

Supplementary information: Randomized semi-quantum matrix processing

Allan Tosta^{1,2}, Thais de Lima Silva¹, Giancarlo Camilo¹, and Leandro Aolita^{1,2}

¹Quantum Research Center, Technology Innovation Institute, Abu Dhabi, UAE

²Federal University of Rio de Janeiro, Caixa Postal 68528, Rio de Janeiro, RJ, Brazil

Supplementary note 1: General framework

1.1 Proof of Lemma 4

Proof. Let U_A be given as in Def. 1. Then

$$U_A^j = \bigoplus_{\gamma} (e^{-i\vartheta_{\gamma} Y_{\gamma}})^j = \bigoplus_{\gamma} \begin{bmatrix} \cos(j \arccos \lambda_{\gamma}) & -\sin(j \arccos \lambda_{\gamma}) \\ \sin(j \arccos \lambda_{\gamma}) & \cos(j \arccos \lambda_{\gamma}) \end{bmatrix}. \quad (1)$$

Since $\cos(j \arccos x) = \mathcal{T}_j(x)$, where $\mathcal{T}_j$ is the j -th Chebyshev polynomial of the first kind, it follows that

$$(\langle 0|_a \otimes \mathbb{1}_s) U_A^j (|0\rangle_a \otimes \mathbb{1}_s) = \bigoplus_{\gamma} \mathcal{T}_j(\lambda_{\gamma}) |\lambda_{\gamma}\rangle\langle\lambda_{\gamma}|_s = \mathcal{T}_j(A). \quad (2)$$

□

1.2 Hadamard test for block encodings

Here we prove Lemmas 5 and 6. Moreover, in Lemma 1 below we show how to extend the construction to Chebyshev polynomials of the second kind. Throughout this section, we use the notation cU to refer to a gate U controlled by the state $|1\rangle$ of the ancilla 1 (the top register in both circuits of Fig. 1, i.e., the Hadamard qubit ancilla). We use $\bar{c}U$ for when the action of the unitary U is controlled by the ancilla 1 being in state $|0\rangle$.

Proof of Lemma 5. Here we prove it for the case $B = \mathbb{1}$, since the other follows in an analogous way. We have that,

$$\mathbb{E}[\mathbf{Had}(U_A^j, |\phi\rangle, |\psi\rangle)] = \text{Tr}\{X_1 G(|+\rangle\langle+|_1 \otimes \varrho_{a,s}) G^{\dagger}\}, \quad (3)$$

with $G = \bar{c}U_{\phi} cU_A^j cU_{\psi}$ and $\varrho_{a,s} = |0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_s$. Computing the expression inside the trace we get

$$\begin{aligned} X_1 G(|+\rangle\langle+|_1 \varrho_{a,s}) G^{\dagger} &= \frac{1}{2} |0\rangle\langle 0|_1 (U_A^j U_{\psi} \varrho_{a,s} U_{\phi}^{\dagger}) + \frac{1}{2} |1\rangle\langle 1|_1 (U_{\phi} \varrho_{a,s} U_{\psi}^{\dagger} U_A^{j\dagger}) + \frac{1}{2} |1\rangle\langle 0|_1 (U_{\phi} \varrho_{a,s} U_{\phi}^{\dagger}) \\ &\quad + \frac{1}{2} |0\rangle\langle 1|_1 (U_A^j U_{\psi} \varrho_{a,s} U_{\psi}^{\dagger} U_A^{j\dagger}). \end{aligned} \quad (4)$$

Therefore,

$$\mathbb{E}[\mathbf{Had}(U_A^j, |\phi\rangle, |\psi\rangle)] = \text{Re}\{\text{Tr}_{a,s}[U_A^j |0\rangle\langle 0|_a |\psi\rangle\langle\phi|_s]\} = \text{Re}\{\text{Tr}_s[\mathcal{T}_j(A) |\psi\rangle\langle\phi|_s]\} = \text{Re}\{\langle\phi|\mathcal{T}_j(A)|\psi\rangle\} \quad (5)$$

□

Proof of Lemma 6. We have that

$$\mathbb{E}[\mathbf{Had}(U_A^j, U_A^l, O)] = \text{Tr}\{OX_1 M(|+\rangle\langle+|_1 \otimes \varrho_{a,s}) M^{\dagger}\}, \quad (6)$$

where $M = \bar{c}U_A^l cU_A^j$ and $\varrho_{a,s} = |0\rangle\langle 0|_a \otimes \varrho_s$. Note that

$$\begin{aligned} X_1 M(|+\rangle\langle+|_1 \varrho_{a,s}) M^{\dagger} &= \frac{1}{2} |0\rangle\langle 0|_1 (U_A^j \varrho_{a,s} U_A^{l\dagger}) + \frac{1}{2} |1\rangle\langle 1|_1 (U_A^l \varrho_{a,s} U_A^{j\dagger}) + \frac{1}{2} |1\rangle\langle 0|_1 (U_A^l \varrho_{a,s} U_A^{l\dagger}) + \frac{1}{2} |0\rangle\langle 1|_1 (U_A^j \varrho_{a,s} U_A^{j\dagger}), \end{aligned} \quad (7)$$

which implies

$$\begin{aligned}\mathbb{E}[\mathbf{Had}(U_A^j, U_A^l, O)] &= \frac{1}{2} \text{Tr}_{a,s}[O(U_A^j \varrho_{a,s} U_A^{l\dagger})] + \frac{1}{2} \text{Tr}_{a,s}[O(U_A^l \varrho_{a,s} U_A^{j\dagger})] \\ &= \frac{1}{2} \text{Tr}_s[O(\mathcal{T}_j(A) \varrho_s \mathcal{T}_l^\dagger(A))] + \frac{1}{2} \text{Tr}_s[O(\mathcal{T}_j(A) \varrho_s \mathcal{T}_l^\dagger(A))] \\ &= \text{Tr}_s[O(\mathcal{T}_j(A) \varrho_s \mathcal{T}_l(A))],\end{aligned}\tag{8}$$

since A and O are Hermitian. $\square$

Supplementary Lemma 1. *If we modify the Hadamard test in Fig. 1.a) by applying a $\bar{\text{cNOT}}_{1,a}$ after any other gate, the random variable $\mathbf{Had}_{alt}(U_A^j, |\phi\rangle, |\psi\rangle)$ satisfies for $B = \mathbb{1}$*

$$\mathbb{E}[\mathbf{Had}_{alt}(U_A^j, |\phi\rangle, |\psi\rangle)] = \text{Re}\{\langle\phi|\sqrt{\mathbb{1} - A^2} \mathcal{U}_{j-1}(A)|\psi\rangle\}\tag{9}$$

and

$$\mathbb{E}[\mathbf{Had}_{alt}(U_A^j, |\phi\rangle, |\psi\rangle)] = \text{Im}\{\langle\phi|\sqrt{\mathbb{1} - A^2} \mathcal{U}_{j-1}(A)|\psi\rangle\}\tag{10}$$

for $B = S^\dagger$.

Proof. Again, here we only show the case $B = \mathbb{1}$ since the other case is analogous. From the description of the circuit in Fig. 1.a) and the added modification we have that

$$\mathbb{E}[\mathbf{Had}_{alt}(U_A^j, |\phi\rangle, |\psi\rangle)] = \text{Tr}[X_1 G_{alt}(|+\rangle\langle+|_1 \varrho_{a,s}) G_{alt}^\dagger]\tag{11}$$

with $G_{alt} = \bar{\text{cNOT}}_{1,a} \bar{c}U_\phi cU_A^j cU_\psi$ and $\varrho_{a,s} = |0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_s$. We have that

$$X_1 G_{alt}(|+\rangle\langle+|_1 \varrho_{a,s}) G_{alt}^\dagger = \frac{|1\rangle\langle 0|_1 |1\rangle\langle 1|_a |\phi\rangle\langle\phi|_s}{2} + \frac{|0\rangle\langle 0|_1 U_A^j (|0\rangle\langle 1|_a |\psi\rangle\langle\phi|_s)}{2} + \frac{|1\rangle\langle 1|_1 (|1\rangle\langle 0|_a |\phi\rangle\langle\psi|_s) U_A^{j\dagger}}{2} +\tag{12}$$

$$+ \frac{|0\rangle\langle 1|_1 U_A^j (|0\rangle\langle 0|_a |\psi\rangle\langle\psi|_s) U_A^{j\dagger}}{2},\tag{13}$$

from which we obtain

$$\text{Tr}[X_1 G_{alt}(|+\rangle\langle+|_1 \varrho_{a,s}) G_{alt}^\dagger] = \text{Re}\{\langle\phi|(1|U_A^j|0\rangle_a)|\psi\rangle_s\}.\tag{14}$$

From Eq. (1) and the trigonometric definition of the Chebyshev polynomial $\mathcal{U}_{j-1}(x)$ we have that

$$(|1\rangle_a \otimes \mathbb{1}_s) U_A^j (|0\rangle_a \otimes \mathbb{1}_s) = \sqrt{\mathbb{1} - A^2} \mathcal{U}_{j-1}(A),\tag{15}$$

proving the statement. $\square$

1.3 Proof of Theorem 1

Proof. We provide different proofs for each P . For $P = 1$, note that

$$\langle\phi|\tilde{f}(A)|\psi\rangle = \sum_{j=0}^k a_j \langle\phi|\mathcal{T}_j(A)|\psi\rangle,\tag{16}$$

and from Lemma 5 for $B = \mathbb{1}$,

$$\text{Re}\{\langle\phi|\tilde{f}(A)|\psi\rangle\} = \sum_{j=0}^k a_j \mathbb{E}[\mathbf{Had}(U_A^j, |\phi\rangle, |\psi\rangle)] = \sum_{j=0}^k \sum_{b \in \{-1,1\}} p_j p(b|j) \|\mathbf{a}\|_1 \text{sgn}(a_j) b.\tag{17}$$

The last line implies that, for $\alpha < S^{(1)}$, $\frac{1}{S^{(1)}} \sum_\alpha \|\mathbf{a}\|_1 \text{sgn}(a_{j_\alpha}) b_\alpha^{(1)}$ with j_α sampled from the distribution p_j and $b_\alpha^{(1)}$ sampled from the conditional distribution $p(b|j)$ by running the Hadamard test in Figure 1a) is an unbiased estimator of $\text{Re}\{\langle\phi|\tilde{f}(A)|\psi\rangle\}$. The case for $B = S^\dagger$ is essentially the same, implying that for $S^1 \geq \alpha < 2S^{(1)} - 1$, $\frac{1}{S^{(1)}} \sum_\alpha \|\mathbf{a}\|_1 \text{sgn}(a_{j_\alpha}) b_\alpha^{(1)}$ is an unbiased estimator of $\text{Im}\{\langle\phi|\tilde{f}(A)|\psi\rangle\}$.

Similarly, for $P = 2$ note that

$$\text{Tr}[O\tilde{f}(A)\varrho\tilde{f}^\dagger(A)] = \sum_{j,l=0}^k a_j a_l \text{Tr}[O\mathcal{T}_j(A)\varrho\mathcal{T}_l(A)],\tag{18}$$

which by Lemma 6 implies that

$$\text{Tr}\left[O\tilde{f}(A)\varrho\tilde{f}^\dagger(A)\right] = \sum_{j,l=0}^k a_j a_l \mathbb{E}\left[\mathbf{Had}(U_A^j, U_A^l, O)\right] = \sum_{j,l=0}^k \sum_b \sum_\omega p_j p_l p(b, \omega|j, l) \|\mathbf{a}\|_1^2 \text{sgn}(a_j a_l) b \omega, \quad (19)$$

where $b \in \{-1, 1\}$ and ω in the spectrum of O . The last line implies that $\frac{1}{S^{(2)}} \sum_\alpha \|\mathbf{a}\|_1^2 \text{sgn}(a_{j_\alpha}) \text{sgn}(a_{l_\alpha}) b_\alpha^{(2)} \omega_\alpha$ with j_α, l_α both sampled from the distribution p_j and $b_\alpha^{(2)}, \omega_\alpha$ sampled from the conditional distribution $p(b, \omega|j, l)$ by running the Hadamard test in Figure 1b) is an unbiased estimator of $\text{Tr}\left[O\tilde{f}(A)\varrho\tilde{f}^\dagger(A)\right]$. $\square$

1.4 Proof of Theorem 2

Now we prove Theorem 2.

Proof. Let $\tilde{f}$ be a ν -approximation to the function f . Then, for any two states $|\psi\rangle, |\phi\rangle$ we have that

$$\left| \langle \phi | f(A) | \psi \rangle - \langle \phi | \tilde{f}(A) | \psi \rangle \right| \leq \|f(A) - \tilde{f}(A)\|_{op} \leq \nu. \quad (20)$$

Here, $\|\cdot\|_{op}$ is the operator norm, which for Hermitian operators coincides with the spectral norm $\|\cdot\|$. From Theorem 1 we have $\mathbb{E}[\tilde{z}^{(1)}] = \langle \phi | \tilde{f}(A) | \psi \rangle$ and thus

$$\begin{aligned} \left| z^{(1)} - \tilde{z}^{(1)} \right| &\leq \left| z^{(1)} - \mathbb{E}[\tilde{z}^{(1)}] \right| + \left| \mathbb{E}[\tilde{z}^{(1)}] - \tilde{z}^{(1)} \right| \\ &= \nu + \left| \tilde{z}^{(1)} - \mathbb{E}[\tilde{z}^{(1)}] \right|, \end{aligned} \quad (21)$$

$\tilde{z}^{(1)}$ will be an ϵ -precise estimator of $z^{(1)}$ if and only if

$$\left| \tilde{z}^{(1)} - \mathbb{E}[\tilde{z}^{(1)}] \right| \leq \epsilon - \nu^{(1)}, \quad (22)$$

where $\nu^{(1)} = \nu$. Notice from the definition of the estimator $\tilde{z}^{(1)}$ that the range of its real and imaginary parts each is $\{-\|\mathbf{a}\|_1, \|\mathbf{a}\|_1\}$. Therefore, Hoeffding's inequality implies that the number of samples $S^{(1)}$ necessary to attain an ϵ -precise estimate with $1 - \delta$ confidence is

$$S^{(1)} = \frac{4\|\mathbf{a}\|_1^2}{(\epsilon - \nu^{(1)})^2} \log\left(\frac{4}{\delta}\right), \quad (23)$$

where the real and imaginary parts are sampled independently. By setting the approximation error $\nu^{(1)}$ to $\epsilon/2$ we arrive at the claimed result.

In the case of problem 2, we have to bound $|z^{(2)} - \tilde{z}^{(2)}|$. First, notice that

$$\begin{aligned} \left| \text{Tr}[O f(A) \varrho f^\dagger(A)] - \text{Tr}[O \tilde{f}(A) \varrho \tilde{f}^\dagger(A)] \right| &\leq \left| \text{Tr}[O(f(A) - \tilde{f}(A)) \varrho f^\dagger(A)] \right| \\ &\quad + \left| \text{Tr}[O(f(A) - \tilde{f}(A)) \varrho (f^\dagger(A) - \tilde{f}^\dagger(A))] \right| \\ &\quad + \left| \text{Tr}[O f(A) \varrho (f^\dagger(A) - \tilde{f}^\dagger(A))] \right|, \end{aligned} \quad (24)$$

where we used the triangle inequality after adding and subtracting the terms $\text{Tr}[O f(A) \varrho \tilde{f}(A)]$, $\text{Tr}[O f(A) \varrho f(A)]$ and $\text{Tr}[O \tilde{f}(A) \varrho f(A)]$.

By means of the Holder tracial matrix inequality

$$|\text{Tr}[AB^\dagger]| \leq |\text{Tr}[A]| \|B\|, \quad (25)$$

we can now show that

$$\begin{aligned} \left| \text{Tr}[O f(A) \varrho f^\dagger(A)] - \text{Tr}[O \tilde{f}(A) \varrho \tilde{f}^\dagger(A)] \right| &\leq \|O\| \|f(A) - \tilde{f}(A)\| \|f^\dagger(A)\| + \|O\| \|f(A) - \tilde{f}(A)\|^2 \\ &\quad + \|O\| \|f(A) - \tilde{f}(A)\| \|f(A)\|, \end{aligned} \quad (26)$$

where we used that $|\text{Tr}[\varrho]| = 1$. Therefore,

$$\left| z^{(2)} - \tilde{z}^{(2)} \right| \leq \nu^{(2)} := \nu \|O\| (2\|f(A)\| + \nu). \quad (27)$$

Similarly to the first case, $\tilde{z}^{(2)}$ will be an ϵ -precise estimate of $z^{(2)}$ if and only if

$$\left| \tilde{z}^{(2)} - \mathbb{E}[\tilde{z}^{(2)}] \right| \leq \epsilon - \nu^{(2)}. \quad (28)$$

Since the range of possible values of $\tilde{z}^{(2)}$ lie in $\{-\|O\|\|\mathbf{a}\|_1^2, \|O\|\|\mathbf{a}\|_1^2\}$, Hoeffding's inequality implies that the number of samples necessary to attain an ϵ -precise estimate with $1 - \delta$ confidence is

$$S_2 = \frac{2\|O\|^2\|\mathbf{a}\|_1^4}{(\epsilon - \nu^{(2)})^2} \log\left(\frac{2}{\delta}\right). \quad (29)$$

The claim then follows by setting the approximation error $\nu^{(2)} = \epsilon/2$. $\square$

Supplementary note 2: Error robustness analysis

In this appendix, we prove Theorem 3. Before the proof, two auxiliary lemmas are presented. In Lemma 2 we show that the action of a depolarizing channel can be commuted with the action of a generic unitary channel. In particular, we consider the error model where the unitary is called many times, and each time it acts it is followed by a depolarizing channel. We show that this is equivalent to acting with all the depolarizing channels first and then all the unitaries. Lemma 3 shows how the presence of a depolarizing noise in the block-encoding oracle affects the result of the Hadamard test in Fig. 1a with a fixed j . With these two lemmas, the proof of Theorem 3 is obtained straightforwardly. At the end of the section, we also analyze the influence on the estimate of Algo. 1 of coherent noise in the block-encoding of A .

Supplementary Lemma 2. *Let U be any unitary operator over $\mathcal{H}_1 \otimes \mathcal{H}_a \otimes \mathcal{H}_s$. Then, if Λ is the global depolarization channel with strength p , we have that, for any $\sigma \in \mathcal{H}_1 \otimes \mathcal{H}_a \otimes \mathcal{H}_s$*

$$(\Lambda \circ U)[\sigma] = (U \circ \Lambda)[\sigma], \quad (30)$$

and, in particular,

$$(\Lambda \circ U)^j[\sigma] = U^j \circ \Lambda^j[\sigma], \quad (31)$$

where $U[\sigma]$ is the action of U as a unitary channel over σ .

Proof. First, we have that

$$\begin{aligned} \Lambda \circ U[\sigma] &= \Lambda[(U\sigma U^\dagger)] \\ &= (1-p)(U\sigma U^\dagger) + p \frac{\mathbb{1}}{D_{\text{tot}}} \\ &= U \left((1-p)\sigma + p \frac{\mathbb{1}}{D_{\text{tot}}} \right) U^\dagger \\ &= U \circ \Lambda[\sigma], \end{aligned} \quad (32)$$

where D_{tot} is the dimension of $\mathcal{H}_1 \otimes \mathcal{H}_a \otimes \mathcal{H}_s$. Therefore, since Λ and U commute as channels, we can pass all U to the left of the Λ copies, proving the lemma. $\square$

Supplementary Lemma 3. *Let $\mathbf{Had}(U_A^j, \Lambda, |\phi\rangle, |\psi\rangle)$ be the random variable obtained by the Hadamard test described in Fig. 1a when we use $\Lambda \circ cU_A$ in the place of cU_A . Then, for $B = \mathbb{1}$*

$$\mathbb{E}[\mathbf{Had}(U_A^j, \Lambda, |\phi\rangle, |\psi\rangle)] = (1-p)^j \text{Re}\{\langle \phi | \mathcal{T}_j(A) | \psi \rangle\}, \quad (33)$$

and for $B = S^\dagger$

$$\mathbb{E}[\mathbf{Had}(U_A^j, \Lambda, |\phi\rangle, |\psi\rangle)] = (1-p)^j \text{Im}\{\langle \phi | \mathcal{T}_j(A) | \psi \rangle\}. \quad (34)$$

Proof. Note that, by Fig. 1.a)

$$\mathbb{E}[\mathbf{Had}(U_A^j, \Lambda, |\phi\rangle, |\psi\rangle)] = \text{Tr}[X_1 (N[|+\rangle\langle+|_1 \varrho_{a,s}])], \quad (35)$$

where

$$N = \bar{c}U_\phi \circ (\Lambda \circ cU_A)^j \circ cU_\psi \circ B, \quad (36)$$

and $\varrho_{a,s} = |0\rangle\langle 0|_a |0\rangle\langle 0|_s$. Now, using Lemma 2 we have that

$$\begin{aligned} N &= \bar{c}U_\phi \circ cU_A^j \circ \Lambda^j \circ cU_\psi \circ B \\ &= \bar{c}U_\phi \circ cU_A^j \circ cU_\psi \circ B \circ \Lambda^j = G \circ \Lambda^j, \end{aligned} \quad (37)$$

where G is the unitary channel $G[\sigma] = G\sigma G^\dagger$ with $G = \bar{c}U_\phi cU_A^j cU_\psi B$. Therefore,

$$N[|+\rangle\langle+|_1 \varrho_{a,s}] = G \circ \Lambda^j[|+\rangle\langle+|_1 \varrho_{a,s}] = (1-p)^j G(|+\rangle\langle+|_1 \varrho_{a,s}) G^\dagger + (1 - (1-p)^j) \frac{\mathbb{1}}{D_{\text{tot}}}. \quad (38)$$

Now, by the proof of lem.5, and the fact that X_1 is trace-less, we have that

$$\mathbb{E}[\mathbf{Had}(U_A^j, \Lambda, |\phi\rangle, |\psi\rangle)] = (1-p)^j \mathbb{E}[\mathbf{Had}(U_A^j, |\phi\rangle, |\psi\rangle)], \quad (39)$$

which is equivalent to the statement of the lemma given any choice of B $\square$

Now, we are ready for the proof of Theorem 3.

Proof of Theorem 3. By definition, we have that

$$\left| \mathbb{E}[\tilde{z}^{(1)}] - \mathbb{E}[\tilde{z}^{(1),\Lambda}] \right| = \left| \sum_{j=0}^k a_j (1 - (1-p)^j) \langle \phi | \mathcal{T}_j(A) | \psi \rangle \right|. \quad (40)$$

By Bernoulli's inequality $1 - (1-p)^j \leq pj$ for $p < 1$, implying

$$\left| \mathbb{E}[\tilde{z}^{(1)}] - \mathbb{E}[\tilde{z}^{(1),\Lambda}] \right| \leq p \left| \sum_{j=0}^k j a_j \langle \phi | \mathcal{T}_j(A) | \psi \rangle \right|. \quad (41)$$

Now, using the fact that $|\langle \phi | \mathcal{T}_j(A) | \psi \rangle| \leq 1$ for any $|\phi\rangle, |\psi\rangle$ and the triangle inequality one gets

$$\left| \mathbb{E}[\tilde{z}^{(1)}] - \mathbb{E}[\tilde{z}^{(1),\Lambda}] \right| \leq p \sum_{j=0}^k j |a_j| = p \|\mathbf{a}\|_1 \mathbb{E}[j]. \quad (42)$$

$\square$

Supplementary Lemma 4 (Noise robustness of QSVT). *Let $U_{\tilde{f}(A)}$ be the block-encoding of $\tilde{f}(A)/\|\tilde{f}\|$ resulting from an ideal QSVT circuit, with $\|\tilde{f}\| = \max_{x \in [-1,1]} |\tilde{f}(x)|$. Let $\tilde{z}_{fq}^{(1)} = \|\tilde{f}\| \mathbf{Had}(U_{\tilde{f}(A)}, |\phi\rangle, |\psi\rangle)$ and $\tilde{z}_{fq}^{(1),\Lambda} = \|\tilde{f}\| \mathbf{Had}(U_{\tilde{f}(A)}, \Lambda, |\phi\rangle, |\psi\rangle)$ be the random variable obtained by a Hadamard test on $U_{\tilde{f}(A)}$ and its noisy counterpart with Λ acting after each oracle query, respectively. Then, $\mathbb{E}[\tilde{z}_{fq}^{(1)}] = \text{Re}\{\langle \phi | \tilde{f}(A) | \psi \rangle\}$ and*

$$\left| \mathbb{E}[\tilde{z}_{fq}^{(1)}] - \mathbb{E}[\tilde{z}_{fq}^{(1),\Lambda}] \right| \leq p_{fq} E_{fq}^{(1)} \leq p_{fq} \|\tilde{f}\| k, \quad (43)$$

with $E_{fq}^{(1)} := k \left| \text{Re}\{\langle \phi | \tilde{f}(A) | \psi \rangle\} \right|$.

Proof. First notice that the Hadamard test will work as in Lemma 5 with U_A^j substituted by $U_{\tilde{f}(A)}^j$. The rest of the proof follows analogously to the proof of Lemma 3 by noticing that the circuit always makes k calls to the noisy oracle. $\square$

So far, we considered only the effect of incoherent error. Now we consider the case where the source of error is a coherent imperfection in the block-encoding oracle of A , as we define below.

Supplementary Definition 1. Given an Hermitian operator A with $\|A\| \leq 1$, an ε -approximate block-encoding of A is a unitary operator $U_{A,\varepsilon}$ satisfying

$$\|(\langle 0|_a \otimes \mathbf{1}_s)U_{A,\varepsilon}(|0\rangle_a \otimes \mathbf{1}_s) - A\|_{tr} < \varepsilon, \quad (44)$$

where $\|\cdot\|_{tr}$ is the trace norm.

Under the approximate oracle assumption, the next lemma shows how the error propagates to our Algo. 1 estimate result.

Supplementary Lemma 5. Let $U_{A,\varepsilon}$ be an ε -approximate block-encoding of A , and let U_A be an exact block encoding of A . If $\tilde{z}^{(1)}$ is as in Eq. (6) and $\tilde{z}^{(1),\varepsilon}$ is the unbiased estimator obtained from the Hadamard test in Fig. 1a when exchanging cU_A^j by $cU_{A,\varepsilon}^j$, with $(k\varepsilon) \leq 1/2$, then

$$\left| \mathbb{E}[\tilde{z}^{(1)}] - \mathbb{E}[\tilde{z}^{(1),\varepsilon}] \right| \leq 2\varepsilon \|\mathbf{a}\|_1 \mathbb{E}[j], \quad (45)$$

Proof. If $U_{A,\varepsilon}$ is an ε -approximate block-encoding then,

$$\left\| (\langle 0|_a \otimes \mathbf{1}_s)U_{A,\varepsilon}^j(|0\rangle_a \otimes \mathbf{1}_s) - \mathcal{T}_j(A) \right\|_{tr} \leq 2j\varepsilon, \quad (46)$$

which implies that

$$\left| (\langle 0|_a \langle \phi|_s)U_{A,\varepsilon}^j(|0\rangle_a |\psi\rangle_s) - \langle \phi|\mathcal{T}_j(A)|\psi\rangle \right| \leq 2j\varepsilon. \quad (47)$$

Therefore, denoting by $\delta\tilde{z}^{(1)} \equiv |\mathbb{E}[\tilde{z}^{(1)}] - \mathbb{E}[\tilde{z}^{(1),\varepsilon}]|$ the error in the estimator $z^{(1)}$ from replacing the exact block-encoding U_A by the faulty $U_{A,\varepsilon}$,

$$\delta\tilde{z}^{(1)} = \left| \sum_{j=0}^k a_j \left((\langle 0|_a \langle \phi|_s)U_{A,\varepsilon}^j(|0\rangle_a |\psi\rangle_s) - \langle \phi|\mathcal{T}_j(A)|\psi\rangle \right) \right| \leq \varepsilon \left| \sum_{j=0}^k j a_j \right| \leq 2\varepsilon \|\mathbf{a}\|_1 \mathbb{E}[j]. \quad (48)$$

□

Supplementary note 3: Function approximation for specific use-cases

Function	Domain	Truncation order (k)	$\ \mathbf{a}\ _1$	$\ f\ _\infty$
x^t ($t \in \mathbb{N}$)	$[-1, 1]$	$\sqrt{2t \log \frac{2}{\nu}}$	$\mathcal{O}(1)$	$\mathcal{O}(1)$
$e^{-\beta x/2}$ ($\beta \in \mathbb{R}$)	$[-1, 1]$	$\sqrt{2 \max \left\{ \frac{e^{2\beta}}{2}, \log \left(\frac{2e^{\beta/2}}{\nu} \right) \right\} \log \left(\frac{4e^{\beta/2}}{\nu} \right)}$	$\mathcal{O}(e^{\beta/2})$	$\mathcal{O}(e^{\beta/2})$
x^{-1}	$[-1, -1/\kappa] \cup [1/\kappa, 1]$	$2\kappa \sqrt{\log \frac{\kappa}{\nu} \log \left(\frac{4\kappa^2}{\nu} \log \frac{\kappa}{\nu} \right)} + 1$	$\mathcal{O}(\kappa \sqrt{\log \frac{\kappa}{\nu}})$	$\mathcal{O}(\kappa)$
$\theta(x)$	$[-1, -\xi] \cup [\xi, 1]$	$\mathcal{O}\left(\frac{1}{\xi} \log \frac{1}{\nu}\right)$	$\mathcal{O}\left(\log \left(\frac{1}{\xi} \log \frac{1}{\nu}\right)\right)$	$\mathcal{O}(1)$

Supplementary Table 1: Chebyshev ν -approximation data (approximation interval $\mathcal{I}$, truncation degree k and coefficient 1-norm $\|\mathbf{a}\|_1$) for our exemplary functions. When a singular point is present at the origin, the approximation interval excludes a ball of fixed radius around it. For comparison, we also provide the ∞ -norm $\|f\|_\infty := \max_{x \in \mathcal{I}} |f(x)|$, which in the matrix function setting plays the role of spectral norm $\|f(A)\|$ of the matrix $f(A)$. Namely, the difference between $\|\mathbf{a}\|_1$ and $\|f\|_\infty$ translates into the difference between sample complexities between our semi-quantum algorithm and the corresponding fully-quantum one.

The following general lemma allows us to compute the 1-norm $\|\mathbf{a}\|_1$ in certain cases.

Supplementary Lemma 6. Let $\mathbf{a} = \{a_0, \dots, a_k\}$ be the vector of Chebyshev coefficients of the polynomial $\tilde{f}$ as in (23). Then $\|\mathbf{a}\|_1 = |\tilde{f}(1)|$ if all a_j 's have the same sign and $\|\mathbf{a}\|_1 = |\tilde{f}(-1)|$ if they alternate signs.

Proof. First notice that if all the coefficients have the same sign (respectively, opposite signs) then either $a_j = |a_j|$ or $a_j = -|a_j|$ for all j (respectively, either $a_j = (-1)^j |a_j|$ or $a_j = (-1)^{j+1} |a_j|$ for all j). The claim then follows immediately from the fact that $\mathcal{T}_j(1) = 1$ and $\mathcal{T}_j(-1) = (-1)^j$ for all j . □

3.1 Monomial function

3.1.1 Polynomial approximation and its parameters

The following lemma gives the Chebyshev approximation to the monomial function that we use to prove the complexity of our algorithm. The lemma is a version of [4, Lemma 3.3] with explicit coefficients.

Supplementary Lemma 7. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be such that $f(x) = x^t$ for some $t \in \mathbb{N}$. Then, the function $\tilde{f} : [-1, 1] \rightarrow \mathbb{R}$ given by the Chebyshev polynomial*

$$\tilde{f}(x) = 2^{1-t} \sum_{\substack{j=0 \\ t-j \text{ even}}}^k \binom{t}{(t-j)/2} \mathcal{T}_j(x), \quad (49)$$

is a ν -approximation of f on $[-1, 1]$, if and only if

$$k \geq \sqrt{2t \log\left(\frac{2}{\nu}\right)}. \quad (50)$$

Proof. See [5], Theorem 3.2. □

Supplementary Lemma 8. *Let $\mathbf{a}_{\text{mon}}$ be the vector of coefficients of the Chebyshev polynomial $\tilde{f}_{\text{mon}}$. Then the one-norm is*

$$\|\mathbf{a}_{\text{mon}}\|_1 = 1 - \nu. \quad (51)$$

Proof. Follows as corollary of Lemma 6. □

Supplementary Lemma 9. *Let $\mathbf{a}_{\text{mon}}$ be the vector of coefficients of the Chebyshev polynomial $\tilde{f}_{\text{mon}}$. Then, the average query complexity of estimating $\tilde{z}^{(1)}$ is*

$$\mathbb{E}[j] \leq \frac{1}{1-\nu} \sqrt{\frac{2t}{\pi} \frac{t}{t-1}}. \quad (52)$$

Proof. First notice that

$$\mathbb{E}[j] = \frac{1}{\|\mathbf{a}_{\text{mon}}\|} \sum_{j=1}^k j |a_j| \leq \frac{1}{\|\mathbf{a}_{\text{mon}}\|} \sum_{j=1}^t j |a_j|, \quad (53)$$

since $k \leq t$ by assumption. The summation on the righthand side can be computed exactly, but its precise value depends on whether t is even or odd. For t even, the only non-vanishing Chebyshev coefficients are the even order ones, namely $a_{2\ell} = 2^{1-t} \binom{t}{(t-2\ell)/2}$, and

$$\sum_{j=1}^t j |a_j| = \sum_{\ell=1}^{t/2} 2^{2-t} \ell \binom{t}{(t-2\ell)/2} = 2^{-t} t \binom{t}{t/2} \leq \sqrt{\frac{2t}{\pi}}, \quad (54)$$

where we have used the inequality $\binom{t}{t/2} \leq \frac{2^t}{\sqrt{\pi t/2}}$. Similarly, for odd t the only non-vanishing coefficients are the odd order ones, with $a_{2\ell-1} = 2^{1-t} \binom{t}{(t-2\ell+1)/2}$, and

$$\sum_{j=1}^t j |a_j| = \sum_{\ell=1}^{(t+1)/2} 2^{1-t} (2\ell-1) \binom{t}{(t-2\ell+1)/2} = 2^{1-t} t \binom{t-1}{(t-1)/2} \leq \sqrt{\frac{2t}{\pi} \frac{t}{t-1}}. \quad (55)$$

Notice that the bound (54) contains the one in (55), and therefore can be taken to hold for all t . The claim then follows by incorporating $\|\mathbf{a}_{\text{mon}}\|$ from Lemma 9. □

3.1.2 Problem statement and complexities of the solution

We consider a reversible and ergodic Markov chain on $N = 2^n$ discrete states with row stochastic transition matrix $\mathcal{P}$ satisfying $\sum_k \mathcal{P}_{jk} = 1$ whose stationary state, i.e the eigenstate corresponding to eigenvalue $\lambda = 1$ has the Gibbs amplitudes $\boldsymbol{\pi} = (e^{-\beta E_0}/Z_\beta, \dots, e^{-\beta E_{D-1}}/Z_\beta)$, with $\{E_l\}_{l \in [D]}$ the eigenvalues of the Hamiltonian H . The eigenvalues λ of $\mathcal{P}$ lie in the interval $(-1, 1]$ with $\lambda = 1$ non-degenerate. The so-called *discriminant matrix* A is defined by $A_{jk} = \sqrt{\mathcal{P}_{jk} \mathcal{P}_{kj}}$. For the sake of our algorithm, the relevant properties of A are Hermiticity and the fact that A has the same spectrum of $\mathcal{P}$ with $\sqrt{\boldsymbol{\pi}}$ the eigenvector corresponding to eigenvalue $\lambda = 1$. In particular, A^t for $t = \mathcal{O}(\Delta^{-1})$ becomes approximately a projector over the state $\sqrt{\boldsymbol{\pi}}$ as we state in the following lemma (here $\Delta = \min_{\{\lambda\}} (1 - |\lambda|)$ is the spectral gap of A).

Supplementary Lemma 10 (Determining the monomial degree). *Let $\nu' > 0$, A a Markov-chain discriminant matrix with spectral gap $0 < \Delta < 1$, and $|\mathbf{y}\rangle$ a spin configuration with energy $E_{\mathbf{y}}$. For any integer $t \geq \frac{\log(1/\nu')}{\Delta}$, it holds that*

$$|\langle \mathbf{y} | A^t | \mathbf{y} \rangle - \langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle| \leq \nu' \quad (56)$$

Proof. Let $A = \sum_{\lambda} \lambda |\lambda\rangle\langle\lambda| = \Pi_{\pi} + \sum_{\lambda \neq 1} \lambda |\lambda\rangle\langle\lambda|$ be the spectral decomposition of A , where λ denote its eigenvalues, $\{|\lambda\rangle\}$ the corresponding eigenstates, and $\Pi_{\pi} = |\sqrt{\pi}\rangle\langle\sqrt{\pi}|$ is the projector onto the "purified" Gibbs state. It follows that

$$|\langle \mathbf{y} | A^t - \Pi_{\pi} | \mathbf{y} \rangle| \leq \sum_{\lambda \neq 1} |\lambda|^t \cdot |\langle \mathbf{y} | \lambda \rangle|^2 \leq (1 - \Delta)^t. \quad (57)$$

Hence, given some tolerated error ν' , one can ensure that $|\langle \mathbf{y} | A^t | \mathbf{y} \rangle - \langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle| \leq \nu'$ for any

$$t \geq \frac{\log 1/\nu'}{\Delta} \geq \frac{\log 1/\nu'}{\log(1 - \Delta)^{-1}}. \quad (58)$$

□

Supplementary Lemma 11. *Let $\tilde{z}^{(1)}$ be an estimate to $\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle$ up to additive error $\epsilon < \langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle / 2$. If $\epsilon \leq \epsilon_r(2e^{\beta E_{\mathbf{y}}} Z_{\beta})$, then $\tilde{Z}_{\beta} := \frac{e^{-\beta E_{\mathbf{y}}}}{\tilde{z}^{(1)}}$ is an estimate for the partition function Z_{β} such that*

$$|\tilde{Z}_{\beta} - Z_{\beta}| < \epsilon_r Z_{\beta}. \quad (59)$$

Proof. The estimate $\tilde{z}^{(1)}$ satisfies $|\tilde{z}^{(1)} - \langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle| \leq \epsilon$. It follows straightforwardly that

$$\frac{e^{-\beta E_{\mathbf{y}}}}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle} \left(1 + \frac{\epsilon}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle}\right)^{-1} < \frac{e^{-\beta E_{\mathbf{y}}}}{\tilde{z}^{(1)}} < \frac{e^{-\beta E_{\mathbf{y}}}}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle} \left(1 - \frac{\epsilon}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle}\right)^{-1}. \quad (60)$$

Now, using that $1 - 2\xi \leq (1 + \xi)^{-1}$ and $1 + 2\xi \geq (1 - \xi)^{-1}$ for $0 \leq \xi \leq 1/2$, we can derive the inequality:

$$\frac{e^{-\beta E_{\mathbf{y}}}}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle} \left(1 - 2\frac{\epsilon}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle}\right) < \frac{e^{-\beta E_{\mathbf{y}}}}{\tilde{z}^{(1)}} < \frac{e^{-\beta E_{\mathbf{y}}}}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle} \left(1 + 2\frac{\epsilon}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle}\right), \quad (61)$$

from which we obtain (59) by using $Z_{\beta} = \frac{e^{-\beta E_{\mathbf{y}}}}{\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle}$ and setting $\epsilon \leq \epsilon_r(2e^{\beta E_{\mathbf{y}}} Z_{\beta})$. □

Supplementary Lemma 12 (Complexity). *Let $\epsilon_r > 0$ and $\nu' = \nu/2 = \frac{\epsilon_r}{12e^{\beta E_{\mathbf{y}}} Z_{\beta}}$. Let $\tilde{z}^{(1)}$ be the estimate obtained by Algorithm 1 with inputs $|\psi\rangle = |\phi\rangle = |\mathbf{y}\rangle$, error $\frac{\epsilon_r}{2e^{\beta E_{\mathbf{y}}} Z_{\beta}}$, and $\tilde{f}(A)$ a ν -approximation to A^t with $t = \frac{\log(1/\nu')}{\Delta}$. Then $\tilde{Z}_{\beta} := \frac{e^{-\beta E_{\mathbf{y}}}}{\tilde{z}^{(1)}}$ is an estimate to the partition function Z_{β} up to relative error ϵ_r . Moreover, the algorithm has maximal query depth $k = \sqrt{\frac{2}{\Delta}} \left(\beta E_{\mathbf{y}} + \log \frac{12 Z_{\beta}}{\epsilon_r} \right)$ and sample complexity $S^{(1)} = 64 e^{2\beta E_{\mathbf{y}}} Z_{\beta}^2 \frac{\log(2/\delta)}{\epsilon_r^2}$.*

Proof. The estimate $\tilde{z}^{(1)}$ has three sources of error relatively to $\langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle$. First, according to Lemma 10, the fact that t is finite causes an error ν' . Second, only an approximation of A^t is implemented, causing an error ν . Finally, the finite number of samples incurs a statistical error ϵ_{st} . Therefore, by triangular inequality, we have $|\tilde{z}^{(1)} - \langle \mathbf{y} | \Pi_{\pi} | \mathbf{y} \rangle| \leq (\nu' + \nu + \epsilon_{st}) := \epsilon$. According to Lemma 11, $\tilde{Z}_{\beta} = \frac{e^{-\beta E_{\mathbf{y}}}}{\tilde{z}^{(1)}}$ is an estimate of Z_{β} with relative error ϵ_r if the total error is chosen to be $\epsilon = \epsilon_r(2e^{\beta E_{\mathbf{y}}} Z_{\beta})$. Since each source of error is controlled individually, for convenience, we take $\nu' = \nu/2 = \frac{\epsilon_r}{12e^{\beta E_{\mathbf{y}}} Z_{\beta}}$ and $\epsilon_{st} = \frac{\epsilon_r}{4e^{\beta E_{\mathbf{y}}} Z_{\beta}}$. The sample complexity is then given by Eq. (8) with $\|\mathbf{a}\|_1 = 1 - \nu$ and error ϵ , while the maximal query depth is given by substituting the expression for t in the expression for the truncation order $k = \sqrt{2t \log(2/\nu)}$. □

3.2 Exponential function

The following Lemma gives the explicit coefficients of the approximation used here and its maximal degree. It is obtained from Ref. [5, Lemma 4.2].

Supplementary Lemma 13 (Chebyshev approximation). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be such that $f(x) = e^{-\beta x/2}$. Then, for $\nu \leq \frac{e^{\beta/2}}{2}$, the polynomial $\tilde{f}_{exp} : [-1, 1] \rightarrow \mathbb{R}$ defined by*

$$\tilde{f}_{exp}(x) := \sum_{j=0}^{k'} \left[\sum_{\substack{l=j \\ l-j \text{ even}}}^t \frac{(-\beta/2)^l}{l!} 2^{1-l} \binom{l}{(l-j)/2} \right] \mathcal{T}_j(x), \quad (62)$$

is a ν -approximation of $f(x)$ on $[-1, 1]$ if $t \geq \left\lceil \max \left\{ \frac{\beta e^2}{2}, \log \left(\frac{2e^{\beta/2}}{\nu} \right) \right\} \right\rceil$, and $k \geq \left\lceil \sqrt{2t \log \left(\frac{4e^{\beta/2}}{\nu} \right)} \right\rceil$.

Proof. Lemma 3.4 in Ref. [5] builds an approximation for the function $e^{-\frac{\beta-\beta x}{2}}$ with error $\nu' \leq 1/2$ in the interval $[-1, 1]$. From this construction, it is straightforward to obtain the approximation for $f(x) = e^{-\beta x/2}$ with error $\nu = e^{\beta/2} \nu' \leq \frac{e^{\beta/2}}{2}$. The explicit expression for the coefficients is obtained from the truncated Taylor series of $e^{-\beta x}$ combined with the truncated Chebyshev representation of monomials x^l . $\square$

Supplementary Lemma 14 (l_1 -norm of coefficients). *Let $\mathbf{a}_{exp}$ be the vector of coefficients of the Chebyshev polynomial $\tilde{f}_{exp}$. Then the one-norm satisfies*

$$e^{\beta/2} - \nu \leq \|\mathbf{a}_{exp}\|_1 \leq e^{\beta/2} + \nu, \quad (63)$$

Proof. First, one can see that, by relabeling the summation index, the coefficients of $\tilde{f}_{exp}$ in Eq. (62) can be rewritten as

$$a_j = 2^{1-\delta_{j,0}} \left(\frac{-\beta}{4} \right)^j \sum_{l=0}^{\lfloor \frac{t-j}{2} \rfloor} \frac{(\beta/16)^l}{l!(j+l)!} \quad (64)$$

This form makes it more clear that the coefficients alternate sign. Hence, Lem.6 implies the inequality because by definition $|\tilde{f}_{exp}(1) - f(1)| = |\tilde{f}_{exp}(1) - e^{\beta/2}| \leq \nu$. $\square$

Supplementary Lemma 15 (Average query depth). *The average query complexity of estimating $\tilde{z}^{(2)}$ with $\tilde{f}_{exp}$ satisfies $2\mathbb{E}[j] \leq \sqrt{2\beta}$.*

Proof. Notice that

$$\begin{aligned} 2\mathbb{E}[j] &= \frac{2}{\|\mathbf{a}_{exp}\|_1} \sum_{j=1}^k j |a_j| \leq 4e^{-\beta/2} \sum_{j=1}^k \left[2 \left(\frac{\beta}{4} \right)^j \sum_{l=0}^{\lfloor \frac{t-j}{2} \rfloor} \frac{(\beta/16)^l}{l!(j+l)!} \right] j \\ &\leq 4e^{-\beta/2} \sum_{j=1}^k \left[2 \left(\frac{\beta}{4} \right)^j \sum_{l=0}^{\infty} \frac{(\beta/16)^l}{l!(j+l)!} \right] j \\ &= 4e^{-\beta/2} \sum_{j=1}^k [2I_j(\beta/2)] j, \end{aligned} \quad (65)$$

where the first inequality comes from applying Lemma 14 with the worst case of ν and Eq. (64). The second inequality is due to the addition of positive terms to the inner sum. In the last equality, we identified the coefficients of the outer sum with the power series of the modified Bessel functions $I_j(\beta/2)$. This also has the interesting feature of connecting the Chebyshev series in [5] with the widely known Jacobi-Anger expansion [1]. Now, we can use the identity $j I_j(\beta/2) = \frac{\beta}{4}(I_{j-1} - I_{j+1})$ to obtain:

$$2\mathbb{E}[j] \leq 2e^{-\beta/2} \beta [I_0(\beta/2) - I_{k+1}(\beta/2)] \leq 2e^{-\beta/2} \beta I_0(\beta/2) \leq \sqrt{2\beta}. \quad (66)$$

$\square$

3.3 Inverse function

The following Lemma gives the explicit coefficients of the approximation used here and its maximal degree. It is obtained from Ref. [2].

Supplementary Lemma 16. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be such that $f(x) = 1/x$. Then, the polynomial $\tilde{f}_{inv} : [-1, 1] \rightarrow \mathbb{R}$ given by*

$$\tilde{f}_{inv}(x) = 4 \sum_{j=0}^k (-1)^j \left[\frac{\sum_{i=j+1}^b \binom{2b}{b+i}}{2^{2b}} \right] \mathcal{T}_{2j+1}(x), \quad (67)$$

is ν -approximation of $1/x$ on $[-1, -1/\kappa] \cup [1/\kappa, 1]$ for any

$$k \geq \sqrt{b \log \left(\frac{4b}{\nu} \right)} \quad \text{and} \quad b \geq \kappa^2 \log \left(\frac{\kappa}{\nu} \right). \quad (68)$$

Supplementary Lemma 17. Let $\mathbf{a}_{\text{inv}}$ be the vector of coefficients of the polynomial $\tilde{f}_{\text{inv}}$. Then $\frac{1}{2}\sqrt{b} - \frac{2bk}{2^{2b}}\binom{2b}{k} \leq \|\mathbf{a}_{\text{inv}}\|_1 \leq \frac{2}{\sqrt{\pi}}\sqrt{b}$. In particular, $\|\mathbf{a}_{\text{inv}}\|_1 \in \Theta(\sqrt{b})$.

Proof. First notice that

$$\begin{aligned} \|\mathbf{a}_{\text{inv}}\|_1 &= \sum_{j=0}^k \left[\frac{4}{2^{2b}} \sum_{i=j+1}^b \binom{2b}{b+i} \right] \\ &= \frac{4}{2^{2b}} \left[\sum_{i=1}^{k+1} \sum_{j=0}^{i-1} \binom{2b}{b+i} + \sum_{i=k+2}^b \sum_{j=0}^k \binom{2b}{b+i} \right] \\ &= \frac{4}{2^{2b}} \left[\sum_{i=1}^{k+1} i \binom{2b}{b+i} + \sum_{i=k+2}^b (k+1) \binom{2b}{b+i} \right] \\ &= \frac{4}{2^{2b}} \left[\sum_{i=1}^b i \binom{2b}{b+i} - \sum_{i=1}^{b-k-1} i \binom{2b}{b+i+k+1} \right]. \end{aligned} \quad (69)$$

One immediately gets the upper bound

$$\|\mathbf{a}_{\text{inv}}\|_1 \leq \frac{4}{2^{2b}} \sum_{i=1}^b i \binom{2b}{b+i} = \frac{2b}{2^{2b}} \binom{2b}{b} \leq \frac{2}{\sqrt{\pi}}\sqrt{b}, \quad (70)$$

where we have used the inequality $\binom{2b}{b} \leq \frac{2^{2b}}{\sqrt{\pi b}}$. Similarly, for the lower bound we get

$$\begin{aligned} \|\mathbf{a}_{\text{inv}}\|_1 &\geq \frac{4}{2^{2b}} \sum_{i=1}^{b-k-1} i \left[\binom{2b}{b+i} - \binom{2b}{b+i+k+1} \right] \\ &= \frac{4}{2^{2b}} \sum_{i=1}^{b-k-1} i \binom{2b}{b+i} \left[1 - \prod_{\ell=0}^k \frac{b-i+1-\ell}{b+i+\ell} \right] \\ &\geq \frac{4}{2^{2b}} \sum_{i=1}^{b-k-1} i \binom{2b}{b+i} \left[1 - \left(\frac{b}{2b-1} \right)^{k+1} \right] \\ &\geq \frac{2}{2^{2b}} \sum_{i=1}^{b-k-1} i \binom{2b}{b+i} \\ &= \frac{b}{2^{2b}} \binom{2b}{b} - \frac{2}{2^{2b}} \sum_{i=b-k}^b i \binom{2b}{b+i} \\ &\geq \frac{1}{2}\sqrt{b} - \frac{2bk}{2^{2b}} \binom{2b}{k}, \end{aligned} \quad (71)$$

where in the second line we have made repeated use of the identity $\binom{n}{m} = \frac{n-m+1}{m} \binom{n}{m-1}$ and the last inequality follows from $\binom{2b}{b} \geq \frac{2^{2b-1}}{\sqrt{b}}$. Since $\binom{2b}{k} \leq \left(\frac{2eb}{k} \right)^k$, it follows that in the asymptotic limit $b \gg k \gg 1$ of interest here the second term is exponentially small in b and, therefore, $\|\mathbf{a}_{\text{inv}}\|_1 \in \Omega(\sqrt{b})$. Together with the upper bound (70), this implies $\|\mathbf{a}_{\text{inv}}\|_1 \in \Theta(\sqrt{b})$. $\square$

Supplementary Lemma 18. Let $\mathbf{a}_{\text{inv}}$ be the vector of coefficients of the Chebyshev polynomial $\tilde{f}_{\text{inv}}$. Then, the average query complexity of estimating $\tilde{z}^{(1)}$ satisfies $\mathbb{E}[j] \leq b/\|\mathbf{a}_{\text{inv}}\|_1$. In particular, $\mathbb{E}[j] \in \mathcal{O}(\sqrt{b})$.

Proof. The proof parallels that of Lemma 17. Namely,

$$\begin{aligned}
\mathbb{E}[j] &= \frac{1}{\|\mathbf{a}_{inv}\|_1} \sum_{j=0}^k (2j+1) \left[\frac{4}{2^{2b}} \sum_{i=j+1}^b \binom{2b}{b+i} \right] \\
&= \frac{4}{2^{2b}\|\mathbf{a}_{inv}\|_1} \left[\sum_{i=1}^{k+1} i^2 \binom{2b}{b+i} + \sum_{i=k+2}^b (k+1)^2 \binom{2b}{b+i} \right] \\
&\leq \frac{4}{2^{2b}\|\mathbf{a}_{inv}\|_1} \sum_{i=1}^b i^2 \binom{2b}{b+i} \\
&= \frac{b}{\|\mathbf{a}_{inv}\|_1}.
\end{aligned}$$

Lemma 17 then implies that $\mathbb{E}[j] \in \mathcal{O}(\sqrt{b})$. □

3.4 Heaviside step function

Supplementary Lemma 19. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be such that $f(x) = \theta(x)$, where $\theta(x)$ is the unit step function. Then, the function $\tilde{f}_s : [-1, 1] \times [-1, 1] \rightarrow \mathbb{R}$ given by the Chebyshev expansion

$$\tilde{f}_s(y-x) = \frac{1}{2} + \sum_{j=0}^k a_j(y) \mathcal{T}_{2j+1}(x) + b_j(y) \sqrt{1-x^2} U_{2j}(x) \quad (72)$$

where

$$\begin{aligned}
a_j(y) &= -\sqrt{\frac{2\sigma}{\pi}} e^{-\sigma} \frac{I_j(\sigma) + I_{j+1}(\sigma)}{2j+1} \sqrt{1-y^2} U_{2j}(y), \\
b_j(y) &= \sqrt{\frac{2\sigma}{\pi}} e^{-\sigma} \frac{I_j(\sigma) + I_{j+1}(\sigma)}{2j+1} \mathcal{T}_{2j+1}(y),
\end{aligned} \quad (73)$$

with I_n being the modified Bessel function of the first kind and U_{2j} is a Chebyshev polynomial of the second-kind, is a ν -approximation of $f(y-x)$ on all $(x, y) \in [-1, 1] \times [-1, 1]$ such that $x \in [-1, -\xi + y] \cup [\xi + y, 1]$ if

$$k = \mathcal{O}\left(\frac{1}{\xi} \log\left(\frac{1}{\nu}\right)\right), \text{ and } \sigma = \mathcal{O}\left(\frac{1}{\xi^2} \log\left(\frac{1}{\nu}\right)\right). \quad (74)$$

Proof. We take the Chebyshev approximation

$$\tilde{f}(x) = \frac{1}{2} + \sqrt{\frac{2\sigma}{\pi}} e^{-\sigma} \sum_{j=0}^k (-1)^j \frac{I_j(\sigma) + I_{j+1}(\sigma)}{2j+1} \mathcal{T}_{2j+1}(x) \quad (75)$$

described in [6] and extend it to the two-variable case using the identity

$$\theta(y-x) = \theta(\sin(\arcsin y - \arcsin x)), \quad (76)$$

that follows from fact that $\sin$ and $\arcsin$ are monotonous functions, and expand

$$\mathcal{T}_{2j+1}(\sin(\arcsin y - \arcsin x)) \quad (77)$$

using trigonometric identities and the trigonometric definitions of $\mathcal{T}_{2j+1}$ and U_{2j} . The conditions on k and σ needed for obtaining the ν -approximation end up being the same as those for the single-variable version. □

Supplementary Lemma 20 (l_1 -norm of coefficients). Let $\mathbf{a}_s(y_*)$ and $\mathbf{b}_s(y_*)$ be the vector of coefficients of the Chebyshev expansion $\tilde{f}_s(y_* - x)$ for $y_* = \frac{1}{\sqrt{2}}$. Then

$$\|\mathbf{a}_s(y_*)\|_1 = \|\mathbf{b}_s(y_*)\|_1 \in \Theta(\log(k)). \quad (78)$$

Proof. Let $\|\mathbf{a}_s(y_*)\|_1 = \|\mathbf{b}_s(y_*)\|_1 = N(k, \sigma)$. To prove the asymptotic growth orders we need to investigate the behavior of the one-norms when $\xi, \nu \rightarrow 0$, which is equivalent to $k, \sigma \rightarrow \infty$. Since, by Lemma 19, σ grows quadratically faster than k as $\xi \rightarrow 0$, we first investigate the behavior of $N(k, \sigma)$ as σ grows. Computing the limit

$$\bar{N}(k) = \lim_{\sigma \rightarrow \infty} N(k, \sigma) = \lim_{\sigma \rightarrow \infty} \sum_{j=0}^k \sqrt{\frac{\sigma}{\pi}} e^{-\sigma} \frac{I_j(\sigma) + I_{j+1}(\sigma)}{2j+1} \quad (79)$$

gives us

$$\bar{N}(k) = \frac{1}{\pi\sqrt{2}} \sum_{j=0}^k \frac{1}{2j+1}, \quad (80)$$

where we have used $\lim_{\sigma \rightarrow \infty} \sqrt{\sigma} e^{-\sigma} I_l(\sigma) = \frac{1}{\sqrt{2\pi}}$. Now, note that

$$1 + \frac{H_k}{2} < \sum_{j=0}^k \frac{1}{2j+1} < \frac{H_{k+1}}{2}, \quad (81)$$

where $H_k = \sum_{j=1}^k \frac{1}{j}$ is the k -th harmonic number. By the properties of H_k , we have that there exists k_0 such that for all $k > k_0$,

$$\frac{1}{2\pi\sqrt{2}} \ln(k+1) < \bar{N}(k) < \frac{\sqrt{2}}{\pi} \ln(k+1), \quad (82)$$

which proves our statements. $\square$

Supplementary Lemma 21. Let $\mathbf{a}_s(y_*)$ and $\mathbf{b}_s(y_*)$ be the vector of coefficients of the Chebyshev expansion $\tilde{f}_s(y_* - x)$ for $y_* = \frac{1}{\sqrt{2}}$. Then, the average query complexity of estimating $\tilde{z}^{(1)}$ is $\mathbb{E}[j] \in \mathcal{O}\left(\frac{1}{\xi} \sqrt{\log\left(\frac{1}{\nu}\right)} / \log\left(\frac{1}{\xi} \log\left(\frac{1}{\nu}\right)\right)\right)$.

Proof. By definition of $\mathbb{E}[j]$, and using the fact that $|a_j(y_*)| = |b_j(y_*)|$ for all j , we have that

$$\begin{aligned} \mathbb{E}[j] &= \frac{1}{\|\mathbf{a}_s(y_*)\|_1} \sum_{j=0}^k \sqrt{\frac{4\sigma}{\pi}} e^{-\sigma} (I_j(\sigma) + I_{j+1}(\sigma)) \\ &\leq \frac{1}{\|\mathbf{a}_s(y_*)\|_1} \sum_{j=0}^{\infty} \sqrt{\frac{4\sigma}{\pi}} e^{-\sigma} (I_j(\sigma) + I_{j+1}(\sigma)) \\ &= \frac{1}{\|\mathbf{a}_s(y_*)\|_1} \sqrt{\frac{4\sigma}{\pi}}, \end{aligned} \quad (83)$$

where in the last step $\sum_{j=0}^{\infty} (I_j(\sigma) + I_{j+1}(\sigma)) = e^{\sigma}$ follows from the Jacobi-Anger expansion. Since $\|\mathbf{a}_s(y_*)\| = \Omega(\log(k))$ for large k , it follows that

$$\mathbb{E}[j] = \mathcal{O}\left(\frac{\sqrt{\sigma}}{\log(k)}\right) = \mathcal{O}\left(\frac{\frac{1}{\xi} \sqrt{\log\left(\frac{1}{\nu}\right)}}{\log\left(\frac{1}{\xi} \log\left(\frac{1}{\nu}\right)\right)}\right). \quad (84)$$

$\square$

Supplementary Lemma 22. Let $\mathbf{a}_s(y_*)$ and $\mathbf{b}_s(y_*)$ be the vector of coefficients of the Chebyshev expansion $\tilde{f}_s(y_* - x)$ for $y_* = \frac{1}{\sqrt{2}}$. Then, the error damping factor for estimating $\tilde{z}^{(1)}$ is $\mathcal{O}\left(\frac{1}{\xi} \sqrt{\log\left(\frac{1}{\nu}\right)} / \log\left(\frac{1}{\xi} \log\left(\frac{1}{\nu}\right)\right)\right)$.

Proof. Follows immediately from Lemmas 19 and 21. $\square$

Supplementary Definition 2. Given $\tilde{F}_{\varrho}(y) = \text{Tr}\left[\varrho \tilde{f}_s(y\mathbb{1} - H)\right]$, its T -part is the expression

$$\sum_{j=0}^k a_j(y) \text{Tr}[\mathcal{T}_{2j+1}(H) \varrho], \quad (85)$$

while its U -part is the expression

$$\sum_{j=0}^k b_j(y) \text{Tr}\left[\sqrt{1 - H^2} U_{2j}(H) \varrho\right]. \quad (86)$$

Supplementary Lemma 23. Let $y_* = \frac{1}{\sqrt{2}}$, and $\tilde{z}^{(1),T}(y)$ and $\tilde{z}^{(1),U}(y)$ be the unbiased estimators of the T -part and the U -part of $\tilde{F}_{\varrho}(y)$ obtained from algorithm 1, for any point y , respectively. Then, for any point y we have that

$$\tilde{F}_{\varrho}(y) = \frac{1}{2} + \mathbb{E}_{w^T(y)} \left[\tilde{z}^{(1),T}(y_*) \right] + \mathbb{E}_{w^T(y)} \left[\tilde{z}^{(1),U}(y_*) \right], \quad (87)$$

where $\mathbb{E}_{w^T(y)}$ and $\mathbb{E}_{w^U(y)}$ are the weighted expected values with weight functions

$$\begin{aligned} w^T(y, j) &= \frac{a_j(y)}{a_j(y_*)}, \\ w^U(y, j) &= \frac{b_j(y)}{b_j(y_*)}. \end{aligned} \tag{88}$$

Proof. To show this, just notice that by the definition of $\mathbb{E}_{w^T(y)}$ and $\tilde{z}^{(1,T)}(y_*)$ that $\mathbb{E}_{w^T(y)} [\tilde{z}^{(1,T)}(y_*)]$ is equal to

$$\begin{aligned} \sum_{j=0}^k \sum_b w^T(y, j) p_j(p(b|j) \|\mathbf{a}_s(y_*)\| \operatorname{sgn}(a_j(y_*))b) &= \sum_{j=0}^k w^T(y, j) a_j(y_*) \langle \psi | \mathcal{T}_{2j+1}(H) | \psi \rangle \\ &= \sum_{j=0}^k a_j(y) \langle \psi | \mathcal{T}_{2j+1}(H) | \psi \rangle \\ &= \mathbb{E} [\tilde{z}^{(1,T)}(y)], \end{aligned}$$

while a similar computation also shows that

$$\mathbb{E}_{w^U(y)} [\tilde{z}^{(1,U)}(y_*)] = \mathbb{E} [\tilde{z}^{(1,U)}(y)]. \tag{89}$$

which proves the lemma. $\square$

Supplementary Corollary 24. *To sample $\tilde{F}_\rho(y)$ for $\mathcal{O}(\log(\xi^{-1}))$ values of y it is only necessary to sample $\tilde{F}_\rho(y_*)$ with $1 - \mathcal{O}(\delta/\log(\xi^{-1}))$ confidence, and for each desired value of y , to multiply each sample by the appropriate weight factor.*

Proof. Follows directly from the previous lemma. $\square$

Supplementary note 4: Extensions

4.1 Extension to QSVT

For simplicity, in the main text, we considered only Hermitian matrices. In that case, there is an operation to build a qubitized oracle whose powers directly lead to the Chebyshev polynomials of the block-encoded matrix. This section shows that our algorithm can be applied to randomize the more general quantum singular value transformation technique.

Let us consider a unitary U_A such that $(\langle 0|_a \otimes \mathbb{1}_s) U_A (|0\rangle_a \otimes \mathbb{1}_s) = A$ as in Eq. (20), except that now the block-encoded matrix A does not need to be Hermitian. In fact, A could even be rectangular. In this case, U_A is a block-encoding of a square matrix built from A by filling the remaining entries with zeroes. The input matrix assumes a singular value decomposition, and we are interested in singular value transformations as in the definition below (Def. 16 in Ref. [3]):

Supplementary Definition 3 (Singular value transformations). *Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be an even or odd function. Let $\sum_{j=1}^d \sigma_j |\tilde{\psi}_j\rangle\langle\psi_j|$ be a singular value decomposition of the $d \times d$ matrix A with right eigenvectors $\{|\psi_j\rangle\}$, left eigenvectors $\{|\tilde{\psi}_j\rangle\}$ and singular values $\{\sigma_j\}$ (including zero values). The singular value transformation of A by f is defined as*

$$f^{(SV)}(A) := \begin{cases} \sum_{j=1}^d f(\sigma_j) |\psi_j\rangle\langle\psi_j|, & f \text{ even}, \\ \sum_{j=1}^d f(\sigma_j) |\tilde{\psi}_j\rangle\langle\psi_j|, & f \text{ odd}. \end{cases} \tag{90}$$

It was shown in Ref. [3] that a quantum singular value transformation (QSVT) could be implemented from calls to U_A and $U_A^\dagger$, and single qubit rotations controlled by $|0\rangle_a$. In Theorem 17 and Corollary 18, the authors state that the singular value transformation by a real polynomial of degree k with even or odd parity is always achieved by some sequence of qubit rotations $\Phi = \{\phi_1, \dots, \phi_k\} \in \mathbb{R}$ and a circuit making k queries to U_A and $U_A^\dagger$, as long as the polynomial is normalized to 1. In particular, to implement the Chebyshev polynomial $T_k^{(SV)}(A)$, the sequence of angles is simply $\phi_1 = (1 - k)\frac{\pi}{2}$ and $\phi_j = \frac{\pi}{2}$ for $j \neq 1$.

Under singular value transformation, Probs. 1 and 2 can be phrased in the same way, except that now the Hermitean A is substituted by a general matrix A and the function is meant to be a singular value transformation according to Def. 3. Moreover, the problems are well-posed only for functions with defined parity. Therefore, among

the use cases presented in the main text, the ones that are extendable to general matrices are the inverse function and matrix monomials.

The randomized QSVT solutions to Probs. 1 and 2 follow as before by sampling polynomial degrees according to their weight in the Chebyshev expansion of the target function. A Hadamard test is then run for each polynomial degree draw, the Chebyshev polynomials being implemented using the circuit in Fig. 1 in [3] with the angle sequence described above. The algorithm works for the same reason as for Hermitian matrices: QSVT implements a block-encoding of the polynomial singular value transformation, while the Hadamard test has the useful feature of always selecting the correct block of a unitary block-encoding.

4.2 Non-normalized matrices

Suppose that, instead of having a normalized matrix A (i.e., satisfying $\|A\| \leq 1$), we have a general Hermitian matrix A as input. We can define the normalized matrix $A' = \frac{A}{\alpha}$, with $\alpha \geq \|A\|$ to be block-encoded instead of A . Therefore, using our framework we can simulate a function $f'(A')$ defined such that

$$f(A) = f(\alpha A') = f'(A'), \quad (91)$$

where $f : \mathcal{D} \rightarrow \mathbb{R}$ and $f' : \frac{\mathcal{D}}{\alpha} \subseteq [-1, 1] \rightarrow \mathbb{R}$ have the same image. Therefore, introducing a sub-normalization will require a Chebyshev expansion for a new function whose domain is also sub-normalized relative to the original. The contribution of the sub-normalization factor α to the sample and query complexities of the randomized algorithm needs to be analyzed on a case-by-case basis. In the following, we analyze the maximal query depth and the l_1 -norm of the Chebyshev coefficients for each use-case we showed in the main text. From them and Theorem 2, the complexities of particular instances of Problem 1 or Problem 2 can be obtained.

In the case of monomials, we have that $f(x) = x^t$ yields $f'(x') = \alpha^t x'^t$. This implies that an ν -approximation to A^n is ν/α^n -approximation to H^n . To attain an error ν in approximating $f(x)$, the approximation in Lemma 7 shall be used with error ν/α^n and also the l_1 -norm of the coefficients will gain a factor α^n . Therefore, The maximal query depth and coefficients are $k = \sqrt{2t \log\left(\frac{2\alpha^t}{\nu}\right)}$ and $\|\mathbf{a}\|_1 = \alpha^t - \nu$, respectively.

In the case of the exponential function, notice that $e^{-\beta A} = e^{-(\beta\alpha)(A/\alpha)}$. Therefore, we just need to re-scale β with a factor α , obtaining $k = \mathcal{O}\left(\sqrt{\alpha\beta \log\left(\frac{e^{\alpha\beta}}{\nu}\right)}\right)$ and $\|\mathbf{a}\|_1 = e^{\alpha\beta}$ from Lemmas 13 and 14, respectively.

In the case of the inverse function, $f(x) = x^{-1}$ needs to be approximated in the domain $\mathcal{D} = [-\|A\|, -1/\|A^{-1}\|] \cup [1/\|A^{-1}\|, \|A\|]$. In this case, the randomized algorithm should implement an approximation to $f'(A') = \alpha^{-1}A'^{-1}$ in the domain $\mathcal{D}/\alpha = [-1, -1/\kappa] \cup [1/\kappa, 1]$, with $\alpha = \|A\|$ and $\kappa = \|A\| \|A^{-1}\|$ as usual. It suffices to apply Lemma 16 to obtain a $\alpha\nu$ -approximation to A'^{-1} with resulting maximal degree $k = \mathcal{O}\left(\kappa \log \frac{\|A^{-1}\|}{\nu}\right)$ and coefficients l_1 -norm

$$\|\mathbf{a}\|_1 = \mathcal{O}\left(\|A^{-1}\| \log \frac{\|A^{-1}\|}{\nu}\right).$$

Lastly, in the case of the Heaviside step function, we have that $f(A) = f'(A')$. However, in order to approximate $\theta(A)$ in the domain $D = (-1, -\xi] \cup [\xi, 1)$, we need to be able to approximate $\theta(A')$ in the larger domain $D = (-1, -\xi/\alpha] \cup [\xi/\alpha, 1)$. Therefore, Lemmas 19 and 20 give $k = \mathcal{O}\left(\frac{\alpha}{\xi} \log \frac{1}{\nu}\right)$ and $\|\mathbf{a}\|_1 = \mathcal{O}\left(\log\left(\frac{\alpha}{\xi} \log \frac{1}{\nu}\right)\right)$.

Supplementary References

- [1] M. Abramowitz, I. A. Stegun, “Handbook of mathematical functions,” Applied mathematics series 55, 62 (1966).
- [2] Andrew M. Childs, Robin Kothari, and Rolando D. Somma. “Quantum algorithm for systems of linear equations with exponentially improved dependence on precision”. In: *SIAM Journal on Computing* 46.6 (Jan. 2017), pp. 1920–1950. ISSN: 0097-5397, 1095-7111. DOI: 10.1137/16M1087072. URL: <http://arxiv.org/abs/1511.02306>.
- [3] András Gilyén et al. “Quantum Singular Value Transformation and beyond: Exponential Improvements for Quantum Matrix Arithmetics”. In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2019 (2019), p. 193. DOI: 10.1145/3313276.3316366. URL: <https://doi.org/10.1145/3313276.3316366>.
- [4] S. Sachdeva and N. K. Vishnoi. “Faster algorithms via approximation theory”. In: *Found. Trends Theor. Comput. Sci.* 9 (2013), pp. 125–210. DOI: 10.1561/04000000065.
- [5] Sushant Sachdeva and Nisheeth Vishnoi. *Approximation Theory and the Design of Fast Algorithms*. arXiv:1309.4882. Sept. 2013. DOI: 10.48550/arXiv.1309.4882. URL: <http://arxiv.org/abs/1309.4882>.

- [6] Kianna Wan, Mario Berta, and Earl T. Campbell. “Randomized Quantum Algorithm for Statistical Phase Estimation”. In: *Physical Review Letters* 129.3 (July 2022). Publisher: American Physical Society, p. 030503. DOI: 10.1103/PhysRevLett.129.030503. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.129.030503>.