Private Quantum Database

Giancarlo Gatti, Floris Geerts, Rihan Hai

MONDRAGON UNIBERTSITATEA · University of Antwerp · TU Delft, Delft University of Technology

We propose a framework to enhance classical databases with quantum-measurement-based functionalities, which allows us to perform private queries without fully revealing data contents (SPIR). Notably, small quantum systems are sufficient to encode and query exponentially large classical databases.

Symmetric Private Information Retrieval (SPIR)

ALICE (SERVER)

Full Name	Phone	Email
Anton Brown	(+31) 123-4567	anton.brown@gmail.com
Brian Smith	(+42) 234-5678	bsmith@yahoo.com
Cynthia Lee	(+51) 345-6789	cynthia.lee@outlook.com
Daniel Kim	(+64) 456-7890	dkim@icloud.com
Eva Martín	(+34) 567-8901	eva.martin@aol.com

BOB (USER) *Only knows the first column*

1. Alice has a database of **R** rows (index **r= 1, 2, … , R**).
2. Bob wants to query row **r** without revealing **r (user privacy)**.
3. Alice does not want to reveal other rows **(data privacy)**.
4. There is no trusted third party. Bob is honest but curious.

This is classically impossible for single server, but quantum methods can address this task.

Private Query Protocol (*R=5* example)

0) Prior agreements (independent of database contents):

A) Select q. system size

# qubits	# rows
1	≤ 3
2	≤ 5
3	≤ 9
4	≤ 17
5	≤ 33
n	$\leq 2^n+1$

B) Map rows to Mutually Unbiased Bases (MUBs)

Row #	Observables
1 (Anton)	$Z\otimes I,\ I\otimes Z$
2	$X\otimes I,\ I\otimes X$
3	$Y\otimes I,\ I\otimes Y$
4	$X\otimes Y,\ Y\otimes Z$
5 (Eva)	$Y\otimes X,\ Z\otimes Y$

(Observable order is an agreement as well)

C) Map possible measurement outcomes to *n*-bit strings

$(+1,+1)\to$ '00' $(+1,-1)\to$ '01'
$(-1,+1)\to$ '10' $(-1,-1)\to$ '11'

D) Define number *k* of quantum state copies to send to Bob (query limit).

1) Alice splits the database rows into *B* *n*-bit batches. Each batch $\mathcal{B}_i$ has *R* rows of *n* bits.

2) For batch *i*, Alice numerically computes an *n*-qubit state $|\psi_i\rangle$ close to the desired *n*-bit string of each MUB.

Example: For batch $\mathcal{B}_i = [[00]\,[00]\,[11]\,[11]\,[11]]$, Alice finds a 2q state where $Z\otimes I,\ I\otimes Z,\ X\otimes I,\ I\otimes X$ measurements often yield **+1**, and the rest often yield **-1**.

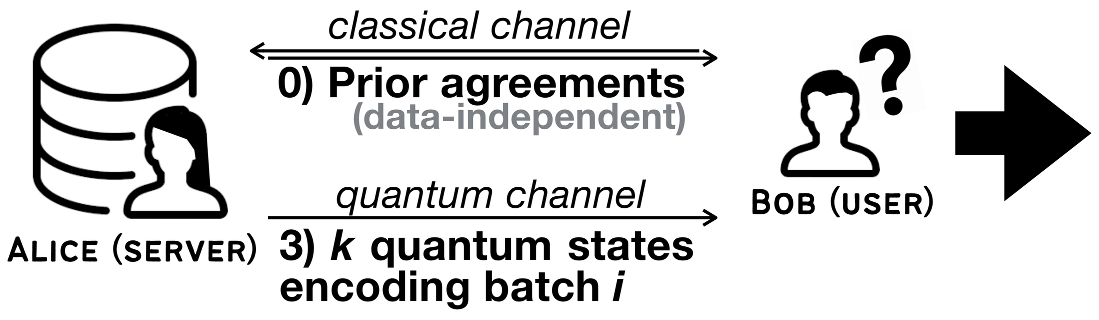
3) Alice sends exactly k copies of state $|\psi_i\rangle$ to Bob.

4) To query *batch* row *r*, Bob measures all copies in basis *r*. The query result is the most common *n*-bit string. He can also choose to query more rows, at the cost of having less samples for each basis, reducing fidelity.

5) Repeat steps 2—4 for each batch, choosing the same row(s). Bob merges the results to reconstruct the selected database rows.

Data privacy metric

Communication scheme

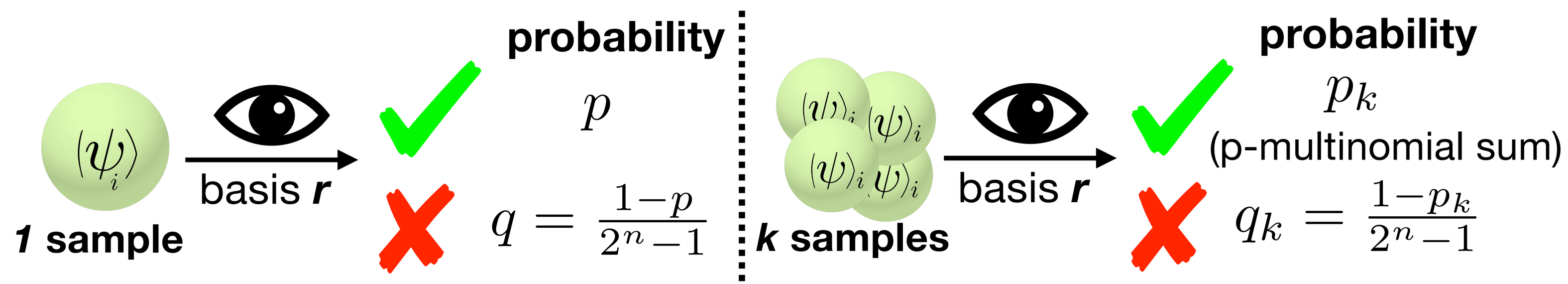


- **User privacy is guaranteed.**
- **Data privacy depends on *k*.**

How can we quantify data privacy?

Data privacy measures how easy it is for Bob to retrieve two rows instead of one. For instance, for ***k=6***, we have

and we need to contrast success probabilities, which depend on the probability ***p*** that a single state projects into the correct ***n***-bit string. Assuming uniformity between states ***i***, rows ***r*** and wrong ***n***-bit strings*, we have



* These uniformities can be enforced by Alice's encoding.

Data privacy metric: $\mathcal{P} = \frac{(p_k)^B}{2(p_{k/2})^B}$ Ratio between expected number of retrieved rows. $\mathcal{P}>1$ is desirable.

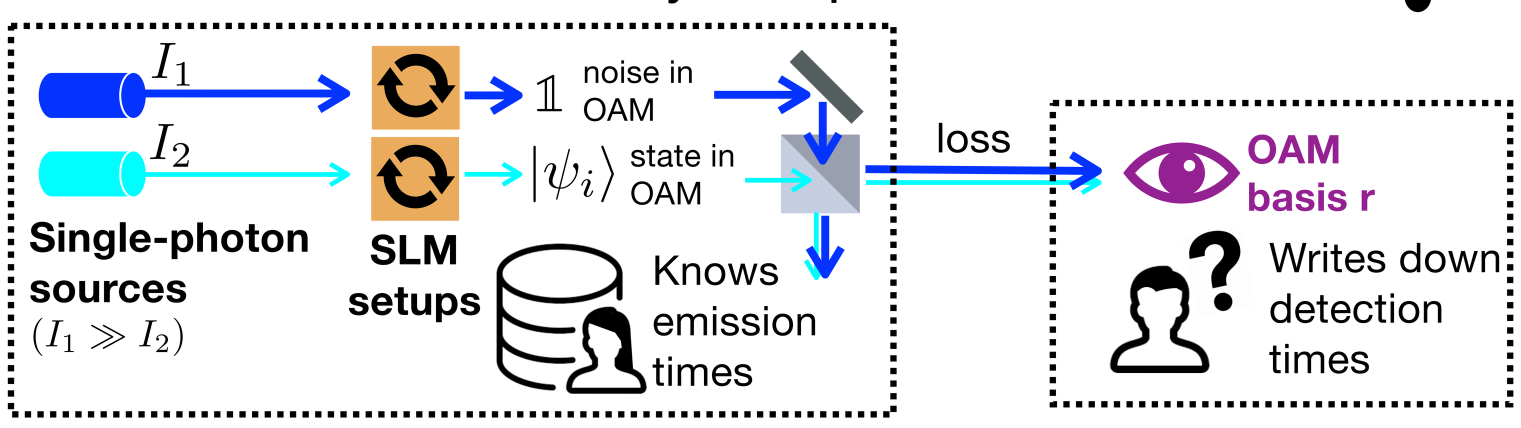
Simulation $(R{=}5,\ n{=}2,\ B{=}65,\ p{=}0.5424^*)$

k	$(p_k)^B$	$(p_{k/2})^B\star$	$\mathcal{P}$	k	$(p_k)^B$	$(p_{k/2})^B\star$	$\mathcal{P}$
1	$\sim 10^{-18}$	$\sim 10^{-18}$	1	41	0.8927	0.1524	2.93
11	0.0006	$\sim 10^{-8}$	$\sim 10^4$	51	0.9709	0.3904	1.24
21	0.1724	0.0004	229.5	61	0.9923	0.6227	0.80
31	0.6429	0.0222	14.49	71	0.9980	0.7869	0.63

* Known average success probability for ***n=2*** Quantum Random Access Codes.
★ Estimated from the average of sample sizes ***(k+1)/2*** and ***(k-1)/2***.

Quantum Channel (WIP)

How can we send exactly ***k*** copies of a state?



1° Bob shares all of his detection times (mostly noise).
2° Alice shares ***k*** of those times which correspond to correct states.
3° Bob postselects the corresponding ***k*** measurement results.

Conclusion

- We propose a protocol to allow a user to query a database privately without forcing the server to fully reveal the database contents (**user privacy** and **data privacy**).
- To do this, the server encodes the database rows into incompatible-observable statistics of quantum states, and sends a limited number of states. Then, the user privately assigns each sample to one row.
- User privacy is guaranteed. Data privacy is quantified via the expected number of retrieved rows. We highlight quantum sample sizes that maximize data privacy and fidelity on a 5-row simulation.

Quantum Information Processing Conference | RIGA 2026

Giancarlo Gatti, Floris Geerts, and Rihan Hai. "Private Quantum Database." *arXiv preprint arXiv:2508.19055* (2025).