exchange, to ensure that progress was habitually reviewed and no issues were left stagnant.

We did encounter a substantial setback towards the end of the project. At first, we envisioned that proof of quantum advantage would come more readily by proving that cPUFs cannot provide quantum selective unforgeability. In fact, we believed the case of cPUF to be the more straightforward to navigate – with intuition indicating that a quantum adversary ought to be able to leverage the classical system. Significant time was spent researching literature to devise an attack, covering a broad range of topics such as universal forgery and key recovery, quantum pseudorandom unitary operators, Merkle puzzles, Simon's algorithm, and permutation analysis. It was late in the project that we recognised the need to take a different approach, conceding that cPUFs *do* provide quantum selective unforgeability and working to prove it, as we have done.

Time management was particularly challenging given the implications of national lockdown during the pandemic. Though in some sense it was more straightforward to schedule regular allocation of project work throughout the week, the profound effect on both mental and physical well-being required greater emphasis on self-care and recreation to maintain work performance. Despite this, I am very pleased to have achieved close coordination with milestones outlined in the project proposal.

Bibliography

[1] G. Alagic, C. Majenz, A. Russell, and F. Song. Quantum-secure message authentication via blind-unforgeability. *ArXiv e-prints: 1803.03761*, 03 2018.

[2] M. V. Anand, E. E. Targhi, G. N. Tabia, and D. Unruh. Post-quantum security of the CBC, CFB, OFB, CTR, and XTS modes of operation. In *Proceedings of the 7th Annual Workshop on Post-Quantum Cryptography*, pages 44–63. Association for Computing Machinery, 02 2016.

[3] M. Arapinis, M. Delavar, M. Doosti, and E. Kashefi. Quantum physical unclonable functions: Possibilities and impossibilities. *arXiv e-prints: 1910.02126*, 10 2019.

[4] M. Arapinis, M. Delavar, M. Doosti, and E. Kashefi. Unforgeability in the quantum world. *Cryptology ePrint Archive, Report 2020/291*, 03 2020.

[5] F. Armknecht, D. Moriyama, A. Sadeghi, and M. Yung. Towards a unified security model for physically unclonable functions. In *2016 Cryptographers' Track at the RSA Conference*, pages 271–287. Springer International Publishing, 02 2016.

[6] D. Boneh, Ö. Dagdalen, M. Fischlin, A. Lehmann, C. Schaffner, and M. Zhandry. Random oracles in a quantum world. In *2011 International Conference on the Theory and Application of Cryptology and Information Security*, pages 41–69. Springer Berlin Heidelberg, 12 2011.

[7] A. Braeken. PUF based authentication protocol for IoT. *Symmetry*, 10:352, 08 2018.

[8] I. Damgård, J. Funder, J. B. Nielsen, and L. Salvail. Superposition attacks on cryptographic protocols. In *2013 International Conference on Information Theoretic Security*, pages 142–161. Springer International Publishing, 11 2013.

[9] J. Delvaux, R. Peeters, D. Gu, and I. Verbauwhede. A survey on lightweight entity authentication with strong PUFs. *ACM Computing Surveys*, 48(2), 10 2015.

[10] X. Dong, Z. Li, and X. Wang. Quantum cryptanalysis on some generalized Feistel schemes. *Science China Information Sciences*, 62(22501), 02 2019.

[11] M. Doosti, N. Kumar, M. Delavar, and E. Kashefi. Client-server identification protocols with quantum PUF. *arXiv e-prints: 2006.04522*, 06 2020.

[12] G. Gianfelici, H. Kampermann, and D. Bruß. A theoretical framework for physical unclonable functions, including quantum readout. *Physical Review A*, 101:042337, 04 2020.

[13] M. Kaplan, G. Leurent, A. Leverrier, and M. Naya-Plasencia. Breaking symmetric cryptosystems using quantum period finding. In *2016 Annual International Cryptology Conference*, pages 207–237. Springer Berlin Heidelberg, 07 2016.

[14] I. L. Cheung M. A. Nielsen. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, University Press, Cambridge, 2010.

[15] I. Marvian and S. Lloyd. Universal quantum emulator. *ArXiv e-prints: 1606.02734*, 06 2016.

[16] K. C. Miao, J. P. Blanton, C. P. Anderson, A. Bourassa, A. L. Crook, G. Wolfowicz, H. Abe, T. Ohshima, and D. D. Awschalom. Universal coherence protection in a solid-state spin qubit. *Science*, 369(6505), 08 2020.

[17] R. S. Pappu. Physical one-way functions. *Science*, 297(5589), 10 2002.

[18] R. Plaga and F. Koob. A formal definition and a new security mechanism of physical unclonable functions. In *2012 International GI/ITG Conference on Measurement, Modelling, and Evaluation of Computing Systems and Dependability and Fault Tolerance*, pages 288–301. Springer Berlin Heidelberg, 03 2012.

[19] M. Rosulek. *The Joy of Cryptography*. Oregon State University, 09 2017.

[20] U. Rührmair, F. Sehnke, J. Sölter, G. Dror, S. Gideon, and J. Schmidhuber. Modeling attacks on physical unclonable functions. In *Proceedings of the 17th ACM Conference on Computer and Communications Security*, pages 237–249. Association for Computing Machinery, 10 2010.

[21] U. Rührmair and M. van Dijk. PUFs in security protocols: Attack models and security evaluations. In *2013 IEEE Symposium on Security and Privacy*, pages 286–300. IEEE, 05 2013.

[22] K. K. Soni and A. Rasool. Cryptographic attack possibilities over RSA algorithm through classical and quantum computation. In *2018 International Conference on Smart Systems and Inventive Technology*, pages 11–15. IEEE, 12 2018.

[23] M. Velema. Classical encryption and authentication under quantum attacks. *arXiv e-prints: 1307.3753*, 07 2013.

[24] B. Škorić. Quantum readout of physical unclonable functions. In *2010 International Conference on Cryptology in Africa*, pages 369–386. Springer Berlin Heidelberg, 05 2010.

[25] B. Škorić. Security analysis of quantum-readout PUFs in the case of challenge-estimation attacks. *Quantum Information & Computation*, 16(1):50–60, 01 2016.

[26] B. Škorić, A. P. Mosk, and P. W. H. Pinkse. Security of quantum-readout PUFs against quadrature based challenge estimation attacks. *International Journal of Quantum Information*, 11:84, 06 2013.

[27] P. Wallden and Elham Kashefi. Cyber security in the quantum era. *Communications of the ACM*, 62(4"):120–129, 4 2019.

[28] X. Xu and W. Burleson. Hybrid side-channel/machine-learning attacks on PUFs: A new threat? In *2014 Design, Automation & Test in Europe Conference & Exhibition*, pages 1–6. IEEE, 03 2014.

[29] Y. Yao, M. Gao, M. Li, and J. Zhang. Quantum cloning attacks against PUF-based quantum authentication systems. *Quantum Information Processing*, 15(8):3311–3325, 05 2016.

[30] H. Zhu. Survey of computational assumptions used in cryptography broken or not by Shor's algorithm, 12 2001. Masters Thesis, McGill University Montréal.

Appendix A

Existing Attacks on QR-PUF ID Protocols

Challenge Estimation

Challenge estimation attacks are deemed the strongest class of classical attacks, subject to investigation by the original QR-PUF protocol [25, 26]. The attack measures the challenge state, then constructs a response based on the measurement outcome and public PUF information. [25, 26] conclude that security against challenge estimation is straightforward to achieve through careful configuration of the setup (details omitted for brevity), and in the case of a multiple-round protocol can be more readily attained by increasing the number of rounds.

Quantum Cloning

Quantum cloning attacks exploit the case in which the QR-PUF unitary transformation is public knowledge, and were demonstrated to outperform challenge estimation [29]. Quantum cloning attacks employ a *quantum cloning machine* to intercept a challenge state and produce a sufficient number of clones on which to perform the PUF unitary such that the adversary can extract enough information to produce an accurate representation of the true response. Investigation into attack probability for a variety of calibrations suggested that both the dimension of quantum states and the number of copies used in the QR-PUF protocol were important resources for protocol security to a varying degree. Thus, through increasing both the attack success probability can be bound arbitrarily close to zero. By tying the dimension of quantum states to the PUF security parameter λ in our model, one can increase security by increasing λ.

Quantum Teleportation

The quantum teleportation attack explored in [24] is executed as follows: transform the challenge state $|\psi\rangle$ into qubits and transfer it to the working memory of a quantum computer; program the quantum computer to perform a computation that simulates applying $\mathrm{U}_{\text{QR-PUF}}$ to the qubits; teleport the result of the computation from qubits back to a response state, and send it to $\mathcal{V}$ over the quantum channel. This attack is perhaps the most direct attempt at impersonation, however, they prevent an adversary from executing by imposing that it is either technically or financially infeasible to acquire both a sufficiently powerful quantum computer and a sufficiently fast method of transforming challenge states into qubits (and the computation results back into response states). This restriction is summarised by *quantum-computational unclonability* of the PUF.

Intercept-Resend

Intercept-resend attacks involve the adversary performing measurements on strategically chosen bases during each round of a multiple-round QR-PUF protocol, and preparing a certain state to send to $\mathcal{V}$ [24]. We omit attack details, though it can be found in full in [24]. The success probability per protocol round is bounded, so the overall probability can be bounded arbitrarily close to zero by increasing the number of rounds. Regardless of the success probability through this reasoning, however, it is known that performing a measurement on a quantum channel disrupts the contents of the transmission itself – which almost guarantees detection. An adversary can not run this attack reliably since they would not be able to fully determine the state of the unknown quantum challenge and therefore can not reliably emulate the correct response.

Summary

We can surmise from these attacks that the primary method of attack on any identification protocol that we are concerned with is impersonation. The general quantum emulation attack, defined in [3], is a generalisation of previously studied superposition and entanglement attacks, employing the quantum emulation algorithm defined in [15] to successfully emulate the correct output from a primitive for a given input – essentially a general quantum method to achieve impersonation. There are a variety of other attacks that have been investigated in the literature, however, they exploit vulnerabilities that arise from practical implementation of PUFs and their composition with adjoining systems, and a number of other aspects not within the scope of this work.

Appendix B

Protocol Cost Analysis

Outlined below are general protocol costs as dependent on PUF security parameter λ. It would be infeasible to outline costs in real-terms as it depends on the efficiency of the technologies involved and the algorithms employed, however, what is important for analysing the efficiency of our protocols is how costs generally increase with λ.

Enrollment

cPUF

– $\mathcal{V}$ performs $n = \text{poly}(\lambda)$ PUF queries of bit-size λ, a classical communication cost of $O(n\lambda) = \text{poly}(\lambda)$.
– $\mathcal{V}$ stores $n = \text{poly}(\lambda)$ CRPs consisting of 2 λ bit-size classical strings, a classical storage cost of $O(2n\lambda) = \text{poly}(\lambda)$.

QR-PUF

– $\mathcal{V}$ prepares $n = \text{poly}(\lambda)$ quantum states from a Hilbert space of dimension λ, a quantum computational cost of $n\mathcal{Q}$.
– $\mathcal{V}$ performs $n = \text{poly}(\lambda)$ PUF queries of quantum states from a Hilbert space of dimension λ, a quantum communication cost of $O(n\lambda) = \text{poly}(\lambda)$.
– $\mathcal{V}$ decodes $n = \text{poly}(\lambda)$ quantum states to their classical description using the received bases, a quantum computational cost of $n\mathcal{D}$.
– $\mathcal{V}$ stores $n = \text{poly}(\lambda)$ CRPs consisting of 2 λ bit-size classical strings, a classical storage cost of $O(2n\lambda) = \text{poly}(\lambda)$.

qPUF

– $\mathcal{V}$ prepares $n = \text{poly}(\lambda)$ quantum states from a Hilbert space of dimension λ, a quantum computational cost of $n\mathcal{Q}$.
– $\mathcal{V}$ performs κ_1 PUF queries for each of $n = \text{poly}(\lambda)$ quantum states from a Hilbert space of dimension λ, a quantum communication cost of $O(\kappa_1 n\lambda) = \text{poly}(\lambda)$.
– $\mathcal{V}$ stores $n = \text{poly}(\lambda)$ CRPs consisting of 1 λ bit-size classical string and κ_1 copies of a quantum state, a classical storage cost of $O(n\lambda) = \text{poly}(\lambda)$ and quantum storage cost of $O(\kappa_1 n\lambda) = \text{poly}(\lambda)$.

Time-Bounded qPUF

– $\mathcal{V}$ prepares 1 state from a Hilbert space of dimension λ, a quantum computational cost of $\mathcal{Q}$.
– $\mathcal{V}$ performs κ_1 PUF queries of a quantum state from a Hilbert space of dimension λ, a quantum communication cost of $O(\kappa_1\lambda)$.
– $\mathcal{V}$ stores 1 CRP consisting of 1 λ bit-size classical string and κ_1 copies of a quantum state, a classical storage cost of $O(\lambda)$ and quantum storage cost of $O(\kappa_1\lambda)$.

Adaptive CRPs qPUF

– $\mathcal{V}$ prepares $m = \text{poly}(\lambda)$ quantum states from a Hilbert space of dimension λ, a quantum computational cost of $m\mathcal{Q}$.
– $\mathcal{V}$ performs κ_1 PUF queries for each of $m = \text{poly}(\lambda)$ quantum states from a Hilbert space of dimension λ, a quantum communication cost of $O(\kappa_1 m\lambda) = \text{poly}(\lambda)$.
– $\mathcal{V}$ stores $m = \text{poly}(\lambda)$ CRPs consisting of 1 λ bit-size classical string and κ_1 copies of a quantum state, a classical storage cost of $O(m\lambda) = \text{poly}(\lambda)$ and quantum storage cost of $O(\kappa_1 m\lambda) = \text{poly}(\lambda)$.

Identification & Verification

cPUF

Verifier

– $\mathcal{V}$ queries $\mathcal{P}$ with one classical string of bit-size λ, a classical communication cost of $O(\lambda)$.

– $\mathcal{V}$ makes one comparison between two classical strings, a classical computation cost of $O(1)$.

Prover

– $\mathcal{P}$ performs 1 PUF query of bit-size λ, and returns 1 classical string of bit-size λ to $\mathcal{V}$, a classical communication cost of $O(2\lambda)$.

QR-PUF

Verifier

– $\mathcal{V}$ queries $\mathcal{P}$ with one quantum state from a Hilbert space of dimension λ, a quantum communication cost of $O(\lambda)$.
– $\mathcal{V}$ prepares 2 quantum states from a Hilbert space of dimension λ, a quantum computational cost of $2\mathcal{Q}$.
– $\mathcal{V}$ performs a projective quantum measurement with cost $\mathcal{M}$.

Prover

– $\mathcal{P}$ performs 1 PUF query of a quantum state from a Hilbert space of dimension λ, and returns 1 state to $\mathcal{V}$, a quantum communication cost of $O(2\lambda)$.

qPUF

Verifier

– $\mathcal{V}$ queries $\mathcal{P}$ with κ_2 copies of a quantum state from a Hilbert space of dimension λ, a quantum communication cost of $O(\kappa_2\lambda)$.
– $\mathcal{V}$ prepares κ_2 copies of a quantum state from a Hilbert space of dimension λ, a quantum computational cost of $\kappa_2\mathcal{Q}$.
– $\mathcal{V}$ performs a quantum state equality test with cost $\mathcal{T}$.

Prover

– $\mathcal{P}$ performs κ_2 PUF queries of a quantum state from a Hilbert space of dimension λ, and returns κ_2 states to $\mathcal{V}$, a quantum communication cost of $O(2\kappa_2\lambda)$.