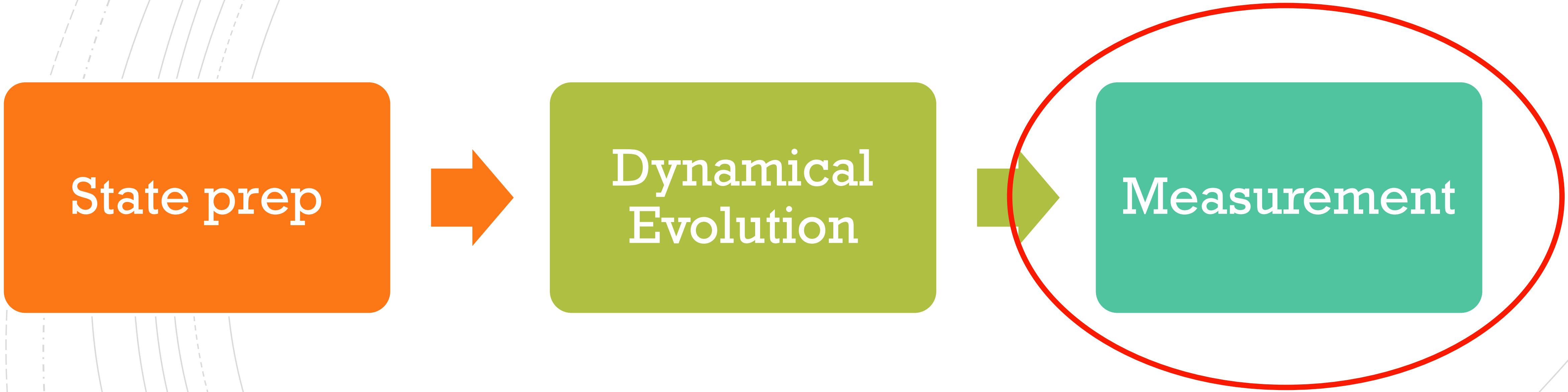
# Polynomial Extrapolation for Trotter

Jacob Watkins, Gumaro Rendon, Nathan Wiebe

ECT* Workshop, June 7th, 2023

# The quantum simulation “workflow”

# The quantum simulation “workflow”



# Two Measurement Examples

## Phase Estimation

Scaling[1]

$$\Theta(\log(1/\alpha)\,/\epsilon)$$

Physics application: eigenvalues[2]

$$e^{-i\,H\,t}|E\rangle = e^{-iEt}|E\rangle$$

## Amplitude Estimation

Scaling[3]

$$\Theta(\log(1/\alpha)\,/\epsilon)$$

Physics application: expectation values[4]

$$\langle O(t)\rangle = Tr(\rho U^{\dagger}(t) O U(t))$$

Here

- $\alpha$ is the failure rate
- $\epsilon$ is the precision


1. Giovannetti, Lloyd, Maccone (2006)
2. Abrams, Lloyd (1998)
3. Suzuki, Uno, et. al (2020)
4. Knill, Ortiz, Somma (2007)

# Two Measurement Examples

## Phase Estimation

Scaling[1]

$$\Theta(\log(1/\alpha)\,/\epsilon)$$

Physics application: eigenvalues[2]

$$e^{-i\,\tilde{H}\,t}|\tilde{E}\rangle = e^{-i\tilde{E}t}|\tilde{E}\rangle$$

## Amplitude Estimation

Scaling[3]

$$\Theta(\log(1/\alpha)\,/\epsilon)$$

Physics application: expectation values[4]

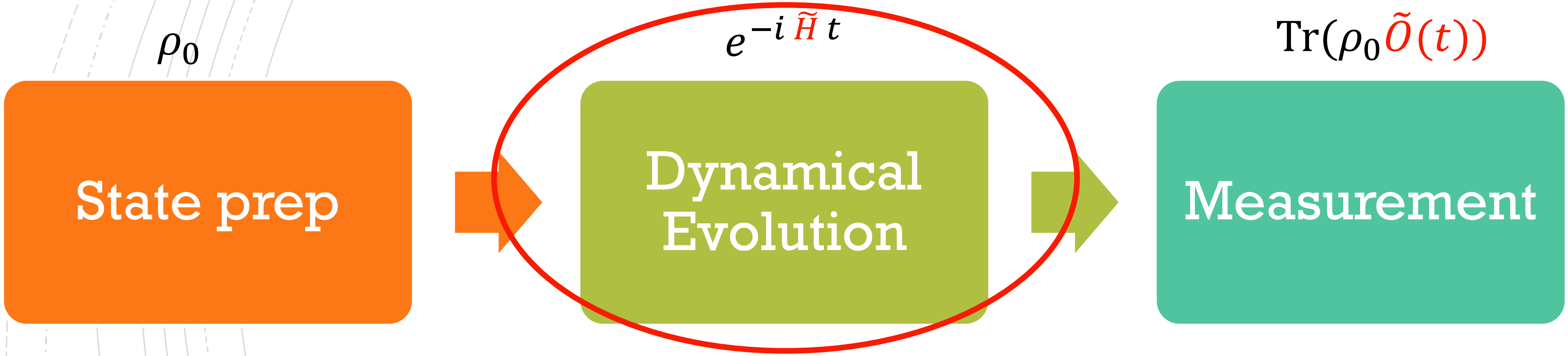
$$\langle\tilde{O}(t)\rangle = \mathrm{Tr}(\rho\tilde{U}^{\dagger}(t)O\tilde{U}(t))$$

Here

- $\alpha$ is the failure rate
- $\epsilon$ is the precision

1. Giovannetti, Lloyd, Maccone (2006)
2. Abrams, Lloyd (1998)
3. Suzuki, Uno, et. al (2020)
4. Knill, Ortiz, Somma (2007)

# The quantum simulation “workflow”



# What are our options?

## Trotter

$$\exp\left(\sum_j -i\, H_j\, t\right) = \prod_j \exp(-i\, H_j t) + O(t^2)$$

✓ Commutator scaling[1]

✓ Zero overhead

✓ Modular, many choices[2, 3]

X Relatively inaccurate: for order $p > 0$

$$N_{\exp} \in O\left(\sqrt[p]{1/\epsilon}\right)$$

## Others

✓ Better accuracy: Cost $\in O(\log 1/\epsilon)$[4, 5]

✓ Asymptotically optimal performance[6] in $t, \epsilon$

X More overhead cost

- X Auxiliary qubits
- X Lots of control gates


1. Childs, Su, Tran, Wiebe, Zhu (2019)
2. Campbell (2019)
3. Ikeda, Abrar, Chuang, Sugiura (2023)
4. Childs, Wiebe (2012)
5. Babbush, Berry, Kivlichan, Wei, Love, Aspuru-Guzik (2016)
6. Low, Chuang (2019)

# Effect on Parameter Estimation

Consider eigenvalue estimation with 1st order Trotter.

$$\exp \sum_j -i\, H_j\, t = \left(\prod_j \exp -i\, H_j\, t\, /r\right)^r + O(t^2/r)$$
$$= \exp -i\widetilde{H}_r t$$

with "effective Hamiltonian" $\widetilde{H}_r$ that has eigenvalues

$$\widetilde{E}_r = E + O(t^2/r).$$

To achieve accuracy $\epsilon$, we require $N \in O(1/\epsilon)$ experiments and $r \in O(t^2/\epsilon)$ Trotter steps. Thus,

$$N_{exp} \propto N\, r \in O(t^2/\epsilon^2),$$

which is quadratically worse than the Heisenberg limit.

# What are our options?

## Trotter

$$\exp\left(\sum_j -i\, H_j\, t\right) = \prod_j \exp(-i\, H_j t) + O(t^2)$$

✓ Commutator scaling[1]

✓ Zero overhead

✓ Modular, many choices[2, 3]

X Relatively inaccurate: for order $p > 0$

~~$N_{\exp} \in O(\sqrt[p]{1/\epsilon})$~~ ← Won't give HL

## Others

✓ Better accuracy: $\text{Cost} \in O(\log 1/\epsilon)$[4, 5] — HL up to logs (not bad)

✓ Asymptotically optimal performance[6] in $t, \epsilon$

X More overhead cost

- X Auxiliary qubits
- X Lots of control gates

1. Childs, Su, Tran, Wiebe, Zhu (2019)
2. Campbell (2019)
3. Ikeda, Abrar, Chuang, Sugiura (2023)
4. Childs, Wiebe (2012)
5. Babbush, Berry, Kivlichan, Wei, Love, Aspuru-Guzik (2016)
6. Low, Chuang (2019)

# Can we achieve near HL scaling using only Trotter and classical resources?

# How do other algorithms do it?

Consider *multiproduct formulas* with Trotter formula $S_p$.

$$\widetilde{U}_{MPF}(t) := \sum_{j=1}^{n} a_j\, S_p(t/k_j)^{k_j} = U(t) + O(t^{n+p-1})$$

Trotter formulas are added *coherently* using the linear combo of unitaries (LCU) technique,[1] or offline with random sampling.[2]

This is a form of Richardson extrapolation to $r \to \infty$.

In the end, we get $O(\log 1/\epsilon)$ simulation cost.

1. Low, Kliuchnikov, Wiebe (2019)
2. Faehrmann, Steudtner,

# But we can do extrapolation *classically* (offline)

$$\widetilde{U}_{MPF}(t) = \sum_j a_j U_j \quad \rightarrow \quad \langle \tilde{O}(t) \rangle = \sum_j a_j \langle O_{k_j}(t) \rangle$$

This has been considered in context of noisy Hamiltonian simulation[1] and linear systems.[2]

This has been demonstrated on IBM hardware with observed improvements.[3]

However, these works lack a theoretical analysis of algorithmic performance.

Other techniques for constructing estimates are less explored.


1. Endo, Zhao, Li, Benjamin, Yuan (2019)
2. Vazquez, Hiptmair, Woerner (2022)
3. Vazquez, Egger, Ochsner, Woerner (2022)

# What is our contribution?

!

- We analyze polynomial interpolation for extrapolating Suzuki-Trotter formulas to zero step size.
- We do a full theoretical analysis of cost in terms of algorithmic errors (no external noise).
- We look specifically at
  - Eigenvalues (via phase estimation) $H|E\rangle = E|E\rangle$
  - Expectation values (via amplitude estimation) $\langle O(t)\rangle$
- We achieve "near" HL scaling: $\tilde{O}(1/\epsilon)$
- Thus, Trotter alone is sufficient for high accuracy estimation tasks relevant to physics. No additional quantum resources needed.

# Set up and notation

# Suzuki-Trotter formulas

Let

$$S_{2k}(t/r)^r = U(t) + O(t^{2k+1}/r^{2k})$$

be the order $2k$ symmetric Suzuki-Trotter formula in $r$ steps.

At lowest order ($k = 1$),

$$S_2(t) = \exp\left(\frac{-i\,H_1 t}{2}\right) \dots \exp\left(\frac{-iH_m t}{2}\right)\,\exp\left(\frac{-iH_m t}{2}\right) \dots \exp\left(\frac{-i\,H_1 t}{2}\right),$$

with higher $k$ defined recursively (with # of terms exponential in $k$).

# $s$ parametrization and effective Hamiltonian

Instead of number of steps $r$, let's consider "dimensionless step size" $s \coloneqq 1/r$.

$$\widetilde{U}_s(t) \coloneqq S_{2k}(st)^{1/s}$$

The above formula suggests extending the definition to real valued $s \in [-1,1]$.

Also, $\widetilde{U}_0 = U$ is the exact propagator.

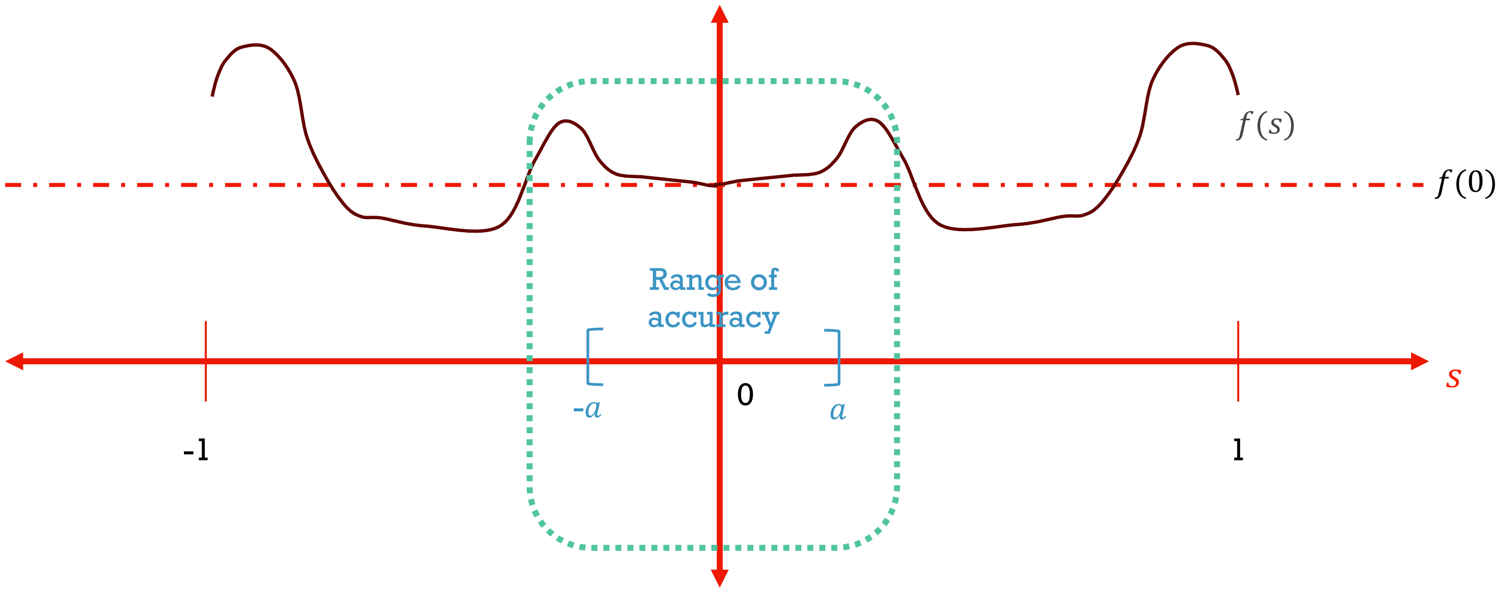
Observe that simulations become more expensive as $s \to 0$.

We can define an effective Hamiltonian $\widetilde{H}_s$ such that

$$\widetilde{U}_s(t) = \exp(-i\, \widetilde{H}_s t)$$

# $s$-dependent observables

To illustrate, let $f(s) \coloneqq \langle \tilde{O}_s(t) \rangle$.



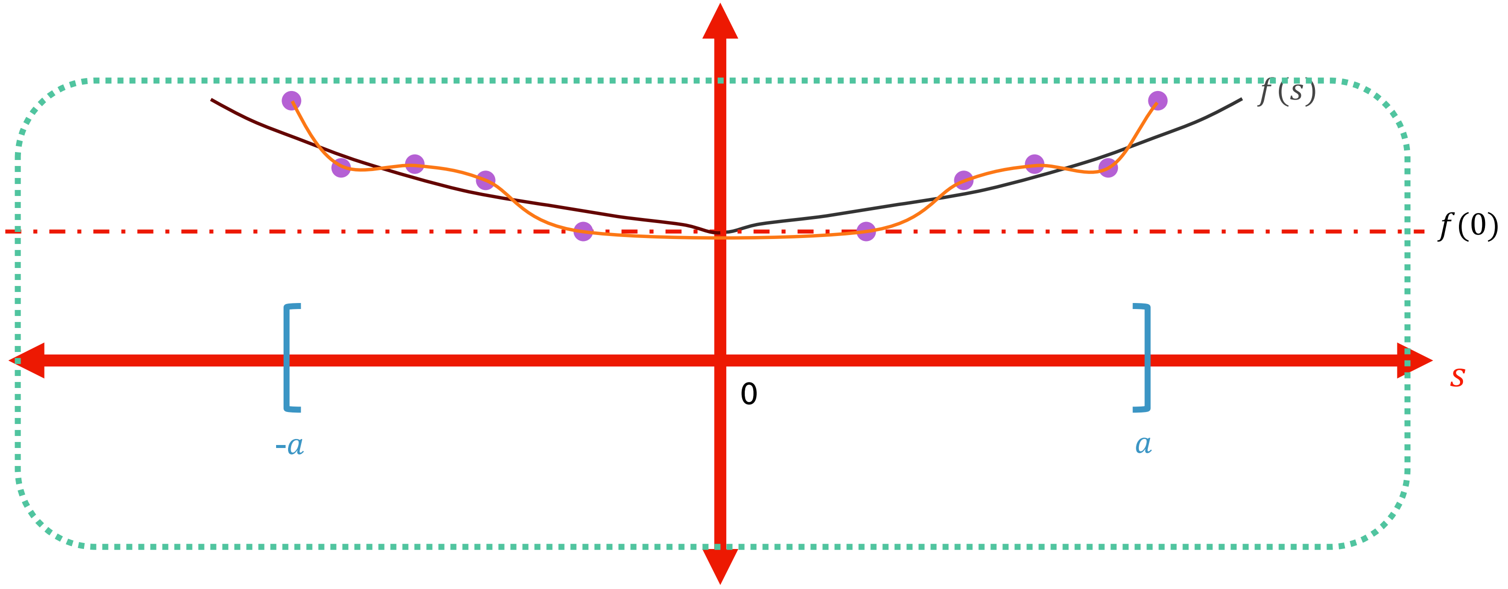

Fact: $f(-s) = f(s)$, meaning $f(s)$ is even (for Suzuki-Trotter).

$a \approx 1/(\|H\|t)$

# Zooming in on accuracy range $[-a, a]$

Let's estimate some values $\tilde{y}_i$ of $f(s_i)$ (choice of $s_i$ matters!).

Then let's construct interpolant $P_{n-1}f(s)$

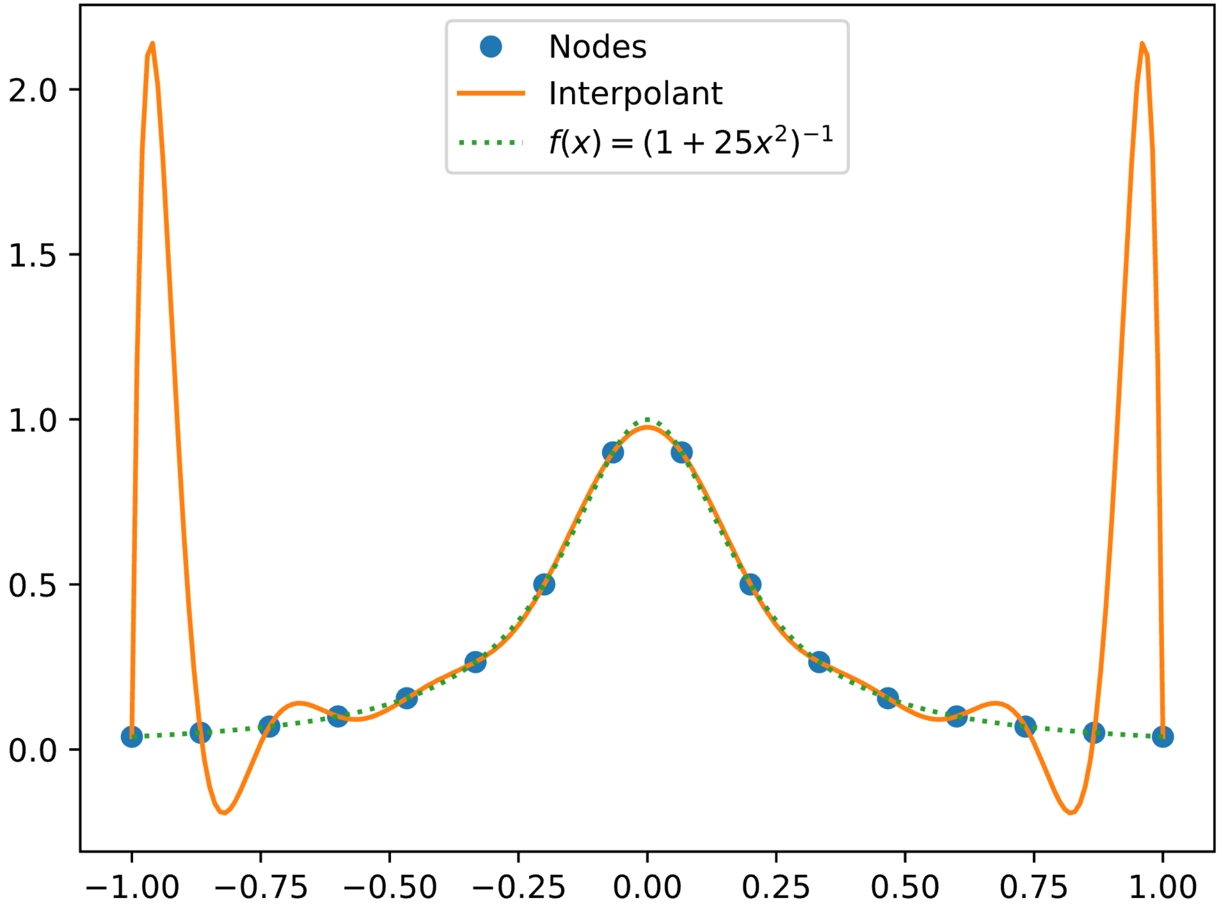



Our estimate for $f(0)$ is then $P_{n-1}f(0)$

# Finer Points

# Choice of nodes matters

Equally spaced nodes lead to wild oscillations




Runge phenomenon (John D. Cook)

# The solution: Chebyshev interpolation

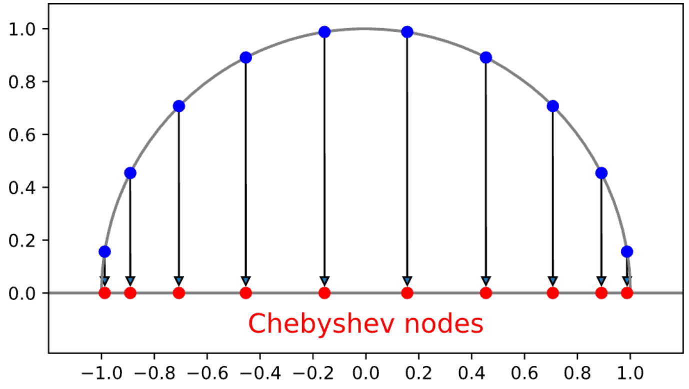
These are simply projections of equally spaced points on the radius $a$ circle.

$$s_k = a\cos\left(\frac{2k-1}{2n}\pi\right), \quad k = 1, \dots, n$$

Equivalently, the zeros of the $n$th Chebyshev polynomial.

There are <u>many</u> good reasons for this choice

- ✓ Guaranteed convergence[1] to $f(s)$ in large $n$ limit (no Runge)
- ✓ Robust to small errors[2] in data $\tilde{y}_i \approx f(s_i)$
- ✓ Polynomial fit accomplished with well-conditioned linear system
- ✓ Nodes anticluster from $s = 0 \Longrightarrow$ cheaper quantum cost




*By Steven G. Johnson (Wikipedia)*

1. For Lipschitz continuous *f*, see Trefethen (2011)
2. *The Chebyshev Polynomials*, Rivlin (1974)

# Theory of Interpolation Error

Using the generalized Mean Value Theorem, we can show

$$E(s) \coloneqq f(s) - P_{n-1}(s) = \frac{f^{(n)}(\xi)}{n!}\omega_n(s)$$

where $\xi \in [-a, a]$ and the nodal polynomial $\omega_n$ is

$$\omega_n(s) \coloneqq \prod_{j=1}^{n}(s - s_j).$$

We care about $s = 0$, Chebyshev nodes, and an upper bound.

From which we show that

$$|E(0)| \leq \max_{s\in[-a,a]}\left(|f^{(n)}(s)|\right)\left(\frac{a}{2n}\right)^n$$

A lot of the work is just upper bounding $f^{(n)}(s)$.

# How hard is this?

- For e.g., phase estimation, $f(s) = \tilde{E}_s = \langle \tilde{E}_s | \tilde{H}_s | \tilde{E}_s \rangle$

$\Rightarrow$Eigenvalue derivatives $f^{(n)}(s)$ found by repeated use of perturbation theory.

- Expectation values: just need derivatives of $e^{-i \tilde{H}_s t}$
- To evaluate and bound these derivatives involves
  - Combinatoric tools, such as Faà di Bruno's formula
  - Plentiful, but tasteful, use of triangle inequality.
- After a long slog, we get an error bound (hard part)
- Then we turn it into an algorithm cost (easier part)

# Sources of error

- Interpolation error
- Imperfect data points
- Imperfect Chebyshev nodes

Turns out to be subdominant, thanks to well-conditioning of Chebyshev!

# Main Results

# First: a crucial lemma on the size of $\widetilde{H}_s$ derivatives

$$H = \sum_{j=1}^{m} H_j \qquad \widetilde{H}_s := \frac{i}{st} \log S_{2k}(st)$$

**Lemma:**

Let $s$ be chosen (small enough) such that

$$2k(5/3)^{k-1} m \max_j \|H_j\| \, st \leq \pi/20\,.$$

Then the following bound holds.

$$\left\|\partial_s^n \widetilde{H}_s\right\| \leq 2t^{-1} n^n \left(2e^2 k (5/3)^{k-1} m \max_j \left\|H_j\right\| t\right)^{n+1}$$

Proof technique: expand logarithm as power series. Turn the crank.

# Main Result 1: Eigenvalue estimation

$$\tilde{H}_s|\tilde{E}_s\rangle = \tilde{E}_s|\tilde{E}_s\rangle$$

Theorem:

Let $s$ be chosen as small as in the previous lemma. Then it is possible to estimate

$$\tilde{E}_0 = E$$

Within precision $\epsilon$ and failure probability at most 1/3, using a number of $e^{-i\,H_j t}$ bounded as

$$N_{exp} \in \tilde{O}\left(\frac{m^2 (25/3)^k \max_j \|H_j\| (1+\Gamma)}{\epsilon}\right)$$

Proof technique, use previous lemma and perturbation theory.
One annoyance: $\Gamma$ depends on **inverse minimal spectral gap.**

# Main Result 2: Expectation value estimation

$$\langle \tilde{O}_s(t) \rangle = \mathrm{Tr}(\tilde{U}_s^\dagger(t) O\, \tilde{U}_s(t)\rho)$$

Theorem:

Again, with $s$ sufficiently small, it is possible to estimate

$$\langle \tilde{O}_0(t) \rangle = \langle O(t) \rangle$$

within precision $\epsilon$ and failure probability at most 1/3 using a number of $e^{-i\,H_j t}$ bounded as

$$N_{exp} \in \tilde{O}\left( \frac{m^3 k^2 (25/9)^{k-1} \left( \max_j \|H_j\| t \right)^2}{\epsilon} \right).$$

Drawback: $O(\,(\|H\|t)^2)$ scaling! Compare to optimal (and achieved) $O(t)$.
This is our bound for *any order* Suzuki Trotter.

# Discussion of Findings

- Our results demonstrate that near-Heisenberg Limited scaling is achievable with polynomial interpolation of Trotter data
- Sadly, Expectation Value algorithm suffers a suboptimal $t^2$ bound
  - Can this be improved? I would guess so.
- Our work is limited in several respects
  - Hardware noise
  - State prep
  - Only Suzuki Trotter, not generic Trotter (but should apply generally)
- Our approach is relatively NISQ friendly, in that it only requires Trotter
- Connections to zero noise extrapolation?
- Connections to lattice QFT?

# Outlook

- Using *only Trotter* + *classical,* we achieve some accuracy gains in important estimation tasks:
  - Eigenvalue estimation
  - Expectation value estimation
- As Trotter simulations become more feasible, we can begin to test polynomial interpolation on real hardware.
- *Broader question: How do we best take advantage of simulation data once we have it?*