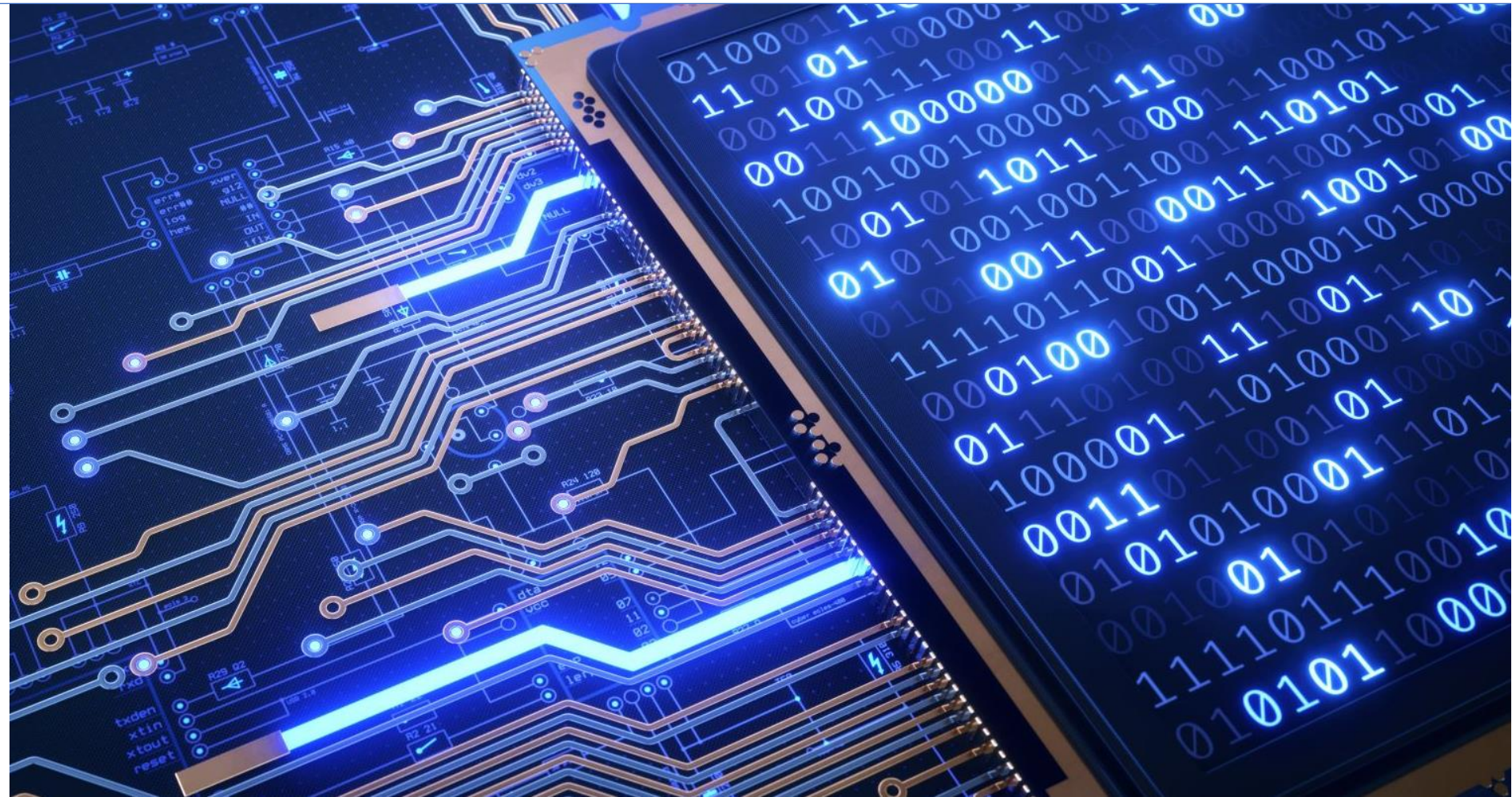
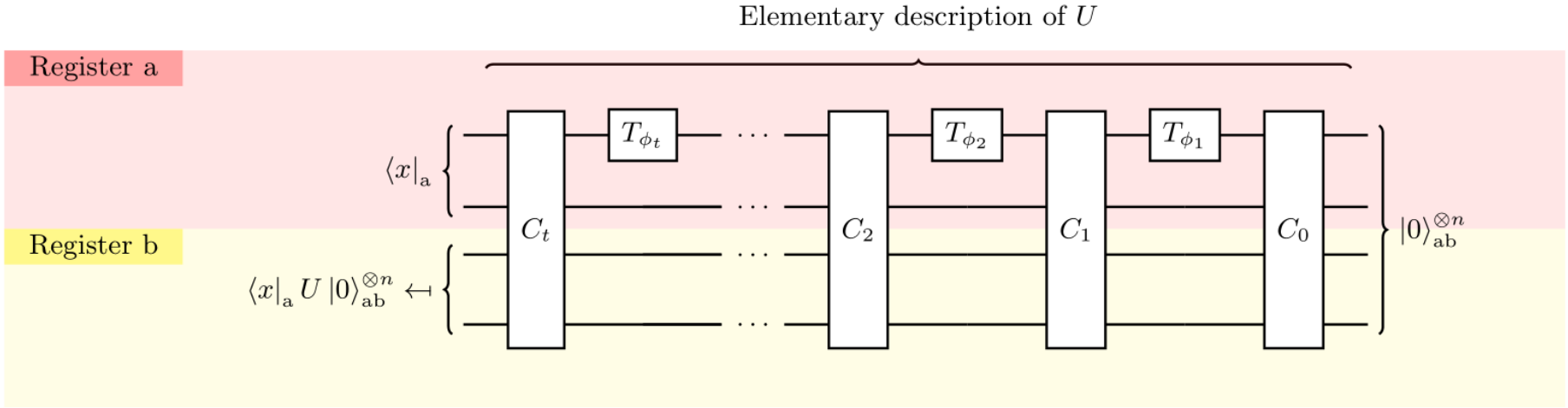
Fast estimation of outcome probabilities for quantum circuits

Hakop Pashayan, Oliver Reardon-Smith, Kamil Korzekwa, Stephen Bartlett



The task: Born rule probability estimation

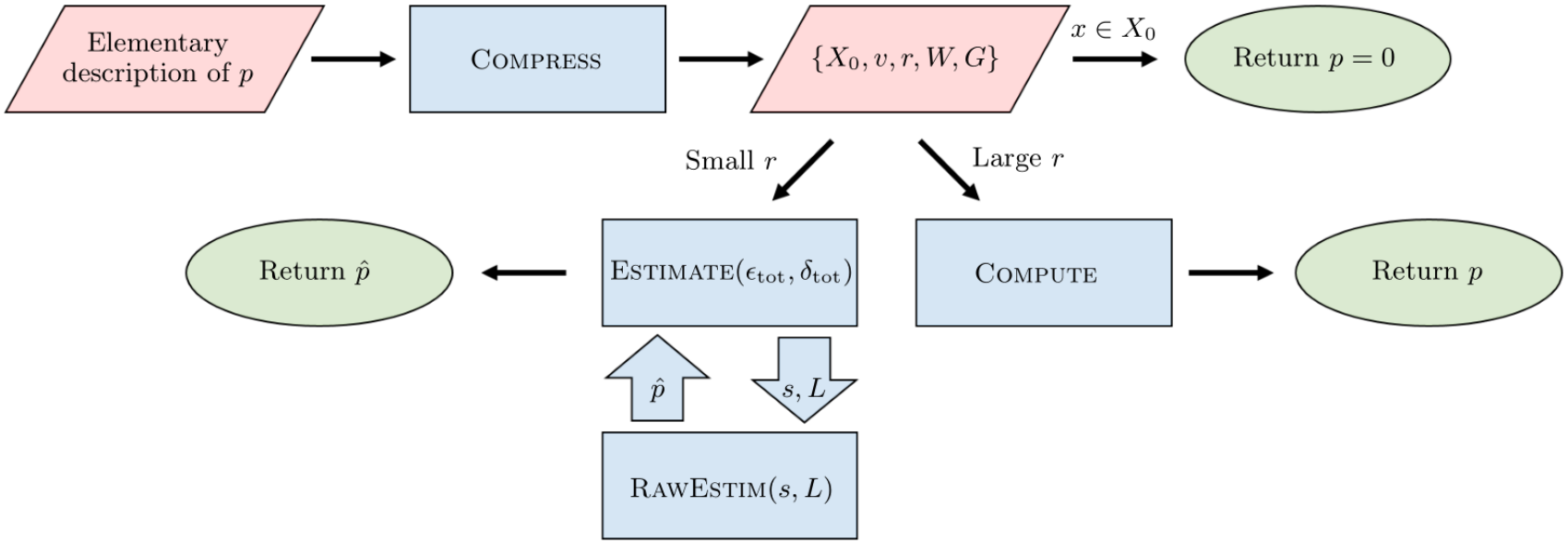


- Allowed gates: Cliffords + T_ϕ where $\phi \in (0, \pi/2)$ can be arbitrary.

$$T_\phi \equiv \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}$$

- For a given $x \in \{0,1\}^w$, project first w qubits in computational basis state $|x\rangle$.
- Want **classical** estimate of $p := \left\| \langle x|_{\mathrm{a}} U |0\rangle_{\mathrm{ab}}^{\otimes n} \right\|_2^2$.

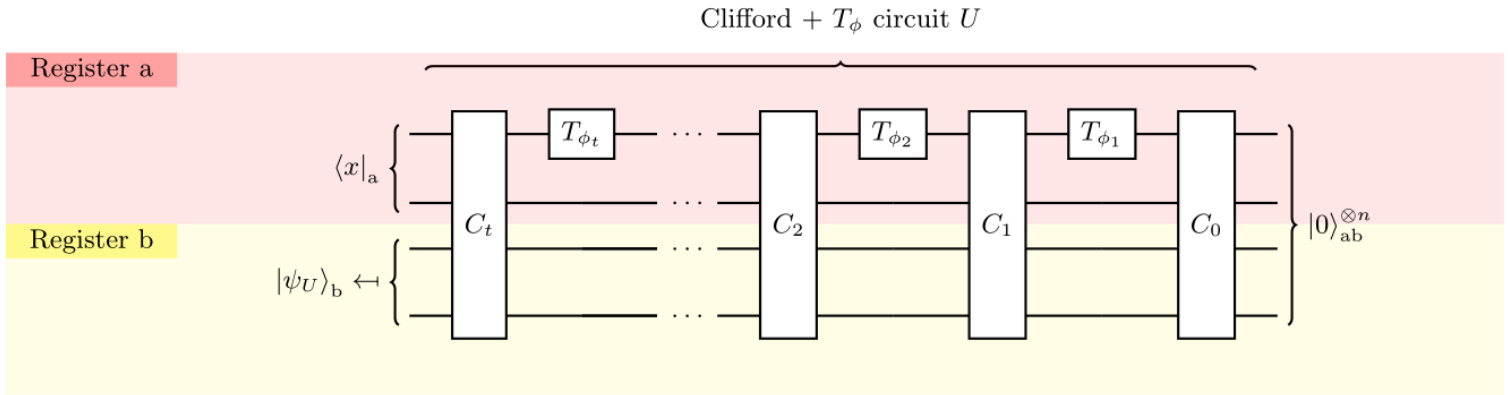
Overview

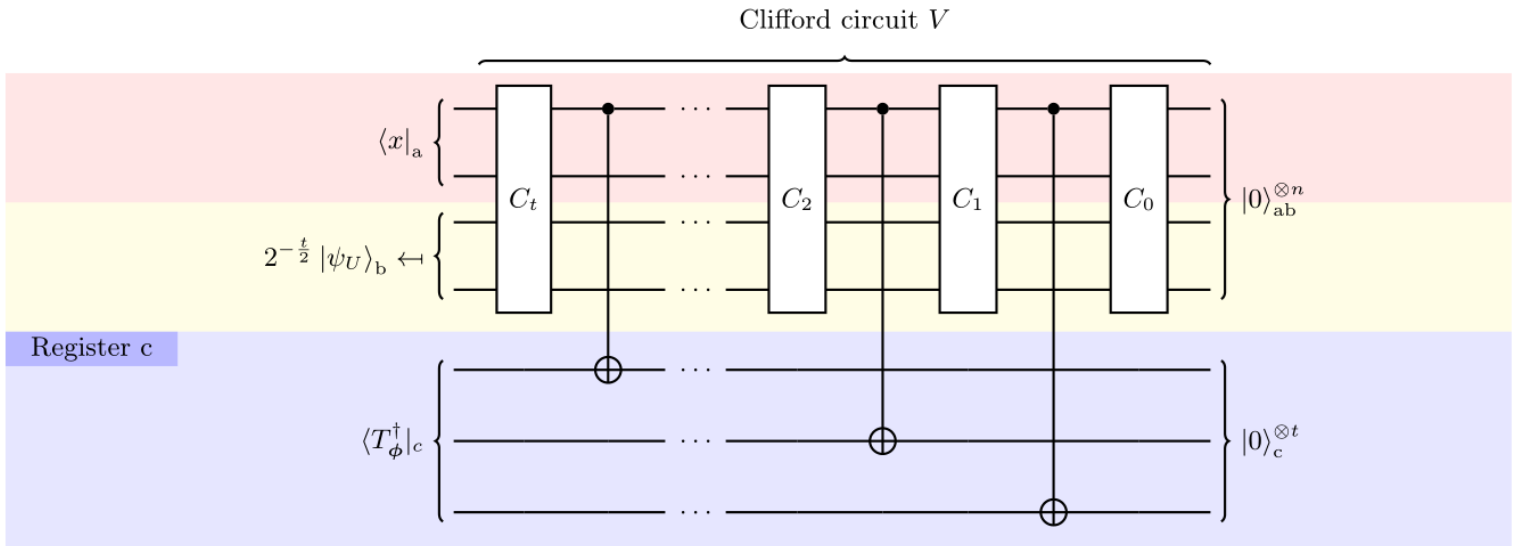


Code for algorithms availible at: https://github.com/or1426/Clifford-T-estimator

1. Compress algorithm
2. Compute algorithm
3. RawEstim algorithm
4. Estimate algorithm

Compress algorithm:





$$
\begin{aligned}
p &= \| |\psi_U\rangle \|_2^2 \\
&= 2^t \mathrm{Tr}\left(\Pi_{G''} |x\rangle\langle x|_{\mathrm{a}} |T_{\boldsymbol{\phi}}^\dagger\rangle\langle T_{\boldsymbol{\phi}}^\dagger|_{\mathrm{c}}\right) \\
&= 2^{t-r+v-w} \mathrm{Tr}\left(\Pi_G |T_{\boldsymbol{\phi}}^\dagger\rangle\langle T_{\boldsymbol{\phi}}^\dagger|\right) \\
&= 2^{t-r+v-w} \left\| \langle 0|^{\otimes t-r} W |T_{\boldsymbol{\phi}}^\dagger\rangle \right\|_2^2
\end{aligned}
$$

$$
|T_{\boldsymbol{\phi}}^\dagger\rangle := |T_{\phi_1}^\dagger\rangle \otimes \ldots \otimes |T_{\phi_t}^\dagger\rangle, \qquad |T_\phi^\dagger\rangle := \frac{1}{\sqrt{2}}(|0\rangle + \exp(-i\phi)\,|1\rangle)
$$

COMPRESS algorithm

Given an 'elementary description of p', COMPRESS **efficiently** outputs:

- $v \in \{0, 1, \ldots, w\}$
- A subset $\{q_1, \ldots, q_v\} \subseteq [w]$ (locations of deterministic outcome qubits) and specifes the outcome.
- Stabilizer generating set $G \in \mathcal{G}(t, t-r)$ $\quad p = 2^{t-r+v-w} \mathrm{Tr}\left(\Pi_G |T_{\boldsymbol{\phi}}^\dagger\rangle\langle T_{\boldsymbol{\phi}}^\dagger|\right)$
- $r \in \{0, 1, \ldots, \min\{t, n-w\}\}$
- Elementary description of Clifford circuit W $\quad p = 2^{t-r+v-w} \left\| \langle 0|^{\otimes t-r} W |T_{\boldsymbol{\phi}}^\dagger\rangle \right\|_2^2$

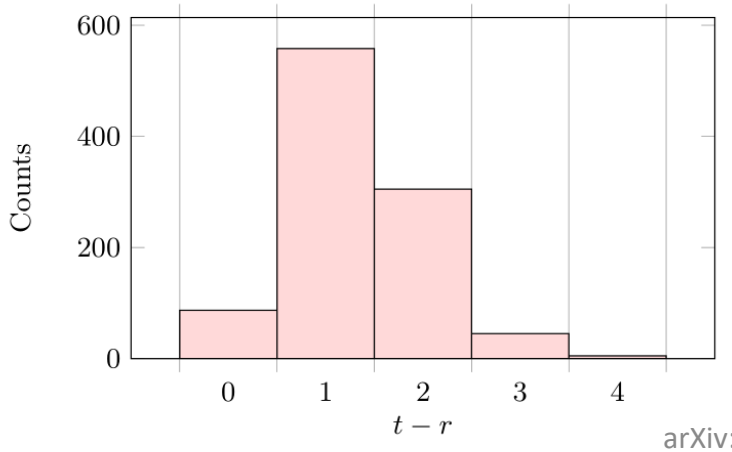
COMPUTE algorithm

- Exactly computes p in time $\tau_{\text{COMPUTE}} = O(2^{t-r}(t-r)t)$.
- Using $G = \{g_1, \ldots, g_{t-r}\}$ from COMPRESS:

$$p = 2^{t-r+v-w}\text{Tr}\left(\Pi_G|T_{\boldsymbol{\phi}}^\dagger\rangle\langle T_{\boldsymbol{\phi}}^\dagger|\right) = 2^{v-w}\langle T_{\boldsymbol{\phi}}^\dagger|\prod_{i=1}^{t-r}(I+g_i)|T_{\boldsymbol{\phi}}^\dagger\rangle$$

- $r \in \{0, 1, \ldots, r_{\max}\}$ where $r_{\max} = \min\{t, n-w\}$.
- r concentrates near $r_{\max}$ for high Clifford count random circuits.

10^3 RCs
$n = 100$
$w = 20$
$t = 80$
$c = 10^5$



$n = 55$
$w = 5$
$t = 80$
$c = 10^5$
runtime $\leq$ 2 hrs

arXiv: 2101.12223

RawEstim algorithm: Concentration

$$p = \left\| 2^{\frac{t-r+v-w}{2}} \langle 0|^{\otimes t-r} W |T_{\boldsymbol{\phi}}^{\dagger}\rangle \right\|_2^2 = \| |\mu\rangle \|_2^2$$

- For $y \in \{0,1\}^t$, we define unnormalized stabilizer states $|\psi(y)\rangle$ and product probability distribution $q(y)$ s.t.

$$\sum_y q(y) \, |\psi(y)\rangle = |\mu\rangle \, .$$

- Concentration: The s-sample average $\left|\overline{\psi}\right\rangle = \frac{1}{s}\sum_{i=1}^{s} |\psi_i\rangle$ concentrates around $|\mu\rangle$:

$$\Pr\left(\left\| \left|\overline{\psi}\right\rangle\!\left\langle\overline{\psi}\right| - |\mu\rangle\langle\mu| \right\|_1 \geq \epsilon \right) \leq 2e^2 \exp\left(\frac{-s(\sqrt{p+\epsilon} - \sqrt{p})^2}{2(\sqrt{m}+\sqrt{p})^2} \right),$$

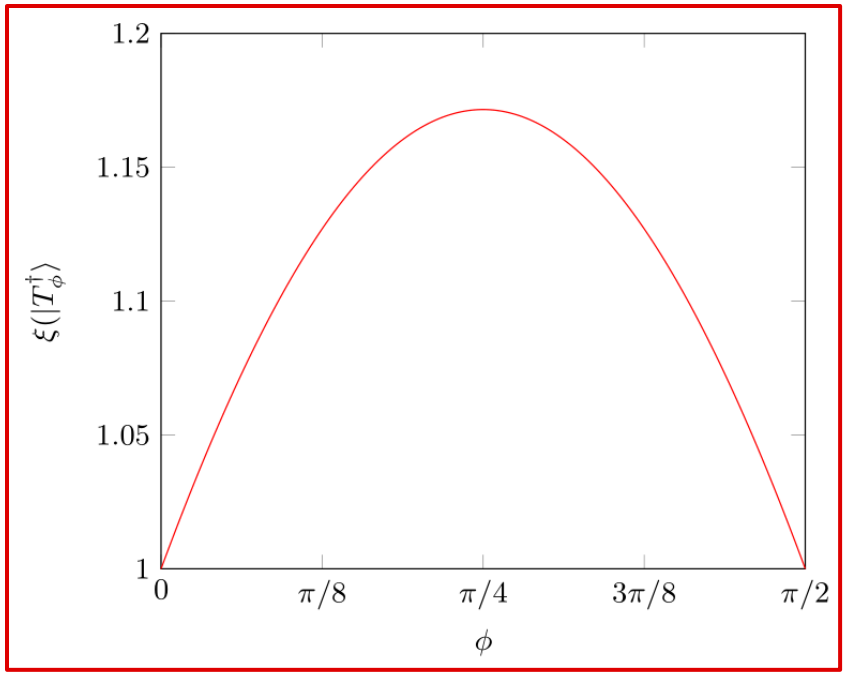
  where $p = \| |\mu\rangle \|_2^2$ and $m \geq \| |\psi(y)\rangle \|_2^2$.

- We show that $\xi^* \geq \| |\psi(y)\rangle \|_2^2$.

T. P. Hayes, A large-deviation inequality for vector-valued martingales, Combinatorics, Probability and Computing (2005).

RawEstim algorithm: FNE and CH-form

- $\hat{p}$ is estimate of $\left\| \left| \overline{\psi} \right\rangle \right\|_2^2$ using L iterations of the BG fast norm estimation algorithm.
- CH-form [BBCCGH] + tricks used to compute $|\psi(y)\rangle \propto \langle 0|^{\otimes t-r} W |\tilde{y}\rangle$
  - Instead of $|\tilde{y}\rangle \mapsto W |\tilde{y}\rangle$ we use $W \left|\tilde{0}^t\right\rangle \mapsto W |\tilde{y}\rangle$
  - Factorization of CH-form: $CHF(|0\rangle \otimes |\varphi\rangle) \mapsto CHF(|\varphi\rangle)$
- Stabilizer extent: $\xi^* := \xi(|T_{\phi_1}^\dagger\rangle) \times \ldots \times \xi(|T_{\phi_t}^\dagger\rangle) < 1.2^t$.



S. Bravyi and D. Gosset, Improved classical simulation of quantum circuits dominated by Clifford gates, Phys. Rev. Lett. **116**, 250501 (2016).

S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, Quantum **3**, 181 (2019).

arXiv: 2101.12223

RawEstim algorithm: performance

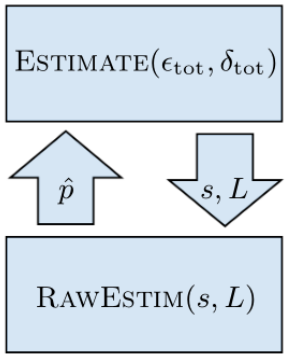
- For given $s, L \in \mathbb{N}^+$ computes $\hat{p}$ in time $\tau_{\text{RawEstim}} = O(st^2(t-r) + sLr^3)$ s.t. $\forall \epsilon_{\text{tot}} > 0, \epsilon \in (0, \epsilon_{\text{tot}})$:

$$\Pr\left(|\hat{p} - p| \geq \epsilon_{\text{tot}}\right) \leq 2e^2 \exp\left(-s\frac{(\sqrt{p+\epsilon} - \sqrt{p})^2}{2(\sqrt{\xi^*} + \sqrt{p})^2}\right) + \exp\left(-L\left(\frac{\epsilon_{\text{tot}} - \epsilon}{p+\epsilon}\right)^2\right)$$
$$=: \delta(p, \epsilon_{\text{tot}}, \epsilon, s, L).$$

- $\delta(p, \epsilon_{\text{tot}}, \epsilon, s, L)$ depends on unknown p
- Can UB $\delta(p, \epsilon_{\text{tot}}, \epsilon, s, L)$ by setting $p = 1$ but this is very conservative.

ESTIMATE algorithm

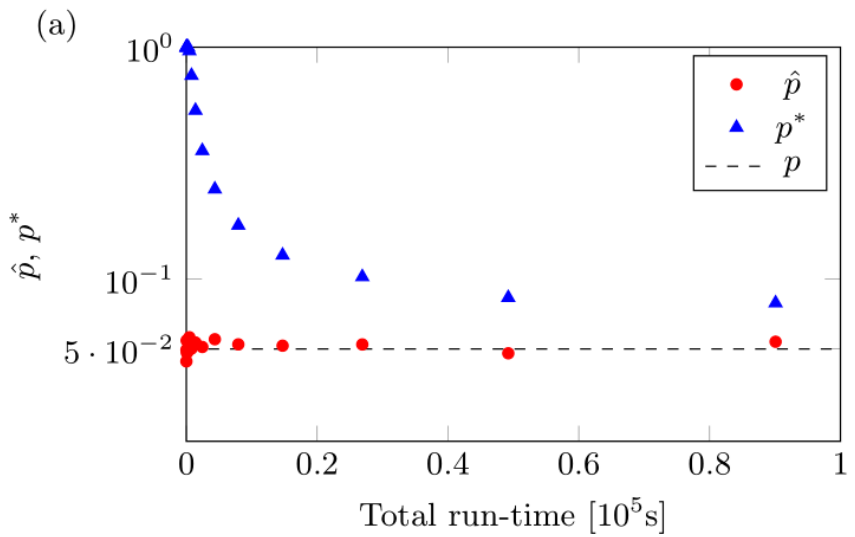
- ESTIMATE purpose: $\Pr\left(|\hat{p} - p| \geq \epsilon_{\text{tot}}\right) \leq \delta_{\text{tot}}$
- ESTIMATE will use RAWESTIM as a sub-routine

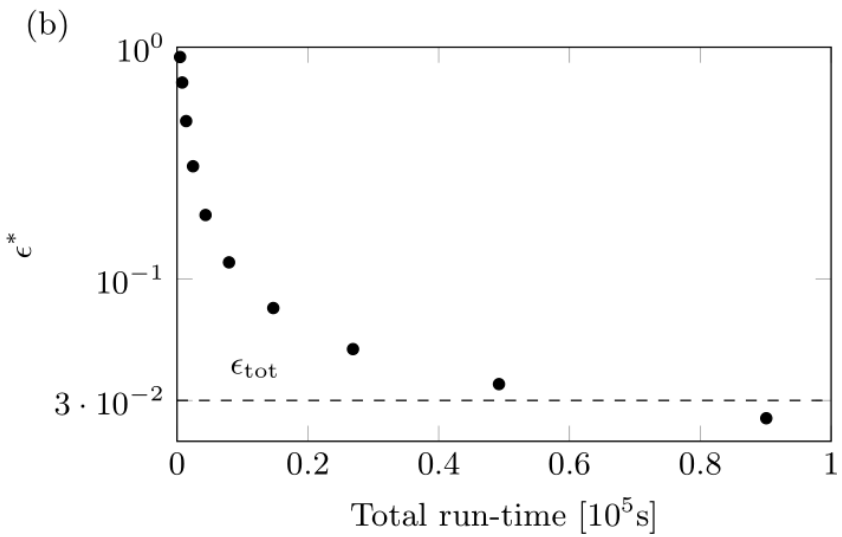


- RAWESTIM gives: $\Pr\left(|\hat{p} - p| \geq \epsilon_{\text{tot}}\right) \leq \delta(p, \epsilon_{\text{tot}}, \epsilon, s, L)$
- Defn: $\epsilon^*(p, \delta_{\text{tot}}, \mathcal{T})$ is the minimal achievable additive error s.t. $\tau_{\text{RAWESTIM}}(s, L) \leq \mathcal{T}$ and FP $\leq \delta_{\text{tot}}$
- $\Pr\left(|\hat{p} - p| \geq \epsilon^*(p, \delta_{\text{tot}}, \mathcal{T})\right) \leq \delta_{\text{tot}}$
- $\epsilon^*(p, \delta_{\text{tot}}, \mathcal{T})$ is monotone increasing in p

ESTIMATE algorithm

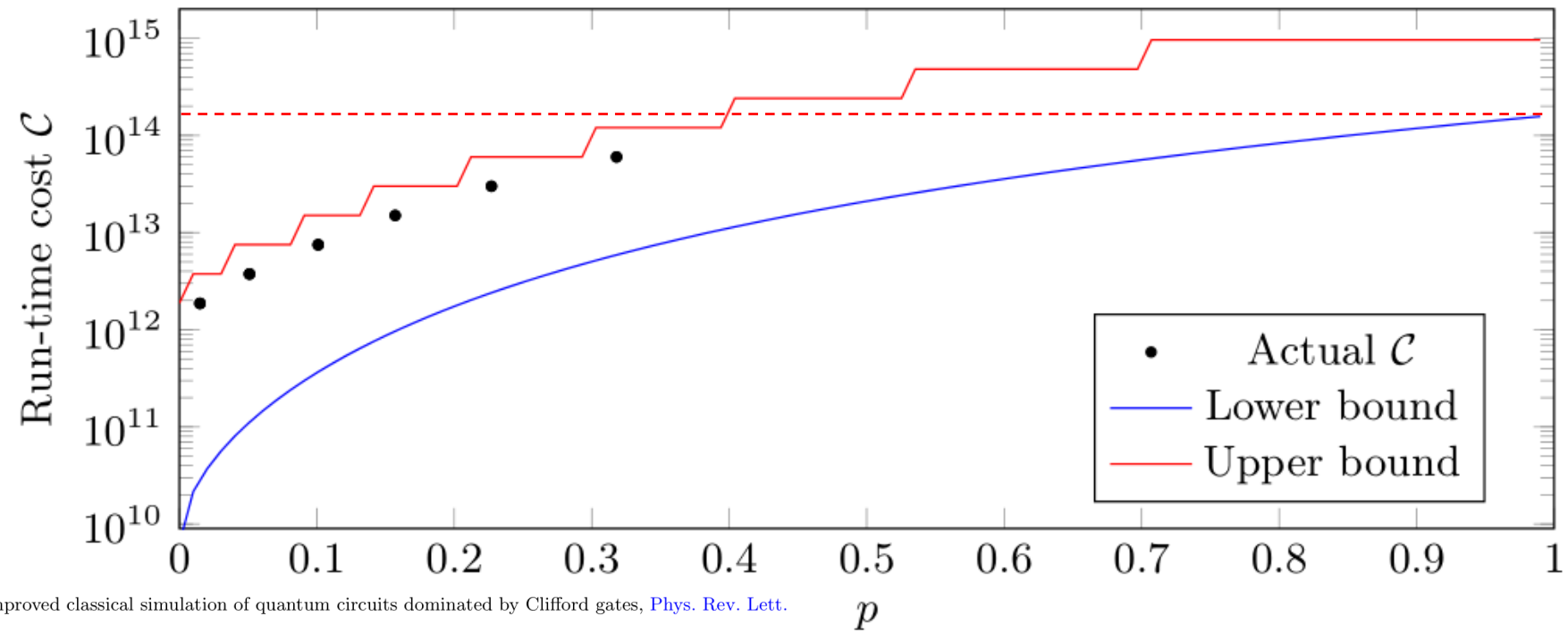
- Idea: for $k = 1, 2, \ldots,$ compute $\epsilon^*(p, \delta_{\text{tot}}, 2^k \mathcal{T}_0)$ until we find $\epsilon^*(p, \delta_{\text{tot}}, 2^k \mathcal{T}_0) \leq \epsilon_{\text{tot}}$
- Idea: for $k = 1, 2, \ldots,$ compute $\epsilon_k^* := \epsilon^*(p_{k-1}^*, \delta_k, 2^k \mathcal{T}_0)$ until we find $\epsilon^*(p^*, \delta_k, 2^k \mathcal{T}_0) \leq \epsilon_{\text{tot}}$
- $p_0^* = 1$, $p_k^* = \hat{p}_k + \epsilon_k^*$ and
- $\delta_k := \frac{6}{\pi^2 k^2} \delta_{\text{tot}}$ ensures $\sum_k \delta_k = \delta_{\text{tot}}$





ESTIMATE algorithm

- ESTIMATE gives: $\Pr\left(|\hat{p} - p| \geq \epsilon_{\text{tot}}\right) \leq \delta_{\text{tot}}$
- Runtime depends on p and is probabilistic.
- RUNTIME gives a probabilistic upper bound of τ_{ESTIMATE} for any possible p.



[1] S. Bravyi and D. Gosset, Improved classical simulation of quantum circuits dominated by Clifford gates, Phys. Rev. Lett. **116**, 250501 (2016).

[2] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, Quantum **3**, 181 (2019).

Outlook

- Can we predict properties of COMPRESS output?
- Practical applications for COMPUTE?
- Can we apply some of these techniqes to improve sampling algorithms?
- Can we adapt ESTIMATE to the mixed state formalism?

Thank you!