Resource-efficient quantum eigenvalue transform with commutator scaling

Arul Rhik Mazumder[1,2] James D. Watson[3] Samson Wang[1]
[1]Caltech [2]Carnegie Mellon University [3]Google Quantum AI

1 Product formulae fit early fault-tolerant hardware

For $H=\sum_\gamma H_\gamma$, $\Lambda := \sum_\gamma \|H_\gamma\|$, a pth-order product formula is *staged*:

$$\mathcal{P}(t) = \prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma} e^{-it\,a_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}$$

Υ sweeps over all Γ terms, each with its own coefficients and ordering π_υ. Then r repetitions:

stage 1 stage Υ

$$|\psi(0)\rangle \; \boxed{e^{-ia_{(1,1)}H_{\pi_1(1)}T/r}} \boxed{\cdots} \boxed{e^{-ia_{(1,\Gamma)}H_{\pi_1(\Gamma)}T/r}} \boxed{\cdots} \boxed{e^{-ia_{(\Upsilon,1)}H_{\pi_\Upsilon(1)}T/r}} \boxed{\cdots} \boxed{e^{-ia_{(\Upsilon,\Gamma)}H_{\pi_\Upsilon(\Gamma)}T/r}} \; |\psi(T)\rangle$$

- **No ancillas, no block encoding** — each $e^{-iH_\gamma t}$ is a native gate.
- **Commutator scaling:** depth tracks $(\alpha_{\mathrm{comm}}^{(p+1)})^{1/p}$, nested-commutator norms, *not* the far larger Λ. No other simulation method does this.
- **But # of steps scale poorly with error:** $r = O\big((\alpha_{\mathrm{comm}}^{(p+1)})^{1/p}T^{1+1/p}\,\varepsilon^{-1/p}\big)$
- **And Υ is not free:** Suzuki gives $\Upsilon_{2k}=2\cdot 5^{k-1}$, so $p=2$ costs 2 operators per term per step but $p=6$ costs 50. Gate depth scales it: $O\big(\Gamma\Upsilon(a_{\max}\Upsilon\lambda_{\mathrm{comm}}T)^{1+1/p}\log(1/\varepsilon)\big)$.

2 Extrapolate the step size to zero

Trotter error is not noise, it is a *power series in the step size* s:

$$\mathrm{Tr}[\rho\,\mathcal{P}(sT)^{1/s}] - \mathrm{Tr}[\rho\,e^{-iHT}] \;=\; \sum_{\ell\ge p} s^\ell \delta_\ell\,, \qquad |\delta_\ell| \le (c\,\lambda_{\mathrm{comm}}T)^{\ell(1+1/p)}.$$

A combination over a schedule $\{s_j\}_{j=1}^m$ with Richardson coefficients $\{b_j\}$ eliminates the leading $m-1$ orders:

$$\mathrm{Tr}[\rho\,e^{-iHT}] \;\approx_{\varepsilon_R}\; \sum_{j=1}^{m} b_j\;\mathrm{Tr}\big[\rho\,\mathcal{P}^{1/s_j}(s_jT)\big].$$

Every term is a *shallower* circuit than the target accuracy would otherwise demand.

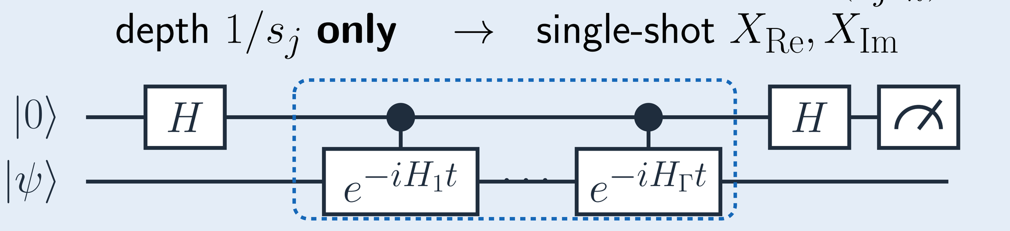
Bring your own schedule: any scheme cancelling $p, p{+}\sigma, \ldots, p{+}(m{-}2)\sigma$ gives gate depth $O\big(\Gamma(\lambda_{\mathrm{comm}}T)^{1+1/p}(4\|\vec b\|/\varepsilon)^{1/(\sigma(m-1)+p)}\big)$. That exponent is never worse than $1/p$.

3 . . . but don't build it coherently — sample it

A linear combination of m circuits normally means **LCU** ancillas, block encoding, amplitude ampl. **We don't build it. We sample it** from its Fourier series $f(H)\approx\sum_k c_k e^{-iHt_k}$

CLASSICAL precompute $p(k,j)=|c_kb_j|/S$, $S=\|\vec c\|_1\|\vec b\|_1$
$\rightarrow$ sample *one* pair (k,j), set $s_j \leftarrow s/q_j$

QUANTUM *one* Hadamard test of $\mathcal{P}^{1/s_j}(s_jt_k)$
depth $1/s_j$ **only** $\rightarrow$ single-shot $X_{\mathrm{Re}}, X_{\mathrm{Im}}$



repeat $1/s_j$ times

CLASSICAL reweight $\mu \leftarrow S\cdot\mathrm{sign}(c_kb_j)(X_{\mathrm{Re}}+iX_{\mathrm{Im}})$
$\rightarrow$ average the μ's $\Rightarrow$ **estimate**

repeat $C_{\mathrm{sample}} = O\big(\frac{\|\vec b\|_1^2}{\varepsilon^2}\big)$

- LKW [5] gives $\|\vec b\|_1 = O(\log m)$ with $m = O(\log\log(1/\varepsilon))$ points, so $\|\vec b\|_1^2 = O\big((\log\log(1/\varepsilon))^2\big)$ — depth $\varepsilon^{-1/p}\rightarrow\log(1/\varepsilon)$ for log-log samples, one ancilla.

4 The result

Estimating a time signal $\mathrm{Tr}[\rho\,e^{-iHT}]$ to additive precision ε:

method	ancillary qubits	max depth per sample $\widetilde{O}(\cdot)$	sample overhead $O(\cdot)$	commutator scaling?
Qubitization [2]	$\lceil\log\Gamma\rceil+3$	$\Gamma\Big(\Lambda T + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)}\Big)$	$1/\varepsilon^2$	
Multiproduct formula [6]	$\widetilde{O}(\log\log(\lambda_{\mathrm{comm}}T/\varepsilon))$	$\Gamma\lambda_{\mathrm{comm}}T\log^2(\lambda_{\mathrm{comm}}T/\varepsilon)$	$1/\varepsilon^2$	✓
Product formulae [1]	1	$\Gamma(\alpha_{\mathrm{comm}}^{(p+1)})^{1/p}T^{1+1/p}\,(1/\varepsilon)^{1/p}$	$1/\varepsilon^2$	✓
Random compiler [3]	1	Λ^2T^2	$1/\varepsilon^2$	
Randomized Extrapolation Trotterization *(this work)*	**1**	$\Gamma(\lambda_{\mathrm{comm}}T)^{1+1/p}\log(1/\varepsilon)$	$\frac{(\log\log(1/\varepsilon))^2}{\varepsilon^2}$	✓

vs. prior extrapolation [4, 8]: Watson–Watkins [4] do $f=e^{-iHt}$ only and we *reduce* to them; Chakraborty et al. [8] reach the same depth for Task 1.
Ours adds: randomization $\Rightarrow$ refined *sample* complexity · Task 2 (distributions) · the four refined scalings below. These apply to [4], [8] too.

5 Not just time signals

Sampling a Fourier series *inside the same randomization loop* gives the eigenvalue transform at the same depth — what **QSVT** does, without its multi-qubit controls or ancilla register:

- **Task 1:** estimate $\mathrm{Tr}[\rho\,U f(H)]$ and $\mathrm{Tr}[f(H)\rho f(H)^\dagger O]$.
- **Task 2:** the *distribution* $p_i = |\langle i|\,Uf(H)\,|\psi\rangle|^2$ to ε in ℓ_2, for $(\log\log(1/\varepsilon))^4/\varepsilon^2$ samples.
- **End-to-end:** every oracle instantiated, including the time-evolution oracle Lin–Tong leave abstract in **Ground state energy**, **Green's functions**, **time-evolved states**.

6 Is λ_{comm} a cop-out?

Everything is hidden in the *extrapolated* commutator factor λ_{comm}.

$$\lambda_{\mathrm{comm}} \;:=\; \sup_{\substack{j\ge p\\ 1\le\ell\le\lfloor j/p\rfloor}} \Bigg[\sum_{\substack{j_1,\ldots,j_\ell\ge p\\ j_1+\cdots+j_\ell=j}} \prod_{\kappa=1}^{\ell} \frac{\alpha_{\mathrm{comm}}^{(j_\kappa+1)}}{(j_\kappa+1)^2}\Bigg]^{\frac{1}{j+\ell}}$$

A supremum over *every* order, where plain Trotter needs $\alpha_{\mathrm{comm}}^{(p+1)}$ at one. Still tightly bounded:

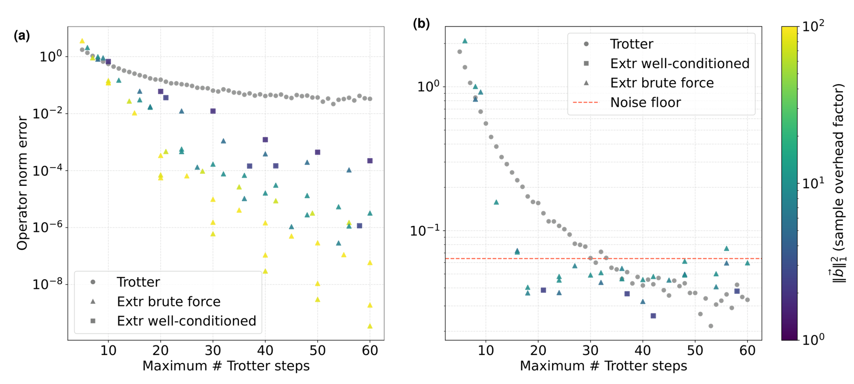
- For plane-wave electronic structure $\alpha_{\mathrm{comm}}^{(j)} = O(n^j) \Rightarrow \lambda_{\mathrm{comm}} = O(n)$, the *same* n-dependence as plain Trotter ($\lambda_{\mathrm{comm}}^{1+1/p}/(\alpha_{\mathrm{comm}}^{(p+1)})^{1/p} \le 1.50,\ 1.15,\ 1.04$ for $p=1,2,4$).

Always $\lambda_{\mathrm{comm}} \le 4\Lambda$. We replace it in special settings:

setting	gate depth becomes $\widetilde{O}(\cdot)$	note
α_{comm} known only at fixed order p	$\Gamma\cdot\max\big\{\Lambda T,\ (\alpha_{\mathrm{comm}}^{(p+1)})^{1/p}T^{1+1/p}\big\}$	needs nothing beyond what plain Trotter already needs
k-local, n qubits, onsite energy g	$\Gamma\,(\widetilde\lambda_{k\text{-local}}T)^{1+1/p}$, $\quad(\widetilde\lambda_{k\text{-local}})^{1+1/p} = O\big(kg(p\Lambda^{1/p}+g\log\frac{ngT}{\varepsilon})\big)$	naive substitution *diverges*, blocking [4, 8]; fixed via [7]
long tail, $H=H_A+H_B$ *(doubly randomized)*	$\Gamma_A(\lambda_{\mathrm{comm}}T)^{1+1/p}+\Lambda_B^2T^2$	interpolates with [3]: *never worse* than it
input state in one symmetry sector (e.g. fermion number η)	$\Gamma\,(\lambda_{\mathrm{comm}}^{(\eta)}T)^{1+1/p}$	fermionic *semi*-norms of nested commutators, not operator norms [10]

Or don't bound λ_{comm} at all: once $s\,(c\,\lambda_{\mathrm{comm}}T)^{1+1/p}\le C<1$ the step is inside the *radius of convergence* and block 2's series decays geometrically — every extra point provably helps, **no a priori bound needed**.

7 It works in practice



8-qubit Heisenberg, $p=2$, exact error. (a) every schedule sits orders of magnitude below Trotter, and far earlier than the bounds predict. **(b)** under measurement noise it hits the noise floor in fewer Trotter steps. Colour: sample overhead $\|\vec b\|_1^2$.

Bounds (not shown), plane-wave $n=100$: beats the best Suzuki–Trotter *at any order* below $\varepsilon\approx 10^{-2}$, at 10–100× samples.

8 Ground state energy

Ansatz overlap $\eta=\mathrm{Tr}[\Pi_0\rho]$, ground energy to precision ε:

method	anc. qubits	max depth per sample $\widetilde{O}(\cdot)$	samples $\widetilde{O}(\cdot)$
QET-U + qubitization [9]	$\lceil\log\Gamma\rceil+3$	$\Gamma\Lambda\,\varepsilon^{-1}$	η^{-1}
QET-U + Trotter [9]	1	$\Gamma(\alpha_{\mathrm{comm}}^{(p+1)})^{1/p}\,\eta^{-1/p}\varepsilon^{-(1+1/p)}$	η^{-1}
Random compiler [3]	1	$\Lambda^2\varepsilon^{-2}$	η^{-2}
this work	**1**	$\Gamma\lambda_{\mathrm{comm}}^{1+1/p}\varepsilon^{-(1+1/p)}$	η^{-2}

Our depth carries no η; Trotter-based competitors pay $\eta^{-1/p}$.

References

[1] Childs et al., "Trotter Error with Commutator Scaling," PRX 2021.
[2] Low, Chuang, "Hamiltonian Sim. by QSP/Qubitization," 2017/2019.
[3] Wan, Berta, Campbell, "Rand. Statistical Phase Estimation," PRL 2022.
[4] Watson, Watkins, "Trotter Error Mitigation," PRX Quantum 2025.
[5] Low, Kliuchnikov, Wiebe, "Well-conditioned multi-product," 2019.
[6] Aftab, An, Trivisa, "Multi-product with explicit comm. scaling," 2024.
[7] Mizuta, "Comm. scaling in multi-product formulas," Quantum 2026.
[8] Chakraborty et al., "QSVT without block encodings," 2025.
[9] Dong, Lin, Tong, "Ground state prep. and energy est. (QET-U)," 2022.
[10] Su 2021; McArdle 2022; Low 2023 — fermionic semi-norms.

9 Open questions

- **What is learnable about $f(H)$ with no control?** Every circuit here is a Hadamard test; the product formula itself can be conjugated instead.
- **Better schedules than brute force?** A simple integer-grid search already beats the LKW grid; block 2's bound holds for *any* schedule.
- **Trotter's other refinements?** Plain product formulae do better on low-energy input states, and beat their own worst-case bounds on average — does either survive extrapolation?

Full paper:
arXiv:2608.13862
arul.rhik@gmail.com
Supported by Caltech's SURF and IQIM, an NSF Physics Frontiers Center (NSF Grant PHY-2317110).