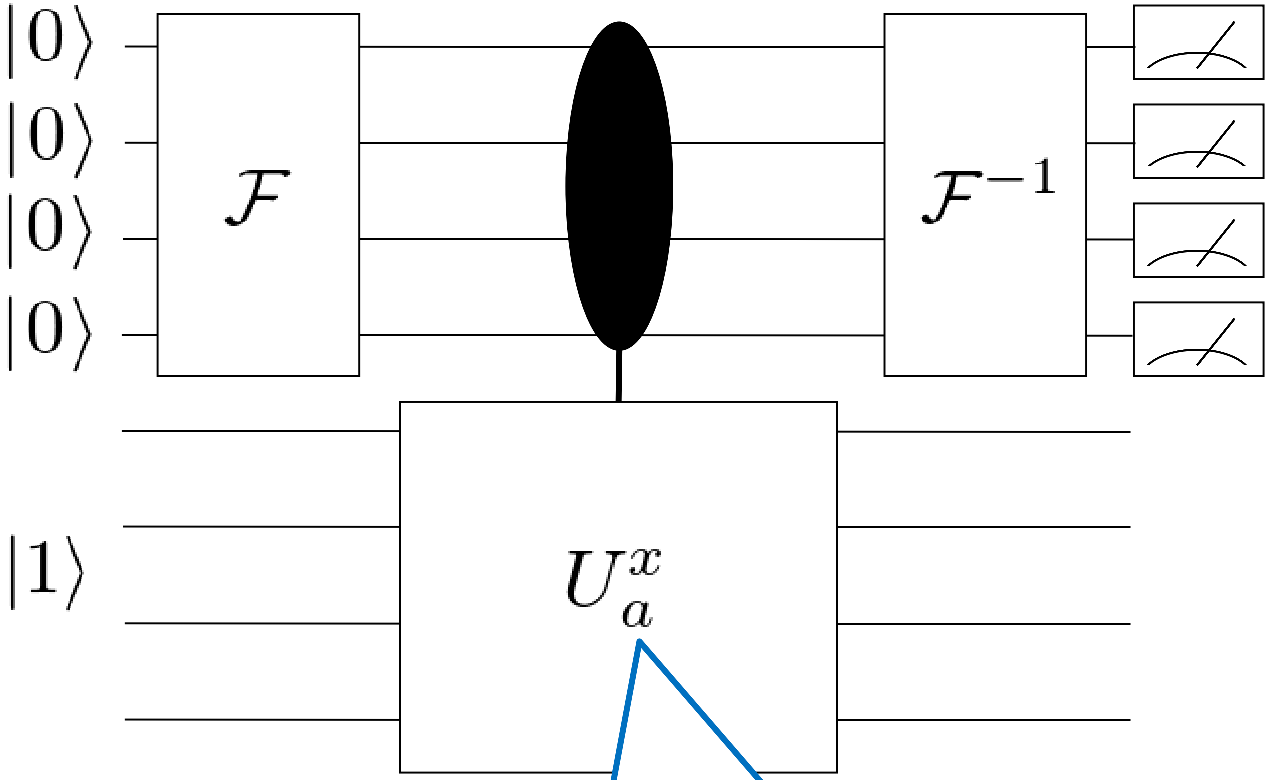
Simulating Quantum Circuits with Sparse Output Distributions

Martin Schwarz[1], Maarten Van den Nest[2]

[1]Vienna Center for Quantum Science and Technology,
Faculty of Physics, University of Vienna, A-1090 Wien, Austria
[2]Max Planck Institut für Quantenoptik,
Hans-Kopfermann-Str. 1, D-85748 Garching, Germany

Quant. Inf. Comp. (*accepted*), arXiv:1310.6749

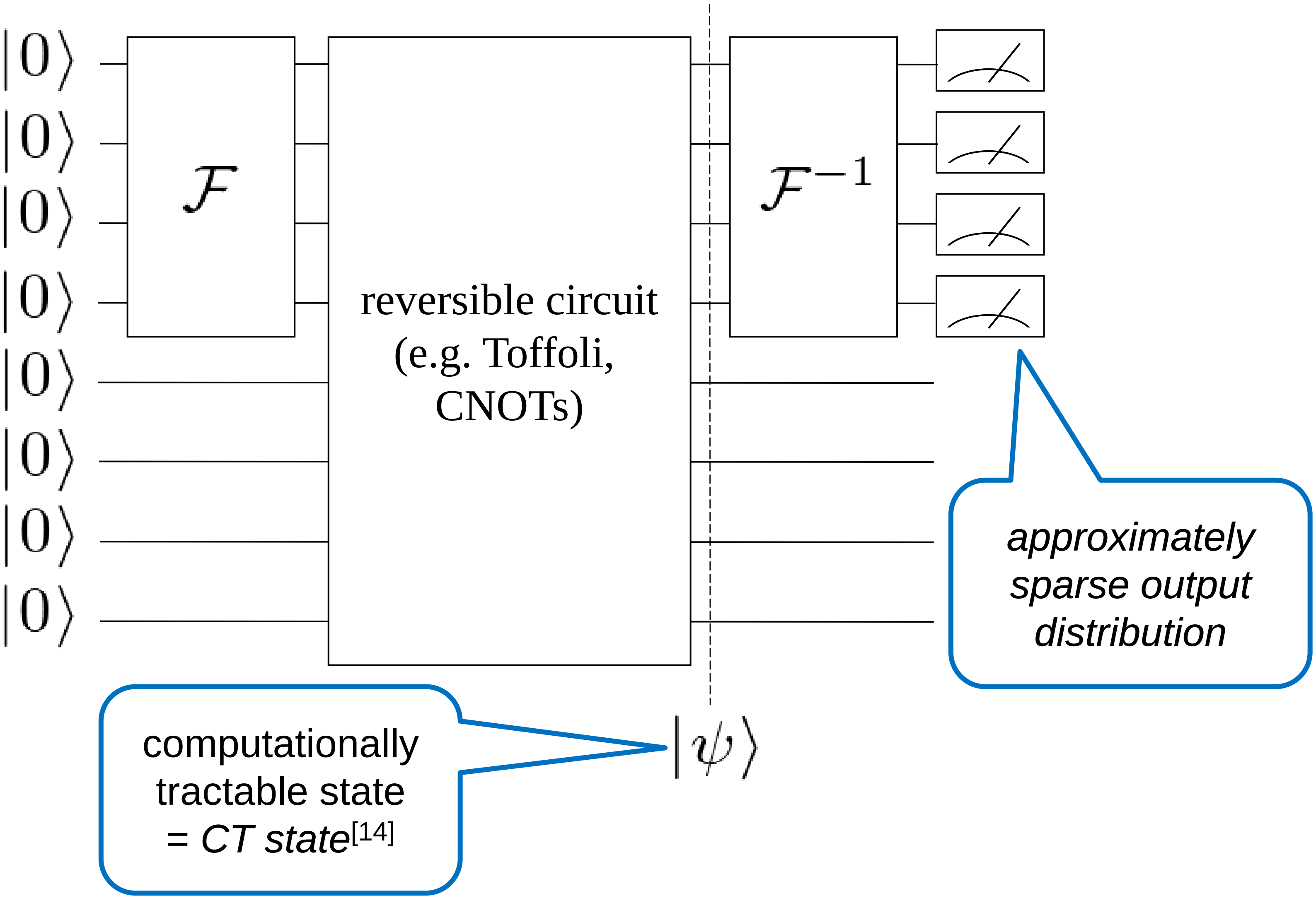
Shor´s algorithm (quantum part)



modular exponentiation

$$\text{c-}U_a^x : |x\rangle|y\rangle \mapsto |x\rangle|ya^x \bmod N\rangle$$

Simulating quantum circuits classically



[14] Van den Nest. Quant. Inf. and Comp., 11(9-10):784–812, 2011.

Computationally Tractable (CT) states

Definition[14]: A state is called *computationally tractable (CT)*, if

(a) $p_x = |\langle x|\psi\rangle|^2$ can be sampled efficiently classically, and if

(b) $\langle x|\psi\rangle$ can be computed efficiently (polynomial in the bit size)

CT states capture two key properties of several important families of simulable quantum states, such as

- MPS with polynomial bond dimension,
- states generated by poly-size Clifford circuits,
- nearest-neighbor matchgate circuits,
- bounded tree-width circuits,
- normalizer circuits over finite Abelian groups (acting on coset states)
- …

[14] Van den Nest. Quant. Inf. and Comp., 11(9-10):784–812, 2011.

Useful lemmas on CT states

CT states have the remarkable property, that overlaps CT states and expectation values of certain operators on CT states can be efficiently computed:

Lemma 12 ([VdN11]). *Let $|\psi\rangle$ and $|\varphi\rangle$ be CT n-qubit states and let A be an efficiently computable basis-preserving n-qubit operation. Then there exists a randomized classical algorithm with runtime poly$(n, 1/\varepsilon, \log\frac{1}{\delta})$ which outputs an approximation of $\langle\psi|A|\varphi\rangle$ with accuracy ε and success probability at least $1-\delta$.*

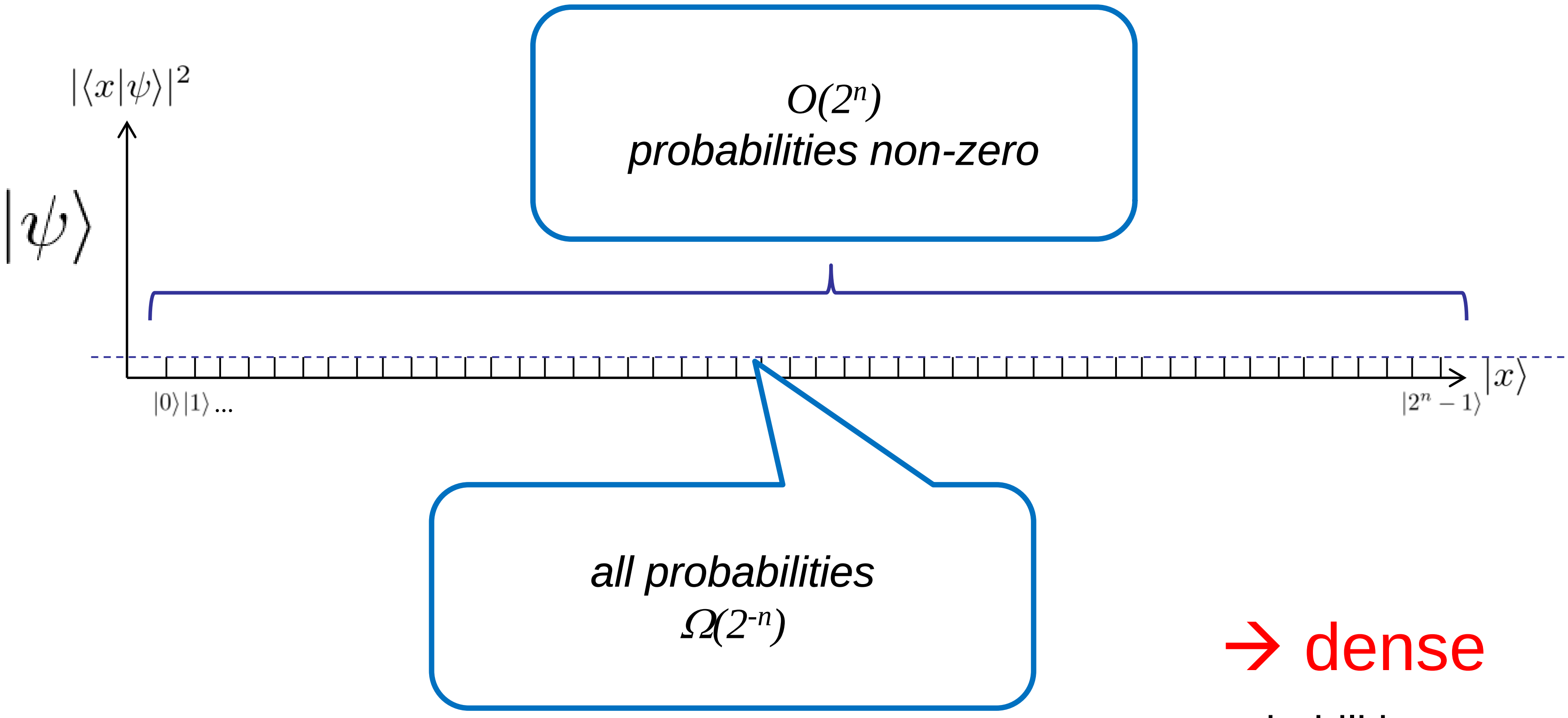
Lemma 13 ([VdN11]). *Let $|\psi\rangle$ and $|\varphi\rangle$ be CT n-qubit states, let $|\xi\rangle$ and $|\chi\rangle$ be CT k-qubit states with $k \leq n$. Then there exists a randomized classical algorithm with runtime poly$(n, 1/\varepsilon, \log\frac{1}{\delta})$ which outputs an approximation of $\langle\varphi|[|\xi\rangle\langle\chi| \otimes \mathbb{1}]|\psi\rangle$ with accuracy ε and success probability at least $1-\delta$.*

Proof: by a sampling argument using a complex-valued Chernoff-bound

Note, that this is exponentially more accurate than estimating the overlap of two explicitly given general state vectors by sampling.

[VdN11] Van den Nest. Quant. Inf. and Comp., 11(9-10):784–812, 2011.

Approximate sparseness

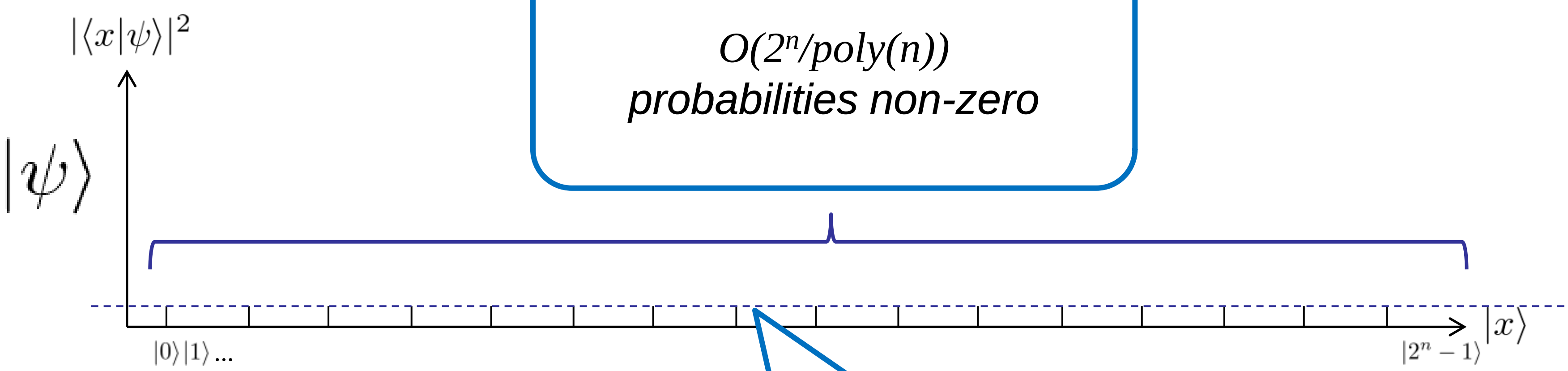


→ dense

probabilities too small to be estimated

Approximate sparseness

$O(2^n/poly(n))$
probabilities non-zero



all probabilities
$\Omega(poly(n)/2^n)$

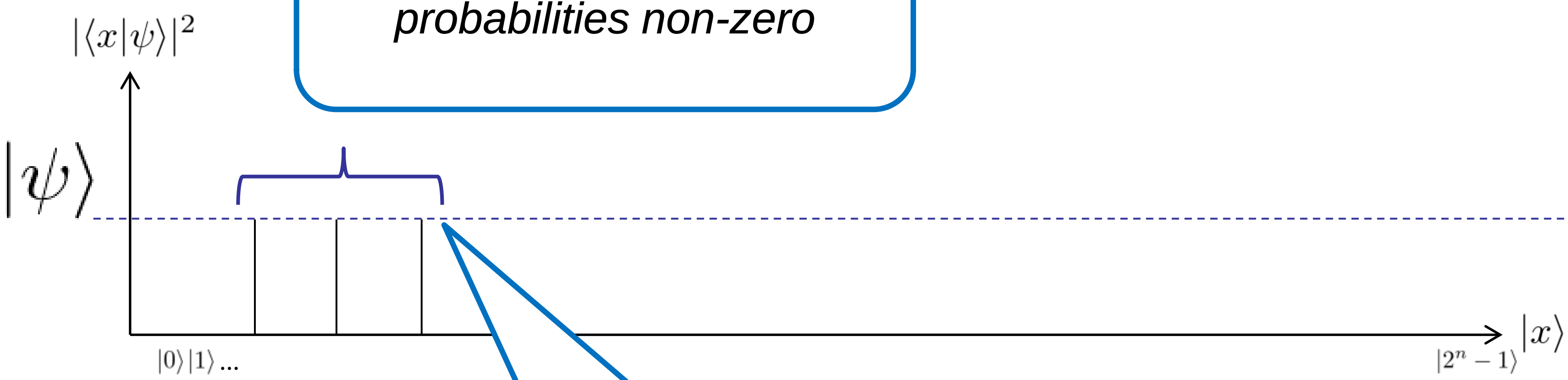
→ dense

non-zero elements
cannot be identified

Shor's algorithm: $\Omega(N/log(N))$ amplitudes non-zero.

Approximate sparseness

O(poly(n))
probabilities non-zero

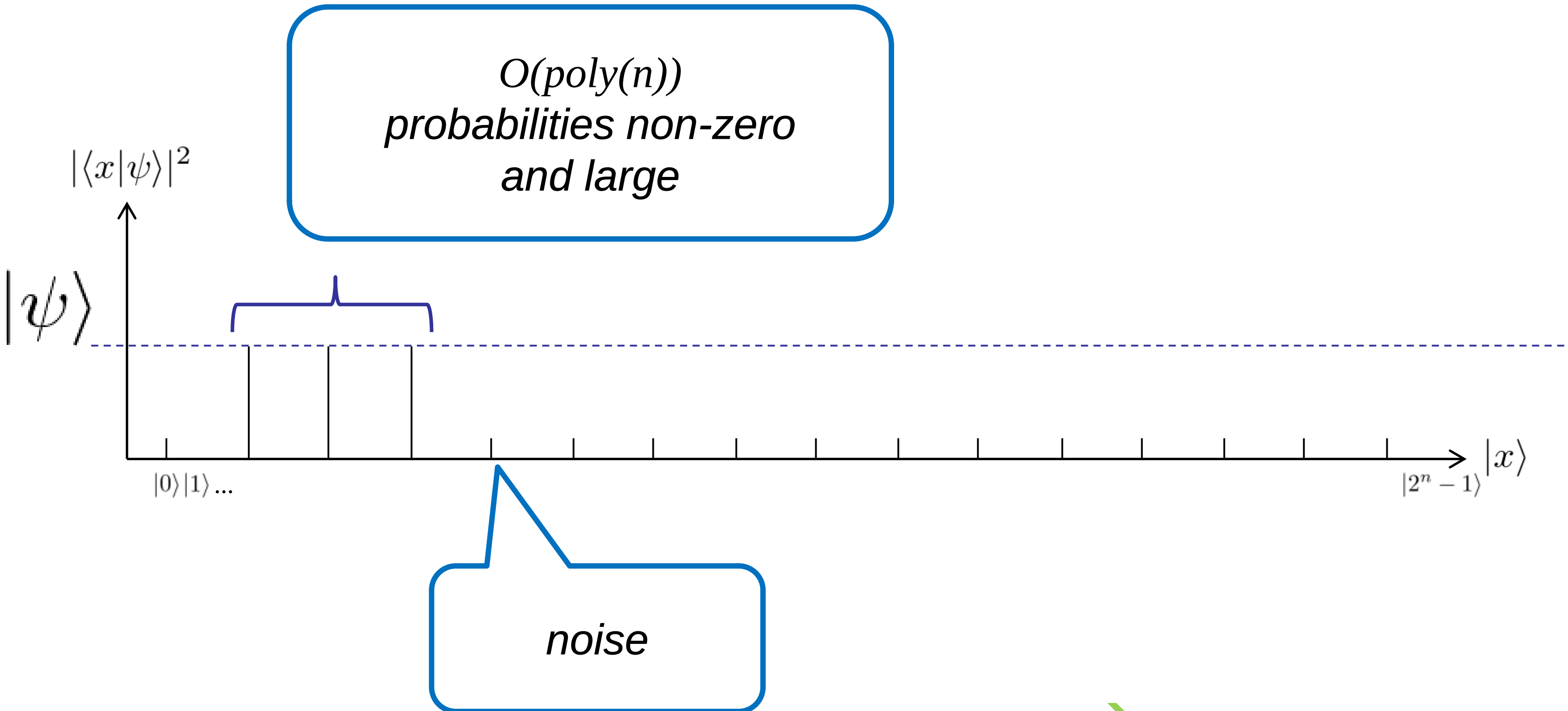


all non-zero probabilities
Ω(1/poly(n))

→ sparse!

elements can be identified!
probabilities can be estimated!

Approximate sparseness



→ approx. sparse

still works with noise

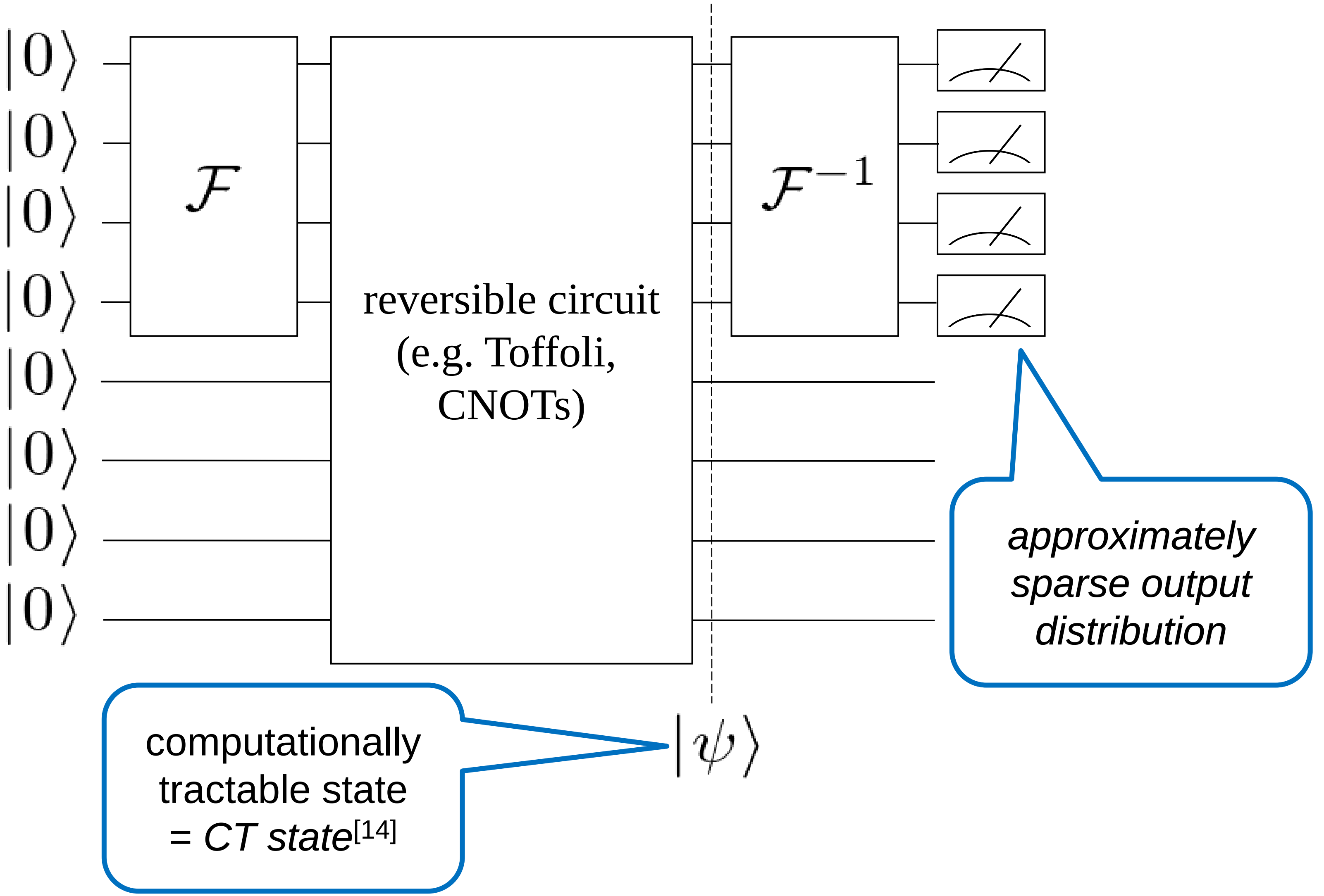
Main result

Theorem. *Consider a unitary n-qubit quantum circuit composed of two blocks $C = U_2 U_1$ with input state $|\psi_{in}\rangle$. Suppose that the following conditions are fulfilled:*
(a) the state $U_1|\psi_{in}\rangle$ obtained after applying the first block is CT,
(b1) the second block U_2 is the QFT modulo 2^n or its inverse, or
(b2) the second block U_2 is a tensor product of unitaries $u_1 \otimes \cdots \otimes u_n$
(c) the final state $|\psi_{out}\rangle = C|\psi_{in}\rangle$ is promised to be $\sqrt{\varepsilon}$-approximately t-sparse for some $\varepsilon \leq 1/6$ and some t.

Then there exists a randomized classical algorithm with runtime $poly(n, t, 1/\varepsilon, \log \frac{1}{\delta})$ which outputs (by means of listing all nonzero amplitudes) an s-sparse state $|\psi\rangle$ which, with probability at least $1 - \delta$, is $O(\sqrt{\varepsilon})$-close to $|\psi_{out}\rangle$, where $s = O(t/\varepsilon)$.

(Theorem is stated for case of amplitudes and 2-norm.
Analogous theorem is true for probabilities and 1-norm.)

Simulating quantum circuits classically



[14] Van den Nest. Quant. Inf. and Comp., 11(9-10):784–812, 2011.

Proof sketch (QFT case)

universität
wien

Lemma (coefficients estimation): Given an approx.-sparse probability distribution over n bits, where *all marginal distributions* of the first m bits are efficiently sampleable. Then w.h.p. a list of the $O(poly(n))$-many bit strings with non-zero probability can be efficently computed and the probabilities can be efficiently estimated to $O(poly(1/\varepsilon))$ accuracy.

Proof: by a binary search / branch-and-bound argument.

The main theorem follows from the coefficient estimation lemma and the next lemma.

Main theorem is a generalization of the **Kushilevitz-Mansour** algorithm or **Goldreich-Levin** theorem to quantum states.

Proof sketch (QFT case)

Lemma (marginal distribution): The *m*-bit marginals of the probability distribution produced by the quantum circuit satisfying the assumptions of the main theorem are efficiently approximable.

Proof sketch:

Generalized Pauli operators:

$$X_d|x\rangle = |x+1\rangle \qquad \mathcal{F}_d^\dagger Z_d \mathcal{F}_d = X_d$$

$$Z_d|x\rangle = e^{\frac{2\pi i}{d}x}|x\rangle \qquad \mathcal{F}_d Z_d \mathcal{F}_d^\dagger = X_d^\dagger$$

We want to estimate $|y_1 \cdots y_m\rangle\langle y_1 \cdots y_m| \otimes I \equiv P(y)$ on a CT state.

Note: $\hat{x} \mod 2^m = \hat{y} \quad \text{iff} \quad \alpha^{\hat{y}} Z^{2^{k-m}} |\hat{x}\rangle = |\hat{x}\rangle \quad \text{with } \alpha := e^{-\frac{2\pi i}{2^m}}.$

Therefore, *P(y)* is the *1*-eigenspace of $M := \alpha^{\hat{y}} Z^{2^{k-m}}$, which can be obtained by the average:

$$P(y) = \frac{1}{2^m} \sum_{u=0}^{2^m-1} M^u.$$

Proof sketch (QFT case)

Thus the marginal probability distribution of the first m qubits of the quantum circuit can be written as

$$p(y) = \langle \mathrm{CT} | [\mathcal{F}^\dagger P(y) \mathcal{F}] \otimes I | \mathrm{CT} \rangle$$

Using $\mathcal{F}^\dagger P(y) \mathcal{F} = \frac{1}{2^m} \sum_{u=0}^{2^m - 1} N^u$ where $N := \alpha^{\hat{y}} X^{2^{k-m}}$ we find that

$$p(y_1 \cdots y_m) = \frac{1}{2^m} \sum_{u=0}^{2^m - 1} \langle \mathrm{CT} | N^u \otimes I | \mathrm{CT} \rangle$$

But each term of the sum is additively approximable by the CT state lemma, thus the sum is additively approximable as well and the lemma follows.

Consequences

- The dense output distribution of Shor‘s algorithm (or its generalizations) is a ***necessary feature*** for the (conjectured) exponential speed-up over classical computers.

- In order to extract meaningful information out of a dense superposition, additional structure (e.g. group structure) must be present, such that $O(poly(n))$ samples suffice to efficiently identify the structure.

Thank you.