On Quantum Obfuscation

Bill Fefferman (QuICS, University of Maryland/NIST)
Joint with Gorjan Alagic (QMATH, University of Copenhagen)

Based on arxiv:1602.01771

Overview

1. Definitions
2. Applications
3. Feasibility?

I. Definitions

- A classical *Black-box* Obfuscator is an algorithm **O**:
  - Input is a circuit C with input length *n*
  - Outputs a circuit **O**(C) so that:
    1. *"Functionality"* of **O**(C) is the same as C
       - C(x)= **O**(C)(x) for all inputs x
    2. *"Efficiency"* is preserved
       - size(**O**(C)) ≤ *poly(n)*
    3. *"Black-box Obfuscation"* property
       - "Anything that can be efficiently learned about **O**(C) can just as well be learned from black-box access to C"
         - For any "adversary algorithm" A there exists "simulator algorithm" S so that for all circuits C:
           - $|\Pr[A(\mathbf{O}(C))=1]-\Pr[S^C(1^{size(C)})=1]| < negl(size(C))$
- What should this mean *quantumly*?
  - The *input circuit*, the *Adversary*, the *Simulator*, and the *Obfuscator* itself should be quantum algorithms
  - The output of the obfuscation, **O**(C) will be a poly(n) qubit quantum state
    - That gains functionality through an "interpreter" algorithm J
      - I.e, all input states σ, $|J(\mathbf{O}(C), \sigma)-C\sigma C^{\dagger}|_{tr} < negl(n)$

II. *Informal Sketches* of Applications

- Transforming Private-key encryption scheme into Public-key encryption scheme
  - ***Idea***: Publish the obfuscation of the private key Encryption algorithm, Enc_k
    - Everyone can encrypt!
    - Only secret key holder decrypts
- Fully homomorphic encryption
  - ***Idea***: Suppose we want to perform some computation on encryptions of two bits
    - Take some public-key encryption scheme, use secret key to construct algorithm that performs the computation
      - By decrypting, applying operation, encrypting outcome
    - Publish the obfuscation of this algorithm along with public key
- Public-key quantum money
  - ***Goal***:
    - A mint "produces" bills in the form of quantum states
    - Everyone can verify authenticity
    - No-one can copy (using no-cloning theorem)

III. Feasibility of obfuscation?

Classical Black-box Impossibility proof (1/3)

- *Theorem [Barak et. al., '01]*: There exist circuits that cannot be Black-box obfuscated.
- Barak et. al., constructs a circuit from which an adversary given **O**(C) gains more information than a simulator could using black-box access to C
- ***Proof idea***:
  - Choose a,b $\in_R \{0,1\}^n$
  - Consider two pairs of circuits:
  1. *First pair*:

$$C_{a,b}(x) = \begin{cases} b & \text{if } x = a \\ 0^n & \text{otherwise.} \end{cases} \qquad D_{a,b}(\mathcal{C}) = \begin{cases} 1 & \text{if } \mathcal{C}(a) = b \\ 0 & \text{otherwise.} \end{cases}$$

  2. *Second pair:*

$$Z(x) = 0^n \text{ for all } x. \qquad D_{a,b}(\mathcal{C}) = \begin{cases} 1 & \text{if } \mathcal{C}(a) = b \\ 0 & \text{otherwise.} \end{cases}$$

  - *Key Point*:
    - Can efficiently distinguish inputs **O**($C_{a,b}$) and **O**($D_{a,b}$) from inputs **O**(Z) and **O**($D_{a,b}$)
      - Run them on each other!
    - But any simulator with black-box access to either pair (who is ignorant of a,b) can't do this!

Classical Black-box Impossibility proof (2/3)

- How to go from pairs of circuits to single circuits?
- Create "combined circuits" that use an additional input bit
  - $F_{a,b}$ is combination of $C_{a,b}$ and $D_{a,b}$
  - $G_{a,b}$ is combination of Z and $D_{a,b}$
- An adversary given as input either $O(F_{a,b})$ or $O(G_{a,b})$ can tell them apart
  - Make a copy of the obfuscation and use this copy to run the obfuscation on itself
- But this doesn't actually work!
  - Can't run a circuit on itself! The input register of is fixed length and not large enough
  - Fixing this requires most of the technical work in the [Barak et. al. '01] proof!

Classical impossibility proof (3/3)

- ***Goal***: need to modify $D_{a,b}$ so that:
  - Adversary can use it to test if given circuit C takes a to b
    - Needs to work even if description of C is longer than input length of $D_{a,b}$
    - Should keep a and b hidden from parties with only black-box access to $D_{a,b}$
- ***Solution***: Construct a new $D'_{a,b}$ that combines three circuits:
  - First circuit outputs encryption of a
  - Second circuit provides ability perform binary gates on encrypted bits
  - Third circuit tests whether a sequence of encryptions consists of the encryptions of the bits of b
- Why does this work?
  - If given **O**(C), we can test if C(a)=b using three new circuits
    - By using the second circuit to homomorphically apply each gate of C to the encryption of a
  - If we only have black-box access to $D'_{a,b}$, cannot learn a and b
    - Follows from *IND-CCA1* security of encryption scheme (which can be constructed from a OWF)
- Shows OWF⇒Black-box obfuscation is impossible
- Can also prove that Efficient Black-box obfuscation⇒OWF (contradiction!)

Adapting to the quantum setting

- ***First case***: The quantum obfuscation has classical outputs
  - Not hard to make "unitary versions" of Barak's circuits $F_{a,b}$ and $G_{a,b}$
  - Run into the same problem as before: how does adversary distinguish the quantum circuit $\mathbf{O}(F_{a,b})$ from the quantum circuit $\mathbf{O}(G_{a,b})$?
      - Similar solution: construct a modified *quantum* circuit that "homomorphically" runs a given *quantum* circuit on encryption of a and checks if the output is an encryption of b!
      - Needs a construction of IND-CCA1 private key encryption on *quantum states* (because our simulated quantum computation at any time is in some quantum state)!
    - For other computational notions of encryption on quantum plaintext see our paper "Computational Security of Quantum Encryption" [F., with Alagic, Broadbent, Gagliardoni, Schaffner, St. Jules] **Also at this QCrypt**!
- ***Second case***: What happens if the obfuscator outputs quantum states?
  - Here the no-cloning theorem forbids us from copying obfuscation state
  - In the case that the output of the obfuscation is "reusable" can still achieve impossibility

Statistical indistinguishability obfuscation?

- *Statistical i.o property*: for functional equivalent C_1,C_2 the obfuscations ρ_1=**O**(C_1) and ρ_2=**O**(C_2) are negligible in trace distance.
- Impossibility of *quantum* statistical I.O (unless **QSZK**=**PSPACE**)
  - Two problems:
    1. "Quantum circuit distinguishability": Given two quantum C_0 and C_1 are they functionally similar (i.e., in diamond norm)?
       - This is **PSPACE**-complete [Rosen and Watrous '05]
    2. "Quantum state distinguishability" Given two efficiently preparable quantum states ρ_0 and ρ_1 are they close in trace norm?
       - This is **QSZK**-complete [Watrous'02]
  - Suppose we have efficient quantum statistical I.O algorithm **O**
    - Given instance of Quantum Circuit Distinguishability, C_0 and C_1
    - Consider obfuscations **O**(C_0) and **O**(C_1)
      - If C_0 and C_1 are functionally the same then **O**(C_0) is close to **O**(C_1) in trace norm (by obfuscation property of **O**)
      - If C_0 and C_1 are functionally different then can show that **O**(C_0) is far in trace norm from **O**(C_1) (using functional equivalence of obfuscation!)
    - So we've reduced Quantum Circuit Distinguishability to Quantum State Distinguishability and **PSPACE**$\subseteq$**QSZK**

Surviving notions of quantum obfuscation

1. Quantum black-box obfuscation with *uncloneable* output
   - Many of our applications survive!
2. Quantum *Computational* Indistinguishability Obfuscation
   - i.e., if C_1 and C_2 are functionally equivalent then $O(C_1)$ and $O(C_2)$ are *computationally indistinguishable* (using definition of [Watrous'08])
   - ***Application***: Quantum "Witness Encryption" for **QMA**
     - Classically the existence of "Witness Encryption" for **NP** would have many useful applications [Garg, Gentry, Sahai and Waters'13]
       - E.g., Public key encryption from PRGs, Identity-based Encryption, Attribute-based Encryption etc…
   - Classically, there are candidate indistinguishability obfuscation constructions e.g., [Garg, Gentry, Halevi, Raykova, Sahai, Waters '13]. Can we find quantum I.O constructions?

Thanks!