

Quantum Krylov Algorithms

Will Kirby

IBM Quantum

william.kirby@ibm.com

4/21/26

Working definition of quantum Krylov algorithms

1. Quantum: calculate $x_k \equiv \langle \psi | U^k | \psi \rangle$, $k = 1, 2, \dots, d-1$

- sometimes also $h_{jk} \equiv \langle \psi | U^{-j} H U^k | \psi \rangle$ in cases where $U = e^{-iH \Delta t}$

2. Classical: arrange results as

$$\mathbf{S} = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & \cdots & x_{d-1} \\ \overline{x_1} & 1 & x_1 & x_2 & \cdots & x_{d-2} \\ \overline{x_2} & \overline{x_1} & 1 & x_1 & \cdots & x_{d-3} \\ \overline{x_3} & \overline{x_2} & \overline{x_1} & 1 & \cdots & x_{d-4} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{x_{d-1}} & \overline{x_{d-2}} & \overline{x_{d-3}} & \overline{x_{d-4}} & \cdots & 1 \end{pmatrix}$$

Why? $\mathbf{S}_{jk} = \langle \psi_j | \psi_k \rangle$ for $|\psi_j\rangle \equiv U^j |\psi\rangle$

3. Then, do something classically with $\mathbf{S}$, potentially also $\mathbf{H} \equiv [h_{jk}]$

Why is this quantum Krylov?

Krylov space: $\text{span}[\underbrace{|\psi_0\rangle, U|\psi_0\rangle, U^2|\psi_0\rangle, \dots, U^{d-1}|\psi_0\rangle}_V]$ for some U

- $\mathbf{S} \equiv [\langle\psi_j|\psi_k\rangle]$ is the metric of this space

Quantum: U is most likely unitary

$(\mathbf{H}, \mathbf{S})$ = project H onto basis V

$$\boxed{V^\dagger} \boxed{\hat{H}} \boxed{V} = \boxed{\mathbf{H}}, \quad \boxed{V^\dagger} \boxed{V} = \boxed{\mathbf{S}}$$

Example #1: ground state energies

Final classical step is solving

$$\mathbf{H}\mathbf{v} = \lambda\mathbf{S}\mathbf{v}^1 \text{ or}$$

$$\mathbf{U}\mathbf{v} = \lambda'\mathbf{S}\mathbf{v} \text{ with } \mathbf{U}_{jk} = \mathbf{S}_{j,k+1}^2$$

then return $\min\{\lambda\}$ or $\min\left\{\frac{i \arg \lambda'}{\Delta t}\right\}$

WHY?

1. Converges fast w.r.t. d , dimension of $\mathbf{S}, \mathbf{H}$
2. Nice properties w.r.t. noise

¹Parrish & McMahon, 1909.08925; ²Klymko *et al.*, 2103.08563

Quantum Krylov diagonalization with errors

$$\text{Ideal } (\mathbf{H}, \mathbf{S}) \xrightarrow{\text{errors}} (\mathbf{H}', \mathbf{S}')$$

Worry #1: ill-conditioning.

- $\mathbf{S}$ has exponentially-growing condition number with D
- $\Rightarrow \mathbf{S}'$ may not even be positive semidefinite.

Resolution: regularization.¹

- Project $(\mathbf{H}', \mathbf{S}')$ onto eigenspaces of $\mathbf{S}'$ with eigenvalues above threshold $\epsilon > 0$.
- Epperly *et al.* showed this step introduces energy error linear in ϵ .
- Epperly *et al.* also bounded total error in lowest energy estimate in presence of noise η ...
- **Goal:** but bound is sublinear in η . Can we improve to linear?

¹Epperly, Lin, Nakatsukasa, 2110.07492

Attempt to improve (partly successful)

$$\text{Ideal } (\mathbf{H}, \mathbf{S}) \xrightarrow{\text{errors}} (\mathbf{H}', \mathbf{S}')$$

Inspiration: common “folklore” claim in quantum computing community...

“The reason quantum Krylov algorithms are robust to errors is that the errors just perturb the subspace, and then you still find the lowest energy in the subspace.”

In other words, for $\mathbf{H} = \mathbf{V}^\dagger H \mathbf{V}$, $\mathbf{S} = \mathbf{V}^\dagger \mathbf{V}$, errors cause $\mathbf{V} \rightarrow \mathbf{V}'$.

This is not true in general!

Main technical result of interest ¹

• generic error-ful $(\mathbf{H}', \mathbf{S}')$ after regularization can be expressed as $(\mathbf{V}'^\dagger H' \mathbf{V}', \mathbf{V}'^\dagger \mathbf{V}')$, such that

1. we can bound $\|H' - H\|$ in terms of the error rate
2. span $\mathbf{V}'$ contains a perturbed approximate ground state of H

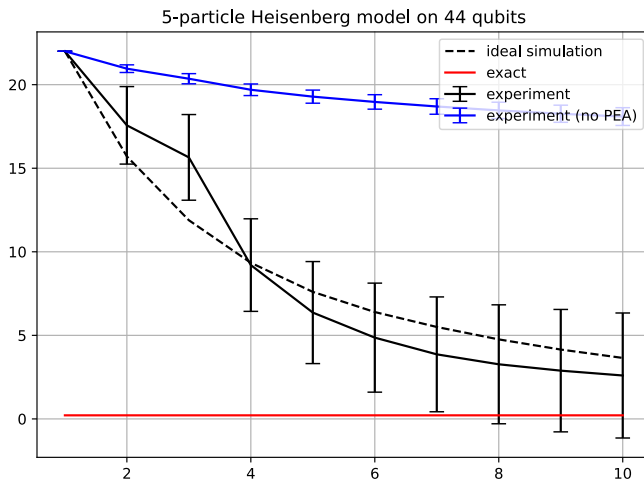
¹WK, 2401.01246

Resulting bounds

$$-O\left(\frac{\|H\|\eta}{\epsilon}\right) \leq \widetilde{E}_0 - E_0 \leq \frac{\|H\|}{|\gamma_0|^2} \cdot O\left(\left(\frac{1}{\Delta} + d\right)\eta + d\epsilon + (1 + \beta)^{-d}\right)$$

where:

- $\gamma_0 \equiv \langle \psi | E_0 \rangle$ is initial state overlap
- $\Delta \equiv E_1 - E_0$ is spectral gap
- $\eta \equiv \Theta\left(\frac{\|\mathbf{H}' - \mathbf{H}\|}{\|H\|} + \|\mathbf{S}' - \mathbf{S}\|\right)$ is catch-all noise rate
- $\epsilon \equiv$ regularization threshold
- $\beta \equiv \Theta\left(\frac{\Delta}{\|H\|}\right)$



<https://arxiv.org/abs/2407.14431>

Thoughts on partial success

- Lower and upper bounds superficially look linear in η ...
- BUT lower bound is $O\left(\frac{\eta}{\epsilon}\right)$ and upper bound is $O(\eta + \epsilon)$.
 - Quantifies idea that threshold is trading off between upper and lower bounds.
 - $\Rightarrow$ best symmetric bound is $O(\eta^{1/2})$ – worse than Epperly *et al.*'s $O(\eta^{2/3})$
- In practice $\epsilon = O(\eta)$ is often found to yield good convergence, leading to open questions:
 - How generally true is that?
 - Is there a tighter lower bound? (Since mine is $O(1)$ in that case!)

Example #2: Quantum Szegő quadrature (QSQ)

Goal: estimate quantities of the form

$$\langle \psi | f(U) | \psi \rangle$$

on a quantum computer.

Inputs for quantum algorithm:

$|\psi\rangle$ = some efficiently preparable state

U = some efficiently implementable unitary

Inputs for classical post-processing:

f = some function

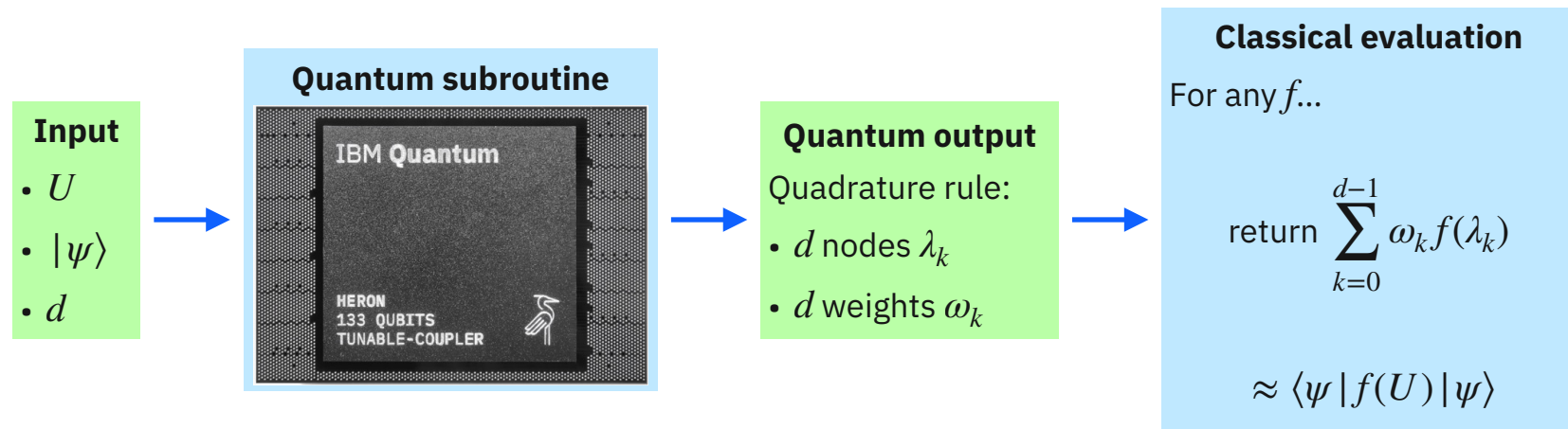
Can extend to

$$\langle \psi | f(U) | \psi' \rangle$$

using exp vals of linear combinations of $|\psi\rangle, |\psi'\rangle$

WK, Yizhi Shen, Daan Camps, Anirban Chowdhury, Katherine Klymko, Roel Van Beeumen, arxiv:2509.19195

Algorithm sketch



Why “quadrature rule”?

Generic quadrature rule: $\int_{\chi} f(x) d\mu(x) \approx \sum_{k=0}^{d-1} \omega_k f(\lambda_k)$

“Optimal” quadrature rules:

- For $\chi = \text{real interval}$, “Gauss quadrature”: d nodes $\Rightarrow$ exact for polynomials up to degree $2d - 1$
- For $\chi = \mathbb{T} = \text{unit circle}$, “Szegő quadrature”: d nodes $\Rightarrow$ exact for Laurent polynomials up to degree $d - 1$


$$f(x) = \sum_{k=-d+1}^{d-1} \alpha_k x^k$$

Where does the integral come from?

Let $|\psi\rangle = \sum_{j=0}^{N-1} \gamma_j |\zeta_j\rangle$ for $|\zeta_j\rangle$ eigenvectors of U

Then $\langle\psi|f(U)|\psi\rangle = \sum_{j=0}^{N-1} |\gamma_j|^2 f(\zeta_j) = \int_{\mathbb{T}} f(z) d\mu(z)$ for $d\mu(z) = \sum_{j=0}^{N-1} |\gamma_j|^2 \delta(z - \zeta_j)$

So idea of “quantum Szegő quadrature,” is to approximate $\langle\psi|f(U)|\psi\rangle$
by approximating above integral using a Szegő quadrature rule,
nodes and weights calculated using quantum computer.

Quantum subroutine: quantum Arnoldi's method

- Ideally, would like Arnoldi's method on $U, |\psi\rangle$
- **Problem:** orthonormalization expensive on quantum
- **Resolution:** don't orthonormalize — form matrix pencil $(\mathbf{U}, \mathbf{S})$ projecting U into Krylov space

$$\mathbf{U} = V^\dagger U V, \quad \mathbf{S} = V^\dagger V \quad \text{for } V = [|\psi\rangle, U|\psi\rangle, U^2|\psi\rangle, \dots, U^{d-1}|\psi\rangle]$$

1. Approximate $(\mathbf{U}, \mathbf{S})$ on quantum computer using Hadamard test or equivalent
2. Classically Gram-Schmidt orthonormalize basis for $\mathbf{U}$ using Gram matrix $\mathbf{S}$ *
3. Formally, result is same upper-Hessenberg $\mathcal{H}$ as from Arnoldi

*Requires regularization

Quantum Arnoldi's method* part 2

Quantum
approximate $(\mathbf{U}, \mathbf{S})$



Regularize $\mathbf{S}$

$$\mathbf{S} \rightarrow \mathbf{S} + \eta$$



Orthonormalize $\mathbf{U}$

for Gram matrix $\mathbf{S}$



$\mathcal{H}$

Resulting $d \times d \mathcal{H}$ “approximates” $2^n \times 2^n U$
— extremal eigenvalues converge — but $\mathcal{H}$ is
only approximately unitary



Project $\mathcal{H}$ to nearest unitary

1. $P^\dagger D Q = \mathcal{H}$ (SVD)

2. $\mathcal{H} \leftarrow P^\dagger Q$



Resulting $d \times d \mathcal{H}$ “approximates” $2^n \times 2^n U$
— extremal eigenvalues converge — and is
unitary — eigenvalues lie in $\mathbb{T}$

* “isometric” Arnoldi’s method

Quantum Szegő Quadrature

- From slide 12: $\langle \psi | f(U) | \psi \rangle = \int_{\mathbb{T}} f(z) d\mu(z)$
- From slide 14: Quantum isometric Arnoldi $\Rightarrow d \times d \mathcal{H}$ with eigenvalues in $\mathbb{T}$ (complex unit circle)
- Let $\lambda_k, v_k =$ eigenvalues, eigenvectors of $\mathcal{H}^T$

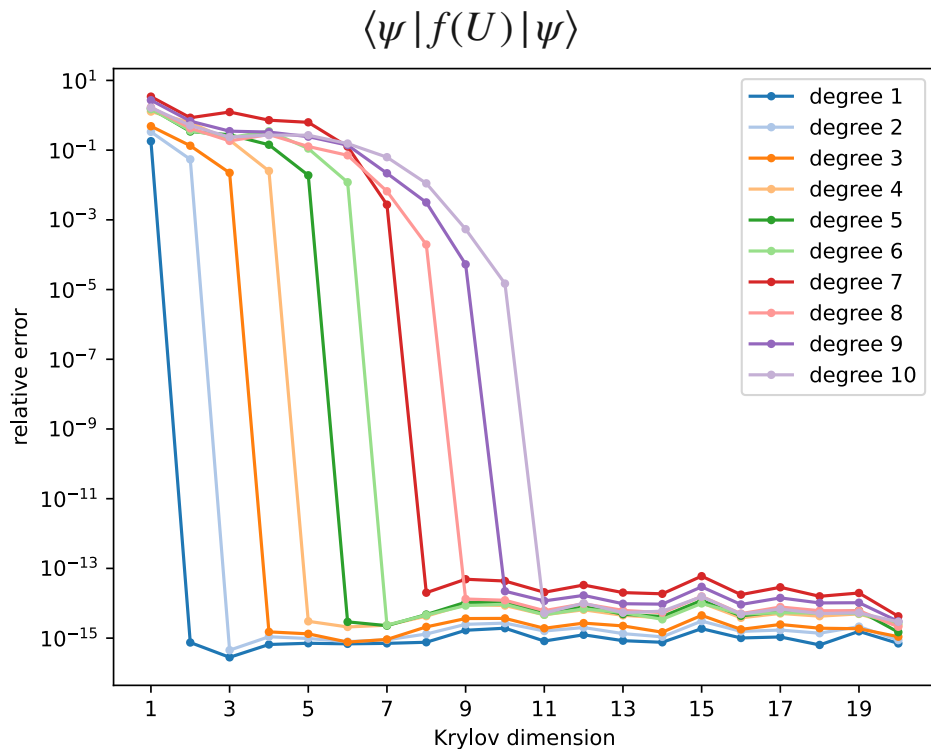
Can prove: a Szegő quadrature rule for the integral above is

$$\int_{\mathbb{T}} f(z) d\mu(z) = \sum_{k=0}^{d-1} |v_{k,0}|^2 f(\lambda_k)$$



zeroth entry in v_k

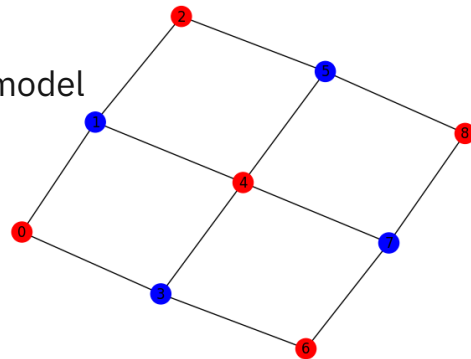
Numerical illustrations (polynomials)



- System: Heisenberg model

- State:

$$|\psi\rangle =$$



- Classically simulate quantum Szegő quadrature
- f = random Laurent polynomials
- Points in plot = mean errors over 10 instances

Beyond polynomials

Theorem: Let $\mathbb{L}_d$ = degree- d Laurent polynomials. If

$$\varepsilon = \min_{p \in \mathbb{L}_{d-1}} \max_{z \in \mathbb{T}} |f(z) - p(z)|$$

then (in the absence of noise) QSQ has error at most 2ε .

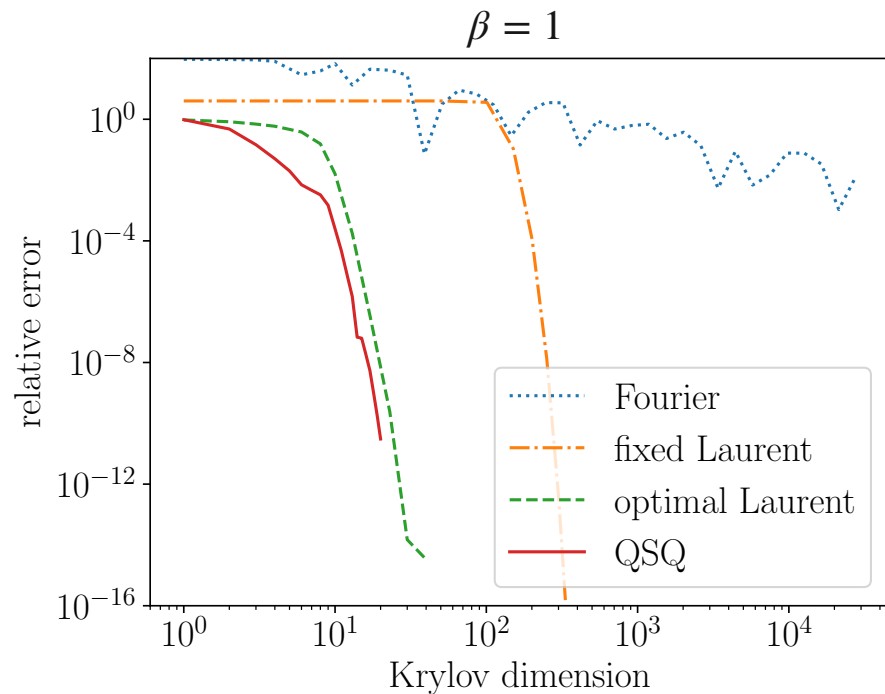
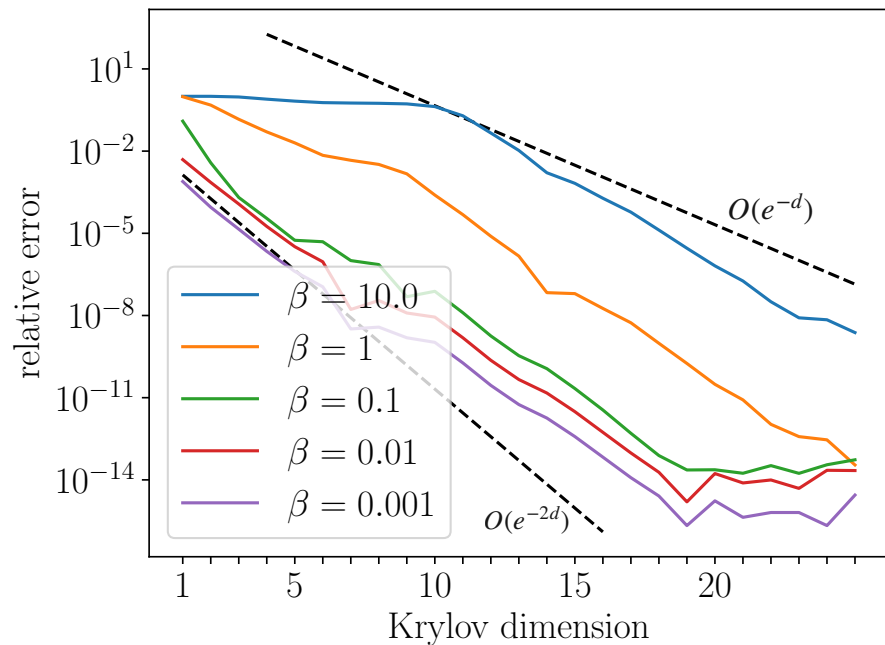
Remarks:

- If we knew optimal $p^*(z) = \sum_{j=-d+1}^{d-1} \alpha_j z^j$, we could use same data to construct it: $\langle \psi_0 | p^*(U) | \psi_0 \rangle = \sum_{j=-d+1}^{d-1} \alpha_j x_j$

 $x_{-j} \equiv \bar{x}_j$
- This would have error at most ε
- Either one might luckily have error zero $\Rightarrow$ not possible to relate directly, hence pointwise ε above

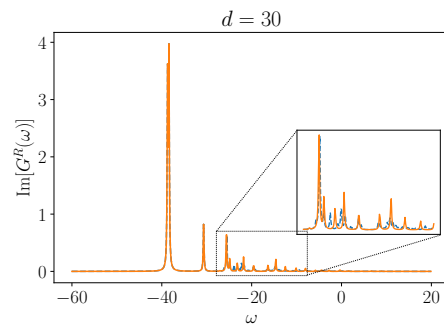
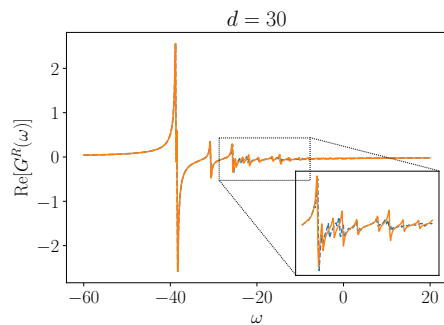
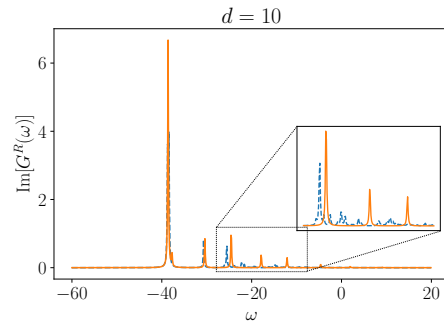
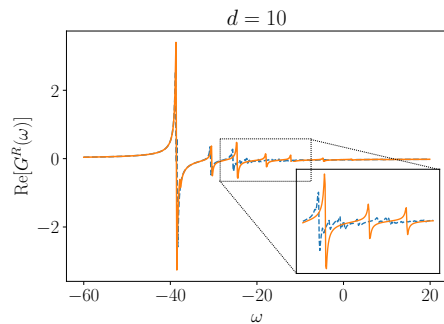
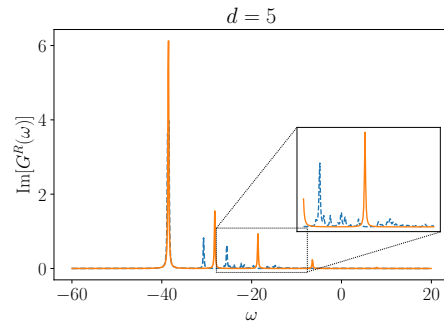
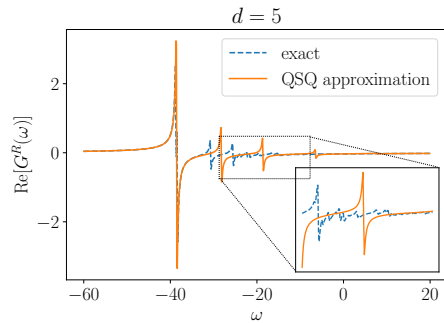
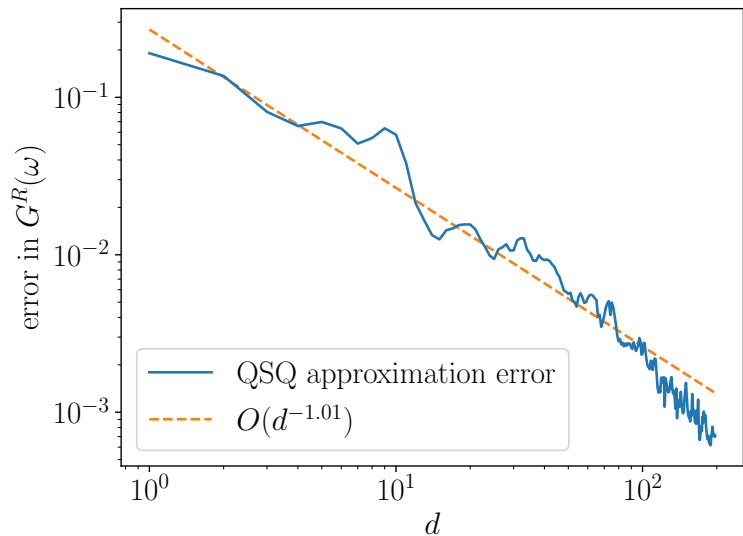
Gibbs state

$$f(U) = U^{i\beta/\Delta t} = e^{-\beta H} \quad (U = e^{-iH\Delta t})$$



Green's function

$$G^R(\omega) = \langle \psi_0 | \frac{1}{H - \omega - i\chi} | \psi_0 \rangle$$



How to think about this tool?

Uses same data as other quantum Krylov algorithms

Like quantum signal processing or QET-U, evaluates functions of operators on Hilbert space

Upsides:

- Does not require knowing function beforehand
 - Can calculate quadrature rule, then use for any function
- Does not require calculating explicit Laurent series
- Noise robust

Downside: no error analysis with regularization and noise

Thank you!

Questions?



Proof sketch (Szegő quadrature)

1. Define $\vec{p}(x) = (p_0(x), p_1(x), \dots, p_{d-1}(x))$ such that orthonormal Krylov basis = $\{p_i(U)v \mid i = 0, \dots, d-1\}$
2. Define *para-orthogonal polynomials* $\tilde{p}_j(x) = p_j(x) + \tau x^j \bar{p}_j(x^{-1})$, $|\tau| = 1$
3. Eigenvalues λ_k of $\mathcal{H}^T$ are zeros of $\tilde{p}_d$, and $\vec{p}(\lambda_k)$ are corresponding eigenvectors *
4. Zeros of $\tilde{p}_d$ are nodes of Szegő quadrature rule *
5. Weights are $\omega_k = \frac{1}{\|\vec{p}(\lambda_k)\|^2}$ (following **, two applications of Christoffel-Darboux for complex unit circle)
6. If v_k is normalized eigenvector of $\mathcal{H}^T$, $\|\vec{\phi}(\lambda_k)\| = \frac{\|\vec{\phi}(\lambda_k)\|}{\|v_k\|} = \frac{|\phi_0(\lambda_k)|}{|v_{k,0}|} = \frac{1}{|v_{k,0}|}$

* W. B. Jones, O. Njåstad, and W. J. Thron, Bulletin of the London Mathematical Society, vol. 21, no. 2, pp. 113–152, 1989.

* W. B. Gragg, Journal of Computational and Applied Mathematics, vol. 46, no. 1, pp. 183–198, 1993.

** Section 2.9 in H. S. Wilf, Mathematics for the Physical Sciences Dover Publications Inc., New York, NY, 1962.

Main ideas for lower bound

Idea #1: work with regularized pair $(\mathbf{H}'', \mathbf{S}'') = (\Pi' \mathbf{H}' \Pi', \Pi' \mathbf{S}' \Pi')$

- $\Pi' =$ projector onto eigenspaces of $\mathbf{S}'$ with eigenvalues above $\epsilon > 0$
- $(\mathbf{H}'', \mathbf{S}'') =$ matrix pair we actually solve to get approximate energies (removing the “projected out” dim’ns)

Main ideas for lower bound

Idea #2: choose a convenient $\mathbf{V}'$ such that $\mathbf{V}'^\dagger \mathbf{V}' = \mathbf{S}''$

- Let $\mathbf{V} = F\sqrt{\mathbf{S}}$ be polar decomposition of $\mathbf{V}$, so F is $2^n \times d$ with orthonormal columns
- Let $\mathbf{V}' = FG\sqrt{\mathbf{S}''}$, with G any $d \times d$ unitary $\Rightarrow \mathbf{V}'^\dagger \mathbf{V}' = \sqrt{\mathbf{S}''}G^\dagger F^\dagger FG\sqrt{\mathbf{S}''} = \mathbf{S}''$

Main ideas for lower bound

Idea #3: choose H' such that $\mathbf{V}'^\dagger H' \mathbf{V}' = \mathbf{H}''$, and $H' = H$ everywhere else (outside $\text{span}(\mathbf{V}')$)

- Corresponding expression: $H' = H + \mathbf{V}' \mathbf{S}''^\dagger (\mathbf{H}' - \mathbf{V}'^\dagger H \mathbf{V}') \mathbf{S}'' + \mathbf{V}'^\dagger$
- Use freedom in G to minimize $\|H' - H\|$
- Good choice turns out to be such that $G\sqrt{\Pi' \mathbf{S} \Pi'} = \text{polar decomp of } \sqrt{\mathbf{S} \Pi'}$

Aside for anyone actually trying to follow the above: $\mathbf{S}''^\dagger \mathbf{V}'^\dagger \cdot \mathbf{V}' = \mathbf{S}''^\dagger \mathbf{S}'' = \Pi'$ and $\Pi' \mathbf{H}' \Pi' = \mathbf{H}''$

Main ideas for lower bound

Proof of lower bound proceeds from

$$\begin{aligned} \|\hat{H}' - \hat{H}\| &= \|\mathbf{V}' \mathbf{S}''^+ (\mathbf{H}' - \mathbf{V}'^\dagger \hat{H} \mathbf{V}') \mathbf{S}''^+ \mathbf{V}'^\dagger\| = \|\sqrt{\mathbf{S}''^+} (\mathbf{H}' - \mathbf{V}'^\dagger \hat{H} \mathbf{V}') \sqrt{\mathbf{S}''^+}\| \\ &\leq \|\sqrt{\mathbf{S}''^+} (\mathbf{H}' - \mathbf{H}) \sqrt{\mathbf{S}''^+}\| + \|\sqrt{\mathbf{S}''^+} \mathbf{V}'^\dagger \hat{H} (\mathbf{V} - \mathbf{V}') \sqrt{\mathbf{S}''^+}\| + \|\sqrt{\mathbf{S}''^+} (\mathbf{V}^\dagger - \mathbf{V}'^\dagger) \hat{H} \mathbf{V}' \sqrt{\mathbf{S}''^+}\| \end{aligned}$$

Bounding each term yields



$$\|\hat{H}' - \hat{H}\| \leq \frac{\|\mathbf{H}' - \mathbf{H}\| + (\sqrt{2} + 1)\|\mathbf{S}' - \mathbf{S}\|\|\hat{H}\|}{\epsilon}$$

Choice of “free” unitary G only matters in second term.

Main ideas for lower bound

$$\|\hat{H}' - \hat{H}\| \leq \frac{\|\mathbf{H}' - \mathbf{H}\| + (\sqrt{2} + 1)\|\mathbf{S}' - \mathbf{S}\|\|\hat{H}\|}{\epsilon}$$

- Weyl's theorem $\Rightarrow$ above upper bounds difference between lowest eigenvalues E'_0 and E_0 of $\hat{H}'$ and $\hat{H}$
- Rayleigh-Ritz $\Rightarrow E'_0$ lower bounds lowest eigenvalue of $(\mathbf{H}'', \mathbf{S}'') = (\mathbf{V}^\dagger \hat{H}' \mathbf{V}', \mathbf{V}^\dagger \mathbf{V}')$
- Combine: $-\|\hat{H}' - \hat{H}\|$ lower bounds ground state energy error of noisy, regularized matrix pair $(\mathbf{H}'', \mathbf{S}'')$ with respect to E_0 , which we are trying to estimate

Upper bound rough ideas

Similar scheme to lower bound:

- work with $(\mathbf{H}'', \mathbf{S}'')$
- Now will use particular point c' in Krylov space to upper bound lowest energy:
 - so, choose convenient $\mathbf{V}', \hat{H}'$ such that $c'^{\dagger} \mathbf{V}'^{\dagger} \mathbf{V}' c' = c'^{\dagger} \mathbf{S}'' c'$, $c'^{\dagger} \mathbf{V}'^{\dagger} \hat{H}' \mathbf{V}' c' = c'^{\dagger} \mathbf{H}'' c'$
- Prove bound on $\|\hat{H}' - \hat{H}\|$ and also show that $\mathbf{V}' c'$ is an approximate ground state
 - Use similar method to Epperly, Lin, and Nakatsukasa (2022).
- Result:

$$E'_0 - E_0 \leq \frac{12\|\hat{H}\|}{|\gamma'_0|^2} \left[\left(D + \frac{\|\hat{H}\|}{\Delta'} + \frac{1}{12} \right) \eta + D\epsilon + 4 \left(1 + \frac{\pi\Delta'}{4\|\hat{H}\|} \right)^{-2D} \right]$$

for $\eta = \max \left(\|\mathbf{S}' - \mathbf{S}\|, \frac{\|\mathbf{H}' - \mathbf{H}\|}{\|\hat{H}\|} \right)$, $\gamma'_0 \approx$ ampl. of ground state in init state, $\Delta' \approx$ spectral gap