# The Quantum Random Oracle Model

Advisor: Serge Fehr

A *cryptographic hash function* is a function $h : \{0,1\}^m \to \{0,1\}^n$ (where typically $n < m$) that is "hard to analyze". Intuitively, we want that the only way to learn information on $h(x)$ for any particular $x \in \{0,1\}^n$ is to actually *compute* $h(x)$, and for every $x$ on which $h$ has not been computed yet $h(x)$ should "look like" a fresh random value in $\{0,1\}^n$.

The *random oracle model* is a way to formally model these intuitive requirements when analyzing the security of cryptographic schemes (like encryption or digital signature schemes) that use a cryptographic hash function $h$ as a building block. In the random oracle model the hash function $h$ is modeled as an *oracle* (referred to as the *random oracle*) that needs to be *queried* on $x \in \{0,1\}^m$ in order for an entity to learn the hash value $h(x)$, and the random oracle's reply $h(x)$ is with respect to a *uniformly random* function $h : \{0,1\}^m \to \{0,1\}^n$.

This model turns out to be very useful for analyzing the security of cryptographic schemes. For instance, it allows us to bookkeep the attackers queries and so we precisely know the hash values that the attacker knows. This is very helpful for proving security. The hope then is that a cryptographic scheme that is proven secure in the random oracle model is still secure "in the real world" where $h$ is a publicly-known cryptographic hash function. Even though this random oracle heuristic cannot be proven, it works very well in practice.

It turns out that the random oracle model needs to be revisited in the context of *quantum attacks*, i.e., when we consider attackers that are equipped with a quantum computer. Such a hypothetical quantum computer would allow the attacker to compute the hash function $h$ in *superposition* over multiple $x$'s, i.e., it can produce states like $\sum_x \alpha_x |x\rangle|h(x)\rangle$, which depend on $h(x)$ for different $x$'s, by computing $h$ once. Therefore, in the random oracle model, the attacker may query the random oracle in superposition as well. However, this then obstructs the typical reasoning. For instance, bookkeeping (i.e. storing copies of) the queries made by the attacker is now not possible anymore by the quantum no-cloning theorem.

A recent article [1] shows an approach for overcoming some of the obstacles caused by allowing the attacker to query the random oracle in superposition. By considering the random oracle to start off with a superposition of all possible functions $h : \{0,1\}^m \to \{0,1\}^n$ (rather than choosing such a function uniformly at random), and by writing its evolving internal quantum state in terms of a conveniently chosen basis for the underlying state space, it becomes possible again to figure out, in some sense, the hash values that the attacker knows.

Unfortunately, the explanations in the article [1] are rather informal. So, the goal of this project is to translate the claims and explanations from [1] into clean mathematical statements and proofs. Depending on the progress, we may also try to find further applications than those considered in [1].

For this project it is necessary to have a fair background in quantum information science, e.g. by having attended a course in quantum computing (like my course last semester). Having some knowledge about cryptography is a plus but not necessary.

## References


1. Mark Zhandry. *How to Record Quantum Queries, and Applications to Quantum Indifferentiability.* Available from `https://eprint.iacr.org/2018/276`.