Computationally binding quantum commitments

Dominique Unruh
University of Tartu

Abstract. We present a new definition of computationally binding commitment schemes in the quantum setting, which we call "collapse-binding". The definition applies to string commitments, composes in parallel, and works well with rewinding-based proofs. We give simple constructions of collapse-binding commitments in the random oracle model, giving evidence that they can be realized from hash functions like SHA-3. We evidence the usefulness of our definition by constructing three-round statistical zero-knowledge quantum arguments of knowledge for all NP languages.

We study the definition and construction of computationally binding string commitment schemes in the quantum setting. A commitment scheme is a two-party protocol consisting of two phases, the commit and the open phase. The goal of the commitment is to allow the sender to transmit information related to a message m during the commit phase in such a way that the recipient learns nothing about the message (hiding property). But at the same time, the sender cannot change his mind later about the message (binding property). Later, in the open phase, the sender reveals the message m and proves that this was indeed the message that he had in mind earlier. We will focus on non-interactive classical commitments, that is, the commit and open phase consists of a single classical message. However, the adversary who tries to break the binding or hiding property will be a quantum-polynomial-time algorithm. At the first glance, it seems that the definition of the binding property in this setting is straightforward; we just take the classical definition but consider quantum adversaries instead of classical ones:

Definition 1 (Classical-style binding – informal) *No quantum-polynomial-time algorithm A can output, except with negligible probability, a commitment c (i.e., the message sent during the commit phase) as well as two openings u, u' that open c to two different messages m, m'.*

Unfortunately, this definition turns out to be inadequate in the quantum setting. Ambainis, Rosmanis, and Unruh [1] show the existence of a commitment scheme (relative to a special oracle) such that: The commitment is classical-style binding. Yet there exists a quantum-polynomial-time adversary A that outputs a commitment c, then expects a message m as input, and then provides valid opening information for c and m. That is, the adversary can open the commitment c to any message of his choosing, even if he learns that message only after committing. This is in clear contradiction to the intuition of the binding property. How is this possible, as Definition 1 says that the adversary cannot produce two different openings for the same commitment? In the construction from [1], the adversary has a quantum state $|\Psi\rangle$ that allows him to compute one opening for a message of his choosing, however, this computation will destroy the state $|\Psi\rangle$. Thus, the adversary cannot compute two openings simultaneously, hence the commitment is classically-binding. But he can open the commitment to an arbitrary message once, which shows that the commitment scheme is basically useless despite being classically-binding.[1]

[1] Note that for classical adversaries, the classical-binding property gives useful guarantees: If an adversary can produce an opening for any message m using some classical algorithm, he can also produce two openings for different messages m, m' by running that algorithm twice.

Prior definitions. A number of definitions occur in the literature that address the above problem. The simplest definition (for bit commitments, i.e., $m \in \{0,1\}$) is what we call the "sum-binding" definition: For a given quantum-polynomial-time adversary, let p_0 denote the probability that the adversary commits and opens to $m = 0$, and let p_1 denote the probability that the adversary commits and opens to $m = 1$. Assume that the adversary learns whether he should open to 0 or 1 only after the commit phase. We call a scheme sum-binding if $p_0 + p_1 \leq 1 + \textit{negligible}$. This definition circumvents the above problems, however, it is only meaningful for bit commitments. A generalization to string commitments is the CDMS-binding property from [2]. However, that definition is parametrized over a specific family of predicates that need to be chosen specific for each application, and it is not clear whether the definition composes in parallel (i.e., if we commit to m_1 and m_2, is this a binding commitment to (m_1, m_2)?) Also, it seems that CDMS-binding commitments are difficult to use in a setting where quantum rewinding is used. For more details and justification, see the full version [5]. Alternatively, perfectly binding commitments or UC commitments can be used to circumvent the above problem. However, for constructing those, we need considerably stronger assumptions, so a weaker definition that that is general and parallel composes would be very useful.

Our contribution. We give a new definition for the computational-binding property for commitment schemes, called "collapse-binding". This definition is composable (several collapse-binding commitments are also collapse-binding together), works well with quantum rewinding (see below), does not conflict with statistical hiding (as perfectly-binding commitments would), allows for short commitments (i.e., the commitment can be shorter than the committed message, in contrast to perfectly-binding commitments, and to extractable commitments in the CRS model). Basically, collapse-binding commitments seem to be in the quantum setting what computationally-binding commitments are in the classical setting.

We show that collision-resistant hash functions are not sufficient for getting collapse-binding or even just sum-binding commitments, at least when using standard constructions, and relative to an oracle. We present a strengthening of collision-resistant hash functions, "collapsing hash functions" that can serve as a drop-in replacement for collision-resistant hash functions. Using collapsing hash functions, we show several standard constructions of commitments to be collapse-binding.

We conjecture that standard cryptographic hash functions such as SHA-3 [3] are collapsing (and thus lead to collapse-binding commitments). We give evidence for this conjecture by proving that the random oracle is a collapsing hash function.

We show that the definition of collapse-binding commitments is usable by extending the construction of quantum proofs of knowledge from [4]. Their construction uses perfectly-binding commitments (actually, strict-binding, which is slightly stronger) to get proofs of knowledge. We show that when replacing the perfectly-binding commitments with collapse-binding ones, we get statistical zero-knowledge quantum arguments of knowledge. In particular, this shows that collapse-binding commitments work well together with rewinding.

Collapse-binding commitments. To explain the definition of collapse-binding commitments, first consider a perfectly-binding commitment. That is, when an adversary A outputs a commitment c, there is only one possible message m_c that A can open c to. Hence, if the adversary A outputs a superposition of messages that he can open c to, that superposition will necessarily be in the state $|m_c\rangle$. Hence, we can characterize perfectly-binding commitments by requiring: when an adversary outputs a superposition of messages that he can open the commitment c to, that superposition will necessarily be a single computational basis vector (i.e., no non-trivial superposition).

To express this more formally, consider the circuit in Figure 1 (a). Here the adversary A outputs a commitment c (classical message). Furthermore, he outputs three quantum registers

S, U, M. S contains his state. M is supposed to contain a superposition of messages, U a superposition of corresponding opening informations. Then we apply the measurement V_c. This measurement measures whether U, M contain matching opening information/message. More formally, V_c measures whether U, M is a superposition of states $|u, m\rangle$ such that u is valid opening information for message m and commitment c. Let $ok = 1$ if the measurement succeeds. Then we feed the registers S, U, M back to the second part B of the adversary. B outputs a classical bit b. As discussed before, a commitment is perfectly-binding iff for all adversaries A, the state of M after measuring $ok = 1$ is a computational basis vector.

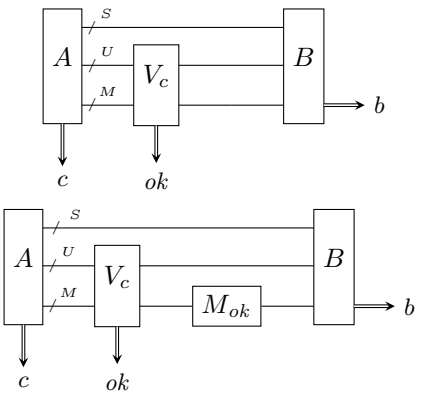

Figure 1: Games from the definition of collapse-binding commitments.

The state of a register is a computational basis vector (or, synonymously: is in a collapsed state) iff measuring that register in the computational basis does not change that state. Consider the circuit in Figure 1 (b). Here we added a measurement M_{ok} on M after V_c. M_{ok} is a complete measurement in the computational basis, but is executed only if $ok = 1$. Since M_{ok} disturbs the state of M iff that state is not a computational basis vector, we can rephrase the definition of perfectly-binding commitments:

A commitment is perfectly-binding iff, for all computationally unlimited adversaries A, B, $\Pr[b = 1]$ is equal in Figures 1 (a) and 1 (b) where b is the output (i.e., guess) of B.[2]

Now we are ready to weaken this characterization to get a computational binding property. Basically, we require that the same holds for quantum-polynomial-time adversaries:

Definition 2 (Collapse-binding – informal) *A commitment is collapse-binding iff, for all quantum-polynomial-time adversaries A, B, $\Pr[b = 1]$ in Figure 1 (a) is negligibly close to $\Pr[b = 1]$ in Figure 1 (b).*

In other words, with a perfectly-binding commitment, the adversary cannot produce a superposition of different messages that are contained in the commitment. But with a collapse-binding commitment, the adversary is forced to produce a state *that looks like it is not a superposition* of different messages. For the purpose of computational security, this will often be as good.

For the other results (in particular constructions and evidence of the usability of the definition), see [5]. Related work is also discussed there.

References.

[1] A. Ambainis, A. Rosmanis, and D. Unruh. Quantum attacks on classical proof systems (the hardness of quantum rewinding). In *FOCS 2014*, pages 474–483. IEEE, 2014. Preprint on IACR ePrint 2014/296.

[2] C. Crépeau, P. Dumais, D. Mayers, and L. Salvail. Computational collapse of quantum state with application to oblivious transfer. In *TCC 2004*, volume 2951 of *LNCS*, pages 374–393. Springer, 2004.

[3] NIST. SHA-3 standard: Permutation-based hash and extendable-output functions. Draft FIPS 202, 2014. Available at `http://csrc.nist.gov/publications/drafts/fips-202/fips_202_draft.pdf`.

[4] D. Unruh. Quantum proofs of knowledge. In *Eurocrypt 2012*, volume 7237 of *LNCS*, pages 135–152. Springer, April 2012. Full version is IACR ePrint 2010/212.

[5] D. Unruh. Computationally binding quantum commitments. IACR ePrint 2015/361, 2015.

[2] Our exposition above was not very rigorous, but it is easy to see that this is indeed an "if and only if".