# Dispelling Myths on Superposition Attacks: Formal Security Model and Attack Analyses


Luka Music[1], Céline Chevalier[2] and Elham Kashefi[1,3]

1 - Département Informatique et Réseaux, CNRS, Sorbonne Université
2 - Université Panthéon-Assas Paris 2
3 - School of Informatics, University of Edinburgh



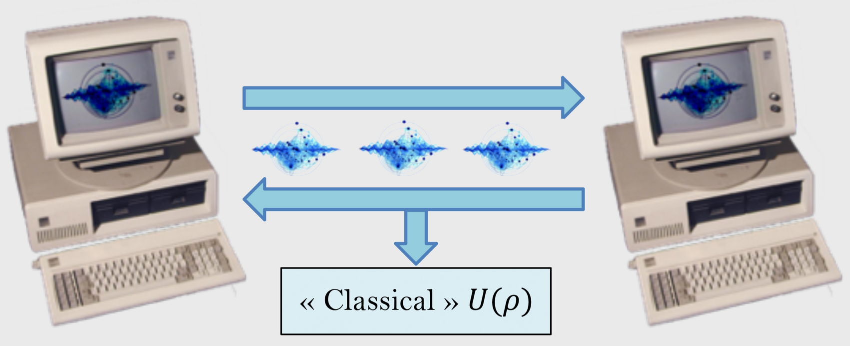
arXiv:2007.00677


## Take-away

- New model for computational security against superposition attacks
- Idea: Superposition-resistance means Adversary can do nothing more than in classical protocol
- Superposition attacks on unconditionally-secure protocols do not translate to computational setting
- Subtle vectors for attacks mean composable frameworks are likely impossible
- Secure protocols exist:
  - Classical One-Time-Pad
  - Yao's 2PC protocol

## Incompleteness of Anterior Models

Quantum Protocols/Classical Functionalities



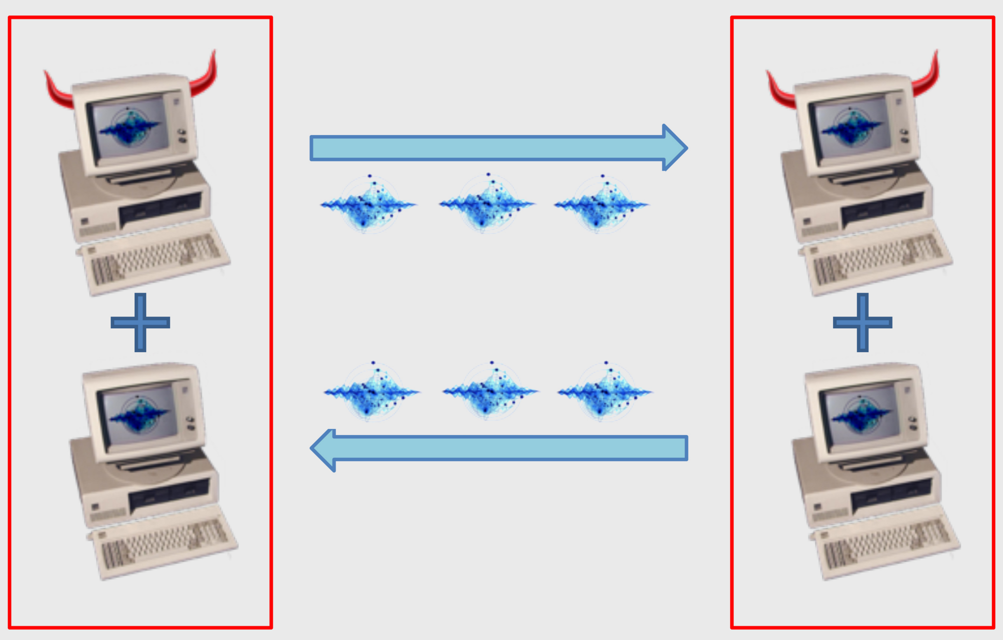

- Initial state: $|\phi\rangle$
- Ancilla for each new message: $|\phi\rangle|0\rangle$
- Classical operations: $U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$
- Result on superposition of inputs:

$$\sum_{x,y}|x\rangle|y\rangle|g^1_{x,y}\rangle|g^2_{x,y}\rangle|f(x,y)\rangle$$

Unwanted Entangled Garbage

- [1] Perfect protocols reduce to: $\sum_{x,y}|x\rangle|y\rangle|f(x,y)\rangle$
- Result: All non-trivial protocols are broken
- Pb: Not applicable to computational setting

### [2]: Corruption Oracle in Superposition

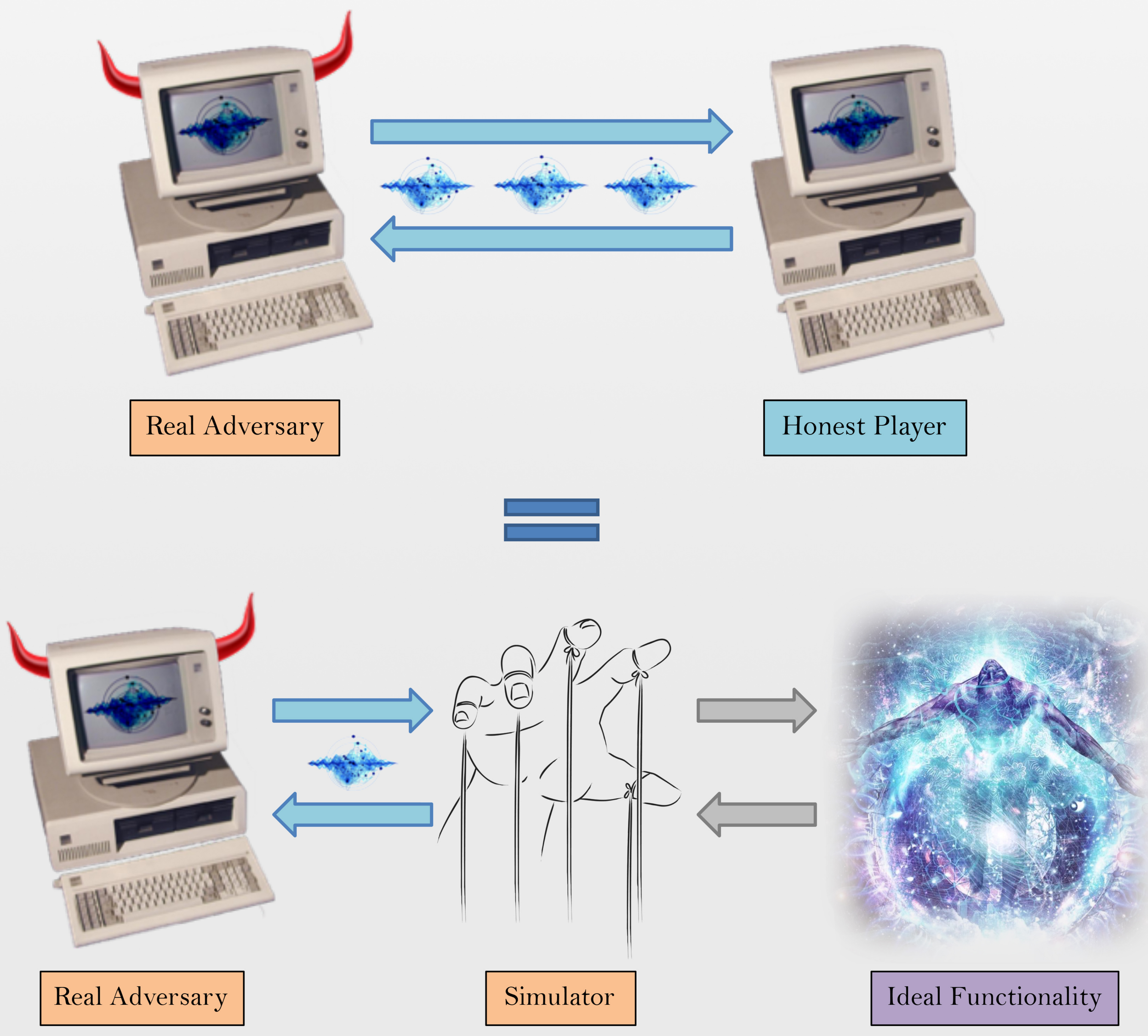


- Result: Non-trivial protocols cannot be simulated
- Pb 1: Not equivalent to static Adversary
- Pb 2: Honest player must have superposed input


[1]: Salvail, Schaner, Sotakova. Quantifying the leakage of quantum protocols for classical two-party cryptography. International Journal of Quantum Information, 13(04):1450041, 2015.

[2]: Damgard, Funder, Nielsen, Salvail. Superposition attacks on cryptographic protocols. Information Theoretic Security, 2014.


## Computational Superposition Security



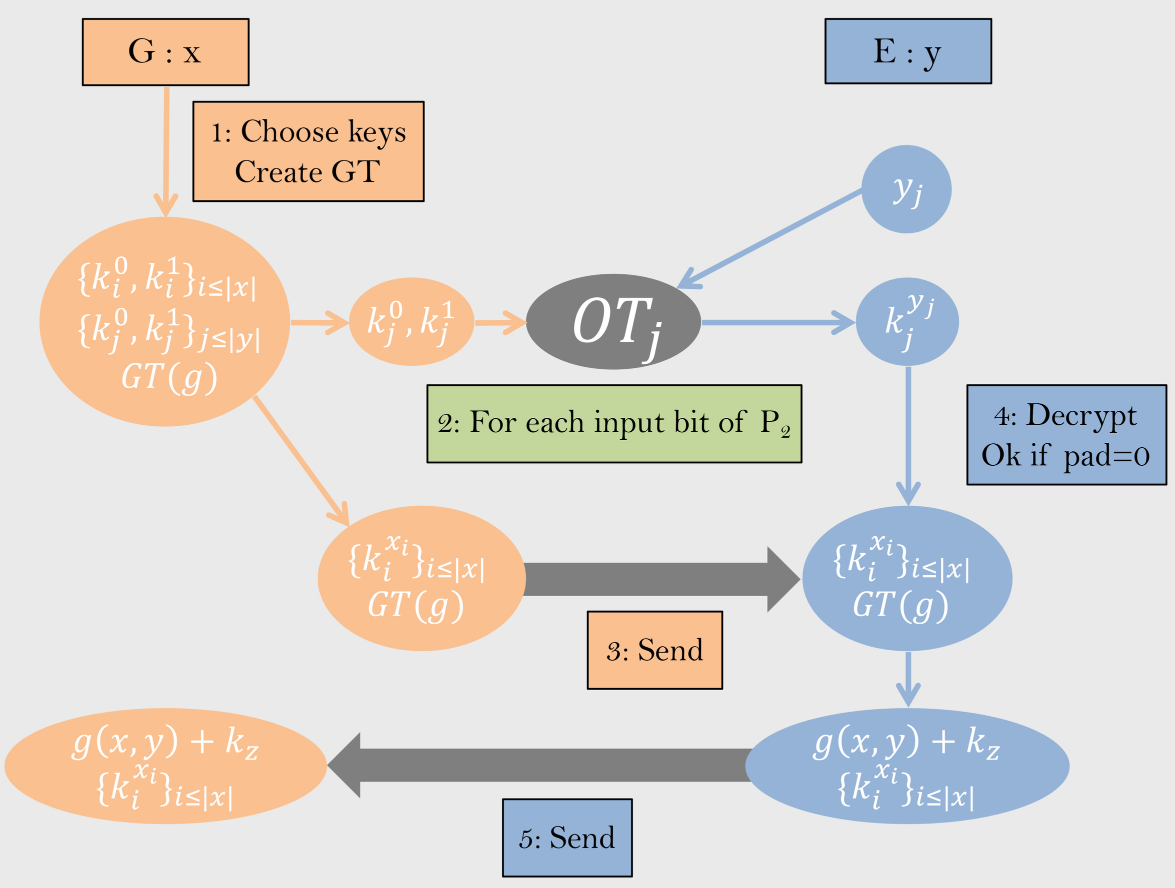

- Adversary fixed at start, honest classical input
- Principle: Superposition-resistance of protocol if it is not affected by adversarial superposition
- Perfect superposition resistance if purely classical messages ⇒ Ideal Functionality purely classical
- Simulator has no superposition access to Ideal Functionality but indistinguishable to Adversary with superposition access

## Yao Two-Party Computation Protocol

- Garbler(x) & Evaluator(y) wish to compute g(x, y)
- Uses Symmetric Encryption & Oblivious Transfer
- Garbled Table (GT) for function $g : \{0,1\}^2 \to \{0,1\}$

$$E_1^{k_z} := \mathsf{Enc}_{k_0^a}\left(\mathsf{Enc}_{k_0^b}(g(0,0) \oplus k_z \parallel 0^p)\right)$$
$$E_2^{k_z} := \mathsf{Enc}_{k_0^a}\left(\mathsf{Enc}_{k_1^b}(g(0,1) \oplus k_z \parallel 0^p)\right)$$
$$E_3^{k_z} := \mathsf{Enc}_{k_1^a}\left(\mathsf{Enc}_{k_0^b}(g(1,0) \oplus k_z \parallel 0^p)\right)$$
$$E_4^{k_z} := \mathsf{Enc}_{k_1^a}\left(\mathsf{Enc}_{k_1^b}(g(1,1) \oplus k_z \parallel 0^p)\right)$$

### Full Protocol




### Modifications to Original Protocol

- GT computed by iterating over function domain
- G sends one copy of its keys for each GT entry
- E sends back G's keys if success
- Modifications do not impact security without superposition access

## Superposition Attack on Yao Protocol

- OT is perfectly classical
- Minimal Oracle Representation: $U_f|x\rangle = |f(x)\rangle$
- MOR exists for AES + CTR symmetric Enc/Dec, no need for ancillas, get same as perfect protocol:

$$\sum_{x,y}|x\rangle|y\rangle|f(x,y)\rangle$$

### Attack Sketch

- G honest until end of OTs, sends superposition of its keys and GT:

$$\frac{|k_{\hat{y}}^y\rangle}{\sqrt{2}}\left(|k_{\hat{x}_0}^x\rangle + |k_{\hat{x}_1}^x\rangle\right)\sum_i(-1)^{k_z}|E_i^{k_z}\rangle$$

- E (1) decrypts in superposition, (2) measures padding and returns if gets $0^p$:

$$\sum_{\hat{x},k_z}(-1)^{k_z}|k_{\hat{x}}^x\rangle|g(\hat{x},\hat{y}) \oplus k_z\rangle|0^p\rangle + |\mathsf{Garbage}\rangle$$

$$\sum_{\hat{x},k_z}(-1)^{k_z}|k_{\hat{x}}^x\rangle|g(\hat{x},\hat{y}) \oplus k_z\rangle$$

- After clean-up (that depends only on keys):

$$(-1)^{g(\hat{x}_0,\hat{y})}|0\rangle|-\rangle + (-1)^{g(\hat{x}_1,\hat{y})}|1\rangle|-\rangle$$

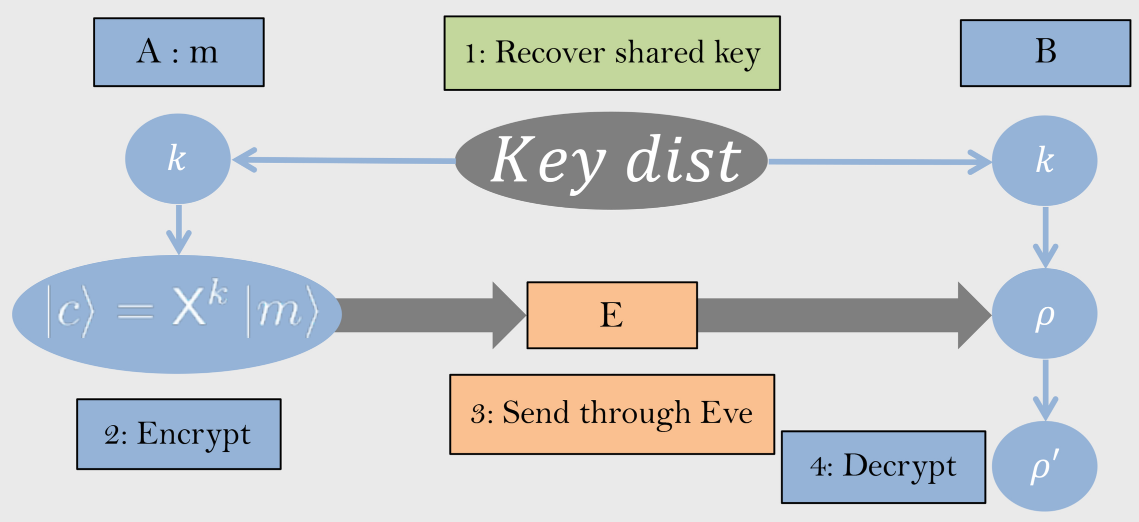
- Finally, apply Hadamard, measure in computational basis

### Attack Result

- Recover XOR of outputs for any two inputs of Adversary's choice (and fixed honest input)
- Success probability, independent of input and function: $p_{\mathcal{A}} = 1 - e^{-1}$
- Attack vector: Keys of G, returned by E after decryption
- Attack principle: make player implement DJ algorithm

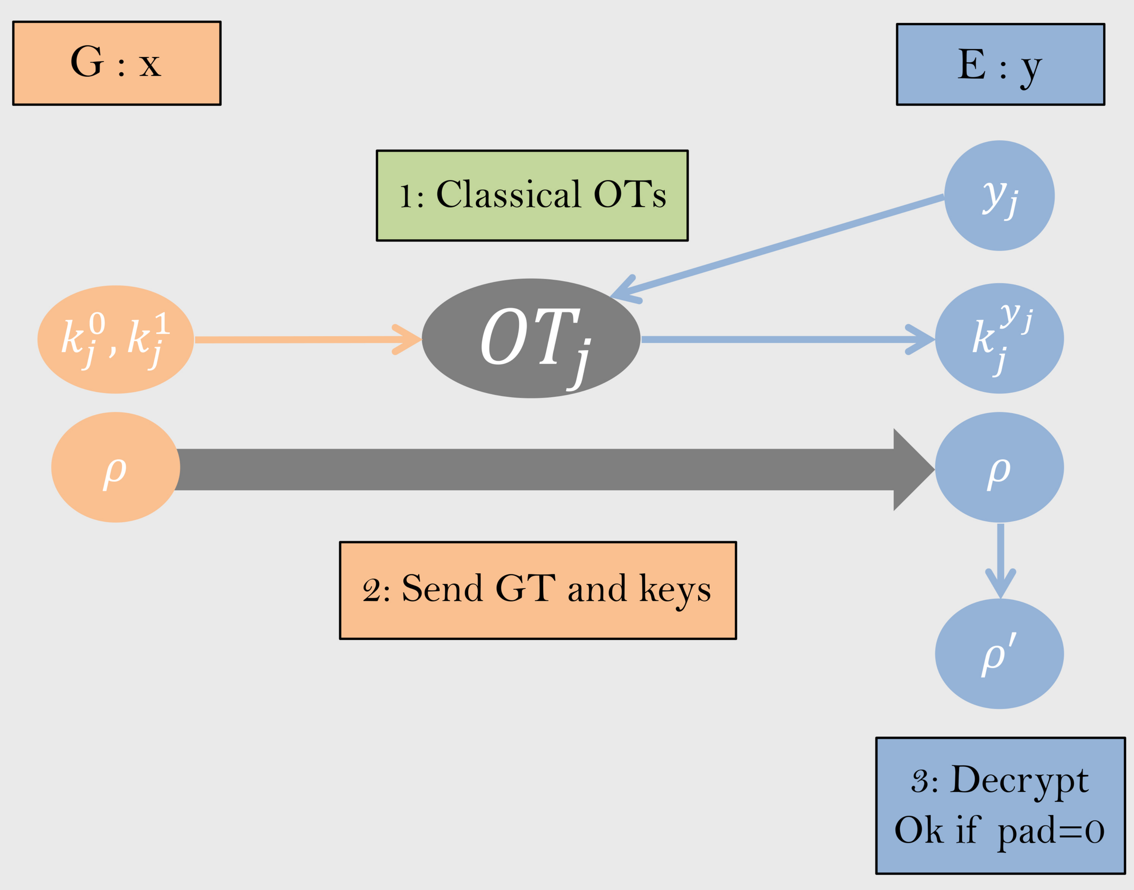
## Positive Security Results

### Classical One-Time Pad




- Local operations give no information to Eve

### Yao's Protocol with E's Output




- Same: Local operations give no information to G