Partial randomization for better approximation

joint work with
1. Angus Lowe & Aram Harrow
2. Jakob Günther, Alexander Schmidhuber, Marek Miller, Matthias Christand & Aram Harrow

Plan

1. Motivation: when randomness helps
2. Truncating quantum states
3. Hamiltonian simulation

1 Motivation: when randomness helps

General principle for quantum computers:

Coherence is better than randomness

e.g. $|\psi\rangle = \alpha |\psi_{good}\rangle + \beta |\psi_{bad}\rangle$

measure $|\psi_{good}\rangle$ with prob $|\alpha|^2$, better: amplitude amplification

1 Motivation: when randomness helps

General principle for quantum computers:

Coherence is better than randomness

e.g. $|\psi\rangle = \alpha |\psi_{good}\rangle + \beta |\psi_{bad}\rangle$

measure $|\psi_{good}\rangle$ with prob $|\alpha|^2$, better: amplitude amplification

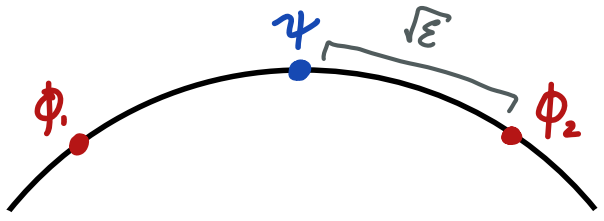
However, often with limited resources, randomness helps!

Examples
- error mitigation
- randomized compilation
- Hamiltonian simulation: qDRIFT

A simple principle

Target state: $|\psi\rangle$

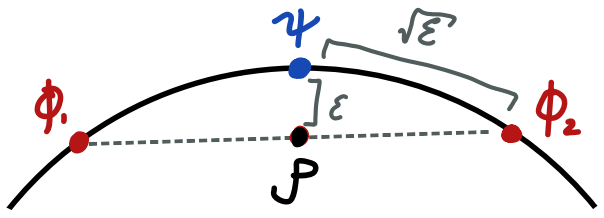
Available states: $|\phi_1\rangle, |\phi_2\rangle$ $\quad |\langle\psi|\phi_i\rangle| \approx 1-\varepsilon$



A simple principle

Target state: $|\psi\rangle$

Available states: $|\phi_1\rangle, |\phi_2\rangle$



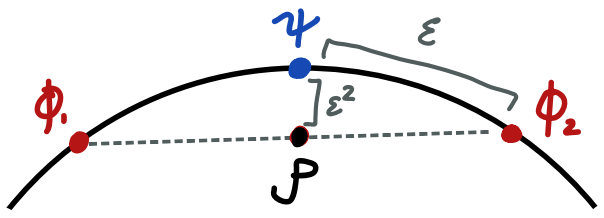
$\rho = \frac{1}{2}|\phi_1\rangle\langle\phi_1| + \frac{1}{2}|\phi_2\rangle\langle\phi_2|$

is $O(\varepsilon)$ close to $|\psi\rangle$

A simple principle

Target state: $|\psi\rangle$

Available states: $|\phi_1\rangle, |\phi_2\rangle$



$\rho = \frac{1}{2}|\phi_1\rangle\langle\phi_1| + \frac{1}{2}|\phi_2\rangle\langle\phi_2|$

is $O(\varepsilon^2)$ close to $|\psi\rangle$

Lesson: mixed state can be better than pure state

The catch: only with respect to trace distance, not fidelity

② Truncating quantum states

$$|\psi\rangle = \sum_{i=1}^{d} \psi_i |i\rangle$$

only k nonzero coefficients

Want: approximate with k-sparse

$$|\psi_{AB}\rangle = \sum_{i=1}^{d} s_i |a_i\rangle|b_i\rangle$$

Want: approximate with Schmidt rank k

What is the best we can do?

② Truncating quantum states

$$|\psi\rangle = \sum_{i=1}^{d} \psi_i |i\rangle$$

only k nonzero coefficients

Want: approximate with k-sparse

$$|\psi_{AB}\rangle = \sum_{i=1}^{d} s_i |a_i\rangle|b_i\rangle$$

Want: approximate with Schmidt rank k

What is the best we can do?

Solution: keep largest k coefficients

② Truncating quantum states

Work with Aram Harrow & Angus Lowe
on the arXiv this week Friday...

$$|\psi\rangle = \sum_{i=1}^{d} \psi_i |i\rangle$$

Want: approximate with k-sparse ← only k nonzero coefficients

Solution: keep largest k coefficients

→ this is the answer for the fidelity

Pure states: $F^2 + T^2 = 1$, $F = \sqrt{1-\varepsilon} \rightarrow T = \sqrt{\varepsilon}$

Fuchs-van Graaf inequalities: $\varepsilon \lesssim T \leq \sqrt{\varepsilon}$

↖ mixed states can get quadratic improvement

② Truncating quantum states

Work with Aram Harrow & Angus Lowe
on the arXiv this week Friday...

$$|\psi\rangle = \sum_{i=1}^{d} \psi_i |i\rangle$$

Want: approximate with k-sparse ← only k nonzero coefficients

Solution: keep largest k coefficients

→ this is the answer for the fidelity

fidelity: F^2 + trace distance: $T^2 = 1$, $F = \sqrt{1-\varepsilon} \rightarrow T = \sqrt{\varepsilon}$

Fuchs-van Graaf inequalities: $\varepsilon \lesssim T \leq \sqrt{\varepsilon}$

↖ mixed states can get quadratic improvement

Note when do you care about T? if you care about expectation values of a fixed observable.

Example

$$|\psi\rangle = \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|1\rangle + \sqrt{\varepsilon}\,|1\rangle \qquad k=2$$

Best for fidelity: $|\phi_1\rangle \sim \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|1\rangle$

$$|\phi_2\rangle \sim \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|2\rangle$$

$$|\psi\rangle\langle\psi| = \begin{pmatrix} 1-2\varepsilon & \sqrt{\varepsilon(1-2\varepsilon)} & \sqrt{\varepsilon(1-2\varepsilon)} \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & \varepsilon \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & \varepsilon \end{pmatrix}$$

Example

$$|\psi\rangle = \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|1\rangle + \sqrt{\varepsilon}\,|1\rangle \qquad k=2$$

Best for fidelity: $|\phi_1\rangle \sim \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|1\rangle$

$$|\phi_2\rangle \sim \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|2\rangle$$

largest term

next-largest...

$$|\psi\rangle\langle\psi| = \begin{pmatrix} 1-2\varepsilon & \sqrt{\varepsilon(1-2\varepsilon)} & \sqrt{\varepsilon(1-2\varepsilon)} \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & \varepsilon \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & \varepsilon \end{pmatrix}$$

$$|\phi_1\rangle\langle\phi_1| \sim \begin{pmatrix} 1-2\varepsilon & \sqrt{\varepsilon(1-2\varepsilon)} & 0 \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad |\phi_2\rangle\langle\phi_2| \sim \begin{pmatrix} 1-2\varepsilon & 0 & \sqrt{\varepsilon(1-2\varepsilon)} \\ 0 & 0 & 0 \\ \sqrt{\varepsilon(1-2\varepsilon)} & 0 & \varepsilon \end{pmatrix}$$

Example

$$|\psi\rangle = \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|1\rangle + \sqrt{\varepsilon}\,|1\rangle \qquad k=2$$

Best for fidelity: $|\phi_1\rangle \sim \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|1\rangle$

$$|\phi_2\rangle \sim \sqrt{1-2\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|2\rangle$$

largest term

$$|\psi\rangle\langle\psi| = \begin{pmatrix} 1-2\varepsilon & \sqrt{\varepsilon(1-2\varepsilon)} & \sqrt{\varepsilon(1-2\varepsilon)} \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & \varepsilon \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & \varepsilon \end{pmatrix}$$

next-largest...

each of these gives error $\sqrt{\varepsilon}$ in trace dist.

$$|\phi_1\rangle\langle\phi_1| \sim \begin{pmatrix} 1-2\varepsilon & \sqrt{\varepsilon(1-2\varepsilon)} & 0 \\ \sqrt{\varepsilon(1-2\varepsilon)} & \varepsilon & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad |\phi_2\rangle\langle\phi_2| \sim \begin{pmatrix} 1-2\varepsilon & 0 & \sqrt{\varepsilon(1-2\varepsilon)} \\ 0 & 0 & 0 \\ \sqrt{\varepsilon(1-2\varepsilon)} & 0 & \varepsilon \end{pmatrix}$$

$\rho = \frac{1}{2}|\phi_1\rangle\langle\phi_1| + \frac{1}{2}|\phi_2\rangle\langle\phi_2|$ has error $O(\varepsilon)$ in trace distance

Example

$$|\psi\rangle = \sqrt{1-\varepsilon}\,|0\rangle + \sqrt{\varepsilon}\,|1\rangle \qquad k=1$$

Best for fidelity: $|\phi\rangle = |0\rangle$

$$|\psi\rangle\langle\psi| = \begin{pmatrix} 1-\varepsilon & \sqrt{\varepsilon(1-\varepsilon)} \\ \sqrt{\varepsilon(1-\varepsilon)} & \varepsilon \end{pmatrix}$$

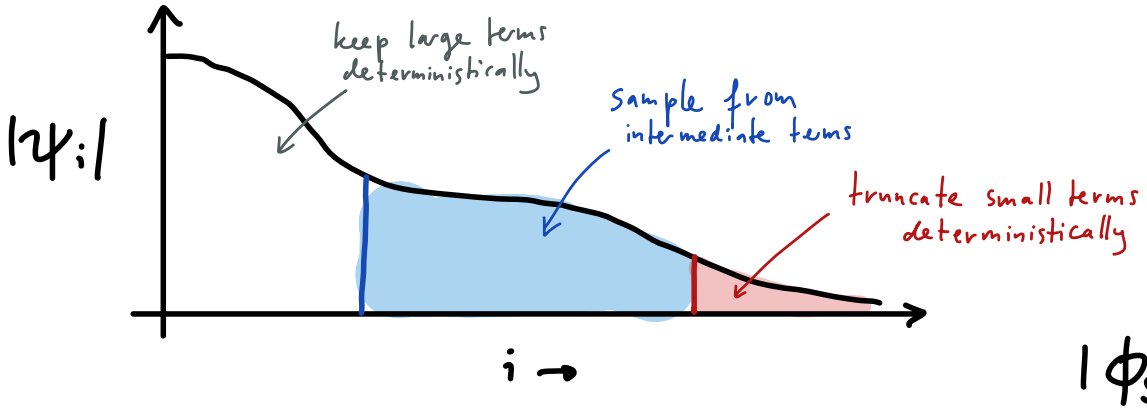
$\rho = \begin{pmatrix} 1-\varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$

Best trace distance: error $O(\sqrt{\varepsilon})$

Results:

if target state is mixed, ⇒ NP-hard

A Show how to exactly compute optimal trace distance approximations for pure states



Note this follows up on & extends previous work of Regula, Johnston and others, mostly focused on robustness

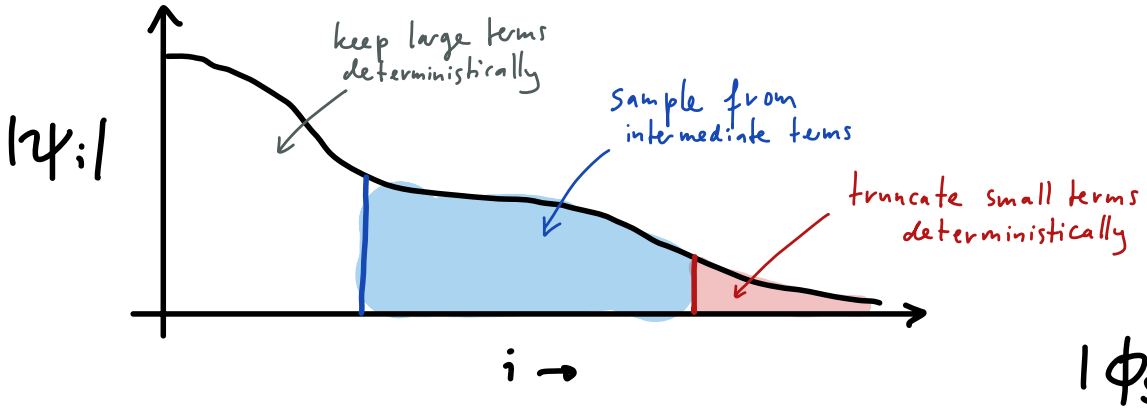
$$|\phi_S\rangle = \alpha \sum_{i=1}^{k-r} \psi_i |i\rangle + \beta \frac{1}{\sqrt{r}} \sum_{i \in S} |i\rangle$$

↖ sample S to have correct marginals

Results:

if target state is mixed, $\Rightarrow$ NP-hard

A Show how to exactly compute optimal trace distance approximations for pure states



Note this follows up on & extends previous work of Regula, Johnston and others, mostly focused on robustness

$$|\phi_S\rangle = \alpha \sum_{i=1}^{k-r} \psi_i |i\rangle + \beta \frac{1}{\sqrt{r}} \sum_{i \in S} |i\rangle$$

↖ sample S to have correct marginals

B Show how to efficiently sample this density matrix

Application

Truncate:

MPS: (or other tensor networks)

k-spares | Schmidt rank k | bond dimension k

no provable optimal global method, but can truncate per bond

previous work by Ferris

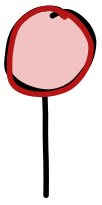
Tensor networks often need to **truncate** bond dimension,

- may be worth to **sample** multiple times from lower bond dim TN
- we care about **observables** (e.g. energy)

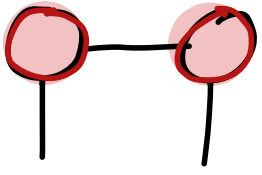
Application

Truncate:

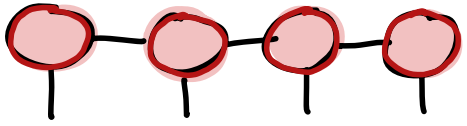
MPS: ← or other tensor networks



k-spares



Schmidt rank k



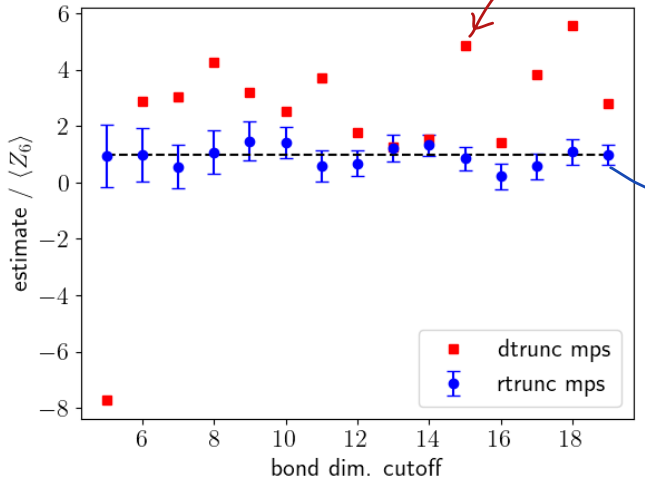
bond dimension k

↓

no provable optimal global method, but can truncate per bond

det. truncation

relative error observable



randomized truncation

3 Hamiltonian simulation & product formulas

$$H = \sum_{\ell=1}^{L} H_\ell - \sum_{\ell=1}^{L} h_\ell P_\ell$$

P_ℓ ← Pauli operator

Want: e^{itH}

Can implement: e^{itP_ℓ} ← Pauli rotation

3 Hamiltonian simulation & product formulas

$$H = \sum_{l=1}^{L} H_l - \sum_{l=1}^{L} h_l P_l$$

P_l: Pauli operator

Want: e^{itH}

Can implement: e^{itP_l} ← Pauli rotation

Solution: $e^{i\delta H} \approx e^{i\delta H_1} \cdots e^{i\delta H_L}$

$$e^{itH} = \left(e^{i\frac{t}{r}H}\right)^r \approx \left(e^{i\frac{t}{r}H_1} \cdots e^{i\frac{t}{r}H_L}\right)^r$$

- choose r so error is small
- # Pauli rotations $\sim L \cdot t^{1+1/p}$ for p-th order

3 Hamiltonian simulation & product formulas

$$H = \sum_{\ell=1}^{L} H_\ell = \sum_{\ell=1}^{L} h_\ell P_\ell$$

P_ℓ: Pauli operator

Want: e^{itH}

Solution: product formula (Trotterization)

Pauli rotations $\sim L \cdot t^{1+1/p}$ for p-th order

↳ downsides:
- scales with L → chemistry: $L = O(N^4)$, N = # qubits
- suboptimal scaling with t
- subtle error dependence ignored here...

3 Hamiltonian simulation & product formulas

$$H = \sum_{\ell=1}^{L} H_\ell = \sum_{\ell=1}^{L} h_\ell P_\ell$$ ← Pauli operator

Want: e^{itH}

Solution: product formula (Trotterization)

Pauli rotations $\sim L \cdot t^{1+1/p}$ for p-th order

↳ downsides:
- scales with L → chemistry: $L = O(N^4)$, N = # qubits
- suboptimal scaling with t
- subtle error dependence ignored here...

Alternative: random product formula (qDRIFT)

normalize by $\lambda = \sum_{\ell=1}^{L} |h_\ell|$ → $H = \lambda \sum_\ell p_\ell P_\ell$ (prob. dist)

(3) Hamiltonian simulation & product formulas

$$H = \sum_\ell p_\ell P_\ell$$

(p_ℓ: prob. dist; P_ℓ: Pauli operator)

Campbell

random product formula (qDRIFT)

$$e^{i\delta H} = I + i\sum_\ell p_\ell P_\ell + O(\delta^2)$$

$$= \sum_\ell p_\ell \underbrace{(I + i\delta P_\ell)}_{\propto\, e^{i \arctan(\delta) P_\ell}} + O(\delta^2)$$

- Sample according to p_ℓ, apply $e^{i \arctan(\delta) P_\ell}$

3 Hamiltonian simulation & product formulas

$$H = \sum_\ell p_\ell P_\ell$$

p_ℓ: prob. dist; P_ℓ: Pauli operator

Campbell

random product formula (qDRIFT)

$$e^{i\delta H} = I + i\sum_\ell p_\ell P_\ell + O(\delta^2)$$

$$= \sum_\ell p_\ell \underbrace{(I + i\delta P_\ell)}_{\propto\, e^{i \arctan(\delta) P_\ell}} + O(\delta^2)$$

When using in a Hadamard test, we get unbiased estimate of $\langle\psi| e^{itH} |\psi\rangle$

- Sample according to p_ℓ, apply $e^{i \arctan(\delta) P_\ell}$
- can also get all higher order terms right

Wan-Berta-Campbell

$$e^{itH} = \left(e^{i\frac{t}{r}H}\right)^r \propto \text{sampling } r \text{ Pauli rotations}$$

crucial: to keep normalization $\Omega(1)$, need

$$r \sim \lambda^2 t^2$$

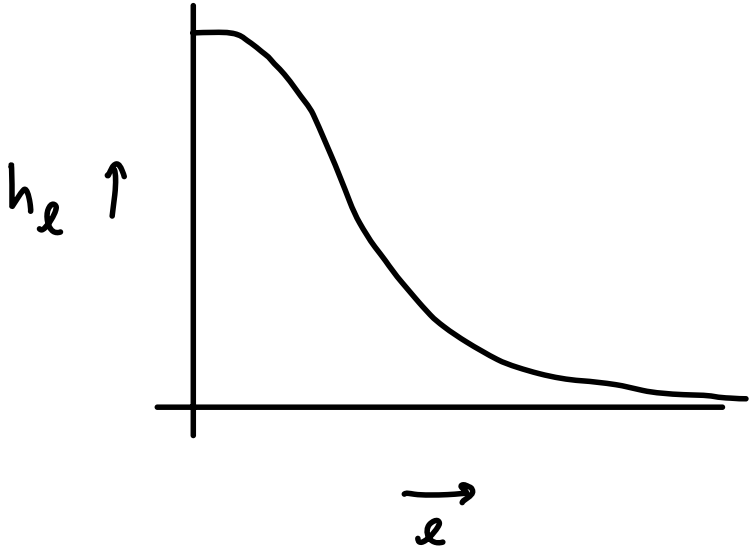
3 Hamiltonian simulation & product formulas

Work with Jakob Günther, Alexander Schmidhuber, Marek Miller, Matthias Christandl & Aram Harrow

Deterministic: $\sim L t^{1+1/p}$

Random: $\sim \lambda^2 t^2$

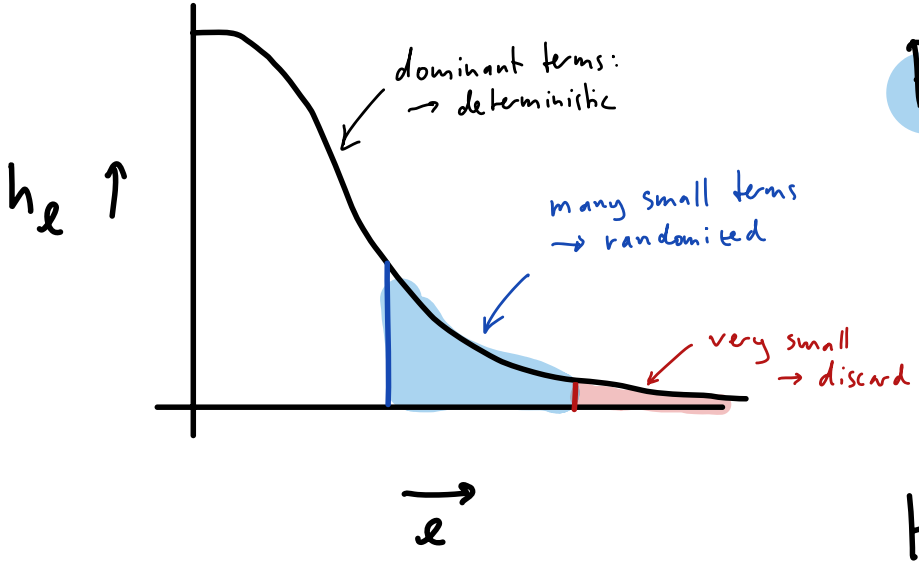
Realistic Hamiltonians: $L \sim N^4$



3 Hamiltonian simulation & product formulas

Deterministic: $\sim L t^{1+1/p}$

Random: $\sim \lambda^2 t^2$



Hagan & Wiebe

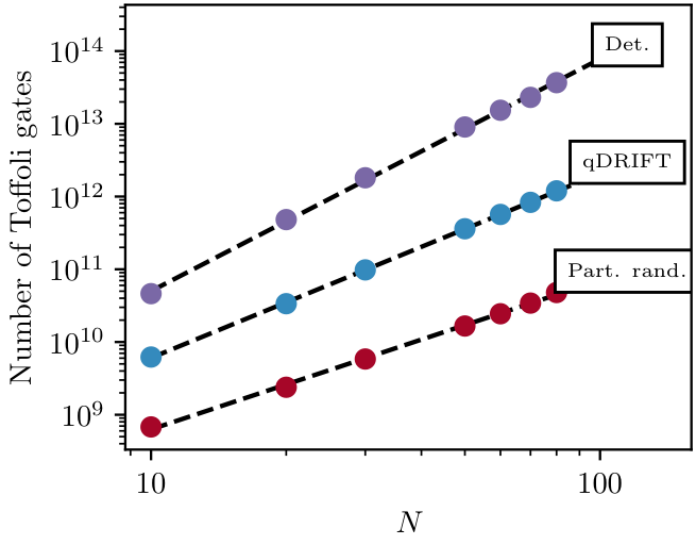
Partially random product formula

$$H = \sum_{\ell=1}^{L'} H_\ell + \lambda_R H_R$$

↖ balance L' and λ_R

Results

1. simple error analysis
2. detailed numerical investigation Trotter error
3. resource estimates for quantum phase estimation ground state energy chemistry examples



← gate count phase estimation ground state energy chain of hydrogen atoms