Supplemental Material
Estimation of Hamiltonian parameters from thermal states

Luis Pedro García-Pintos,[1,*] Kishor Bharti,[2,3] Jacob Bringewatt,[2] Hossein Dehghani,[2,4] Adam Ehrenberg,[2] Nicole Yunger Halpern,[2,5] and Alexey V. Gorshkov[2]

[1] *Theoretical Division (T4), Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[2] *Joint Center for Quantum Information and Computer Science and Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA*
[3] *A*STAR Quantum Innovation Centre (Q.InC), Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, # 16-16 Connexis, Singapore 138632, Republic of Singapore*
[4] *Department of Physics, University of Maryland, College Park, Maryland 20742, USA*
[5] *Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA*

APPENDICES

I. THE QUANTUM FISHER INFORMATION MATRIX

In this section, we derive the closed-form expression for the quantum Fisher information matrix of Hamiltonian parameters for thermal states.

For full-rank states, the quantum Fisher information matrix has elements [1]

$$\mathcal{F}_{lm} := 2 \sum_{jk} \frac{\mathrm{Re}\left[\langle j|\, \partial_l \rho \,|k\rangle \langle k|\, \partial_m \rho \,|j\rangle\right]}{p_j + p_k} \,. \tag{A1}$$

We have defined $\partial_l := \frac{\partial}{\partial \mu_l}$. The matrix characterizes the precision with which multiple parameters μ_l can be estimated. Let $\mathcal{N}$ denote the number of measurements performed. The multiparameter Cramér-Rao bound says that [1]

$$\mathrm{cov}(\hat{\vec{\mu}}) \geq \frac{1}{\mathcal{N}} \mathcal{F}^{-1} \,. \tag{A2}$$

This bound is asymptotically saturable if and only if

$$\mathrm{Tr}\left(\rho[L_l, L_m]\right) = 0. \tag{A3}$$

---

* lpgp@lanl.gov

The symmetric logarithmic derivative is defined by $\partial_l \rho = \frac{1}{2}\{L_l, \rho\}$.

Throughout this appendix, we omit the temperature dependence from the partition-function notation: $Z \equiv Z_\beta$. Since $\rho = e^{-\beta H}/Z = \sum_j e^{-\beta\omega_j} |j\rangle\langle j| /Z$, the derivative in Eq. (A1) is

$$\partial_l \rho = \frac{1}{Z}\partial_l e^{-\beta H} - \rho\frac{\partial_l Z}{Z} = \frac{1}{Z}\left[\partial_l e^{-\beta H} - \rho \operatorname{Tr}\left(\partial_l e^{-\beta H}\right)\right]. \tag{A4}$$

We must calculate the matrix elements of $\partial_l e^{-\beta H}$. Using the Taylor series

$$e^{-\beta H} = \sum_{n=0}^{\infty}(-\beta)^n \frac{H^n}{n!}, \tag{A5}$$

we obtain

$$\begin{aligned}
\langle j| \partial_l e^{-\beta H} |k\rangle &= \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \langle j| \partial_l H^n |k\rangle \\
&= \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} \langle j| \sum_{m=0}^{n-1} H^m A_l H^{n-m-1} |k\rangle \\
&= \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} \sum_{m=0}^{n-1} \omega_j^m \omega_k^{n-m-1} \langle j| A_l |k\rangle \\
&= \langle j| A_l |k\rangle \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} \sum_{m=0}^{n-1} \omega_j^m \omega_k^{n-m-1} \\
&=: \langle j| A_l |k\rangle \, \Gamma_{jk}.
\end{aligned} \tag{A6}$$

We have defined

$$\Gamma_{jk} \coloneqq \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} \sum_{m=0}^{n-1} \omega_j^m \omega_k^{n-m-1} \tag{A7}$$

as a function of the temperature and of the Hamiltonian's spectrum.

We can re-express Γ_{jk} using the formula for an infinite geometric series: if $\omega_j \neq \omega_k$, then

$$\Gamma_{jk} = \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} \frac{\omega_j^n - \omega_k^n}{\omega_j - \omega_k} = \frac{e^{-\beta\omega_j} - e^{-\beta\omega_k}}{\omega_j - \omega_k} = Z\frac{(p_j - p_k)}{\omega_j - \omega_k}, \qquad \text{for} \quad \omega_j \neq \omega_k. \tag{A8}$$

If $\omega_j = \omega_k$, then

$$\begin{aligned}
\Gamma_{jk} &= \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} \sum_{m=0}^{n-1} \omega_j^{n-1} = \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} n\omega_j^{n-1} \\
&= -\beta \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{(n-1)!}\omega_j^{n-1} = -\beta e^{-\beta\omega_j} \\
&= -\beta Z p_j, \qquad \text{for} \quad \omega_j = \omega_k.
\end{aligned} \tag{A9}$$

Using Eqs. (A9) and (A6) we can evaluate the first term in Eq. (A4):

$$\frac{\operatorname{Tr}\left(\partial_l e^{-\beta H}\right)}{Z} = \frac{1}{Z}\sum_j \langle j| A_l |j\rangle \Gamma_{jj} = -\beta\frac{1}{Z}\sum_j \langle j| A_l |j\rangle e^{-\beta\omega_j} = -\beta\langle A_l\rangle. \tag{A10}$$

We denote thermal averages by $\langle A_l\rangle \coloneqq \operatorname{Tr}(A_l\rho)$.

Substituting from Eq. (A10) into Eq. (A4) yields

$$\partial_l \rho = \frac{1}{Z}\partial_l e^{-\beta H} + \beta\langle A_l\rangle\rho. \tag{A11}$$

Furthermore, substituting into Eq. (A6) from Eqs. (A8) and (A9) yields

$$\langle j|\,\partial_l e^{-\beta H}\,|k\rangle = \langle j|\,A_l\,|k\rangle\, Z\frac{(p_j - p_k)}{\omega_j - \omega_k}\,, \qquad \text{for} \qquad \omega_j \neq \omega_k, \tag{A12}$$

and

$$\langle j|\,\partial_l e^{-\beta H}\,|k\rangle = \langle j|\,A_l\,|k\rangle\,\Gamma_{jj} = -\langle j|\,A_l\,|k\rangle\,\beta Z p_j \qquad \text{for} \qquad \omega_j = \omega_k. \tag{A13}$$

Let $\delta A_l := A_l - \langle A_l\rangle = A_l - \text{Tr}\,(\rho A_l)$. If $\omega_j \neq \omega_k$, then

$$\langle j|\,\partial_l\rho\,|k\rangle = \langle j|\,A_l\,|k\rangle\,\frac{(p_j - p_k)}{\omega_j - \omega_k} + \beta\langle A_l\rangle\,\langle j|\,\rho\,|k\rangle = \langle j|\,\delta A_l\,|k\rangle\,\frac{(p_j - p_k)}{\omega_j - \omega_k}\,, \qquad \text{for } \omega_j \neq \omega_k, \tag{A14}$$

whereas, if $\omega_j = \omega_k$,

$$\langle j|\,\partial_l\rho\,|k\rangle = -\langle j|\,A_l\,|k\rangle\,\beta p_j + \beta\langle A_l\rangle p_j\delta_{jk} = -\langle j|\,\delta A_l\,|k\rangle\,\beta p_j = -\langle j|\,\delta A_l\,|k\rangle\,\beta\frac{p_j + p_k}{2}\,, \qquad \text{for } \omega_j = \omega_k. \tag{A15}$$

Thus, the quantum Fisher information matrix in Eq. (A1) becomes

$$\begin{aligned}
\mathcal{F}_{lm} &:= 2\sum_{jk}\frac{\text{Re}\,[\langle j|\,\partial_l\rho\,|k\rangle\langle k|\,\partial_m\rho\,|j\rangle]}{p_j + p_k} \\
&= 2\sum_{\omega_j\neq\omega_k}\frac{(p_j - p_k)^2}{(p_j + p_k)(\omega_j - \omega_k)^2}\text{Re}\left[\delta A^l_{jk}\delta A^m_{kj}\right] + \sum_{\omega_j=\omega_k}\beta^2\frac{p_j + p_k}{2}\text{Re}\left[\delta A^l_{jk}\delta A^m_{kj}\right] \\
&= 2\beta^2\sum_{\omega_j\neq\omega_k}p_j\frac{(1 - p_k/p_j)^2}{(1 + p_k/p_j)\ln^2(p_k/p_j)}\text{Re}\left[\delta A^l_{jk}\delta A^m_{kj}\right] + \beta^2\sum_{\omega_j=\omega_k}\frac{p_j + p_k}{2}\text{Re}\left[\delta A^l_{jk}\delta A^m_{kj}\right].
\end{aligned} \tag{A16}$$

In Appendix II, we use this expression to upper- and lower-bound $\mathcal{F}_{ll}$.

II. BOUNDS ON THE QUANTUM FISHER INFORMATION

In this section, we upper- and lower-bound the diagonals of the quantum Fisher information matrix. That is, we prove Eqs. (6) and (7) from the main text. By Eq. (A16), the quantum Fisher information about a parameter μ_l is

$$\mathcal{F}_{ll} = 2\beta^2\sum_{\omega_j\neq\omega_k}p_j\frac{(1 - p_k/p_j)^2}{(1 + p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2 + \beta^2\sum_{\omega_j=\omega_k}\frac{p_j + p_k}{2}\left|\delta A^l_{jk}\right|^2. \tag{A17}$$

Upper bound in terms of $(\Delta A_l)^2$

If $x := p_k/p_j$, the first term in the quantum Fisher information [Eq. (A17)] depends on $\frac{(1-x)^2}{(1+x)\ln^2(x)}$. It will be convenient to upper-bound this fraction as $\frac{(1-x)^2}{(1+x)\ln^2(x)} \leq (1+x)c_1$, for some c_1 to be determined. Shifting the $(1+x)$ from the inequality's right-hand side to the left-hand side, we form a fraction $\frac{(1-x)^2}{(1+x)^2\ln^2(x)}$ that is maximized at $x = 1$. Furthermore, p_k/p_j comes closest to 1 for energy eigenstates whose energies are as close as possible: $x_{\text{max}} := e^{-\beta\min_{j,k}\{\omega_k - \omega_j\}} \equiv e^{-\beta g_{\text{min}}}$. We have defined $g_{\text{min}} := \min_{j,k}\{\omega_j - \omega_k\}$ as the Hamiltonian's minimum energy gap. Combining these observations, we choose

$$c_1(g_{\text{min}}) := \frac{(1 - x_{\text{max}})^2}{(1 + x_{\text{max}})^2\ln^2(x_{\text{max}})} = \frac{(1 - e^{-\beta g_{\text{min}}})^2}{(1 + e^{-\beta g_{\text{min}}})^2}\frac{1}{\beta^2 g^2_{\text{min}}} = \frac{\tanh^2(\beta g_{\text{min}}/2)}{\beta^2 g^2_{\text{min}}}\,. \tag{A18}$$

The limiting values of c_1, as a function of temperature, are

$$c_1(g_{\text{min}}) \approx \begin{cases} \frac{1}{\beta^2 g^2_{\text{min}}} & \text{for} \quad \beta g_{\text{min}} \gg 1 \\ \frac{1}{4} & \text{for} \quad \beta g_{\text{min}} \ll 1. \end{cases} \tag{A19}$$

Applying this choice and the general bound above to Eq. (A17), we bound the quantum Fisher information about a parameter μ_l:

$$\begin{aligned}
\mathcal{F}_{ll} &= 2\sum_{\omega_j\neq\omega_k}\beta^2 p_j\frac{(1-p_k/p_j)^2}{(1+p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2\\
&\leq 2c_1(g_{\min})\beta^2\sum_{\omega_j\neq\omega_k}p_j\left(1+\frac{p_k}{p_j}\right)\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2 \qquad \text{(A20a)}\\
&= 4c_1(g_{\min})\beta^2\sum_{\omega_j\neq\omega_k}p_j\left|\delta A^l_{jk}\right|^2+4c_1(g_{\min})\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2+[1-4c_1(g_{\min})]\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2\\
&= 4c_1(g_{\min})\beta^2\,\mathrm{Tr}\left(\rho\,[\delta A_l]^2\right)+[1-4c_1(g_{\min})]\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2\\
&= 4c_1(g_{\min})\beta^2\left(\Delta A_l\right)^2+[1-4c_1(g_{\min})]\beta^2\left(\Delta A^D_l\right)^2. \qquad \text{(A20b)}
\end{aligned}$$

We have defined $\Delta A=\sqrt{\langle A^2\rangle-\langle A\rangle^2}$ as the standard deviation of an operator A in the thermal state. Also, $A^D_l\coloneqq\sum_{\omega_j=\omega_k}\langle j|\,A_l\,|k\rangle\,|k\rangle\langle j|$ is the sum of the block-diagonal elements of the matrix that represents A_l relative to the energy eigenbasis. Since $0\leq c_1(g_{\min})\leq 1/4$ in Eq. (A20b), also,

$$\mathcal{F}_{ll}\leq\beta^2\left(\Delta A_l\right)^2. \qquad \text{(A21)}$$

We have proved Eq. (6a) in the main text. Bounds (A20b) and (A21) are saturated if A_l is diagonal relative to the energy eigenbasis.

Lower bound in terms of $(\Delta A_l)^2$

A similar derivation implies a lower bound on $\mathcal{F}_{ll}$. The function $\frac{(1-x)^2}{(1+x)^2\ln^2(x)}$ is minimized at $x=0$ and in the limit as $x\to\infty$. Moreover, x has a minimum value of $x_{\min}\coloneqq e^{-\beta\|H\|_s}$, where $\|H\|_s\coloneqq\max_j\omega_j-\min_j\omega_j$, and a maximum value of $x_{\max}\coloneqq e^{\beta\|H\|_s}$. Since $c_1(-\|H\|_s)=c_1(\|H\|_s)$,

$$\frac{(1-x)^2}{(1+x)\ln^2(x)}\geq(1+x)c_1(\|H\|_s). \qquad \text{(A22)}$$

Using Eqs. (A22) and (A17) leads to

$$\begin{aligned}
\mathcal{F}_{ll} &= 2\sum_{\omega_j\neq\omega_k}\beta^2 p_j\frac{(1-p_k/p_j)^2}{(1+p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2\\
&\geq 2c_1(\|H\|_s)\sum_{\omega_j\neq\omega_k}\beta^2 p_j\left(1+\frac{p_k}{p_j}\right)\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2\\
&= 4c_1(\|H\|_s)\beta^2\sum_{\omega_j\neq\omega_k}p_j\left|\delta A^l_{jk}\right|^2+4c_1(\|H\|_s)\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2+[1-4c_1(\|H\|_s)]\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2\\
&= 4c_1(\|H\|_s)\beta^2\,\mathrm{Tr}\left(\rho\left(\delta A_l\right)^2\right)+[1-4c_1(\|H\|_s)]\beta^2\sum_j p_j\left|\delta A^l_{jk}\right|^2\\
&= 4c_1(\|H\|_s)\beta^2\left(\Delta A_l\right)^2. \qquad \text{(A23)}
\end{aligned}$$

We have proved Eq. (6b) in the main text.

Upper bound in terms of $(\Delta A_l)^2-\frac{1}{2}\|[\sqrt{\rho},A_l]\|_2^2$

We can obtain a distinct upper bound that depends on the Wigner-Yanase skew information. Beginning with Eq. (A17), we split the sum over $\omega_j\neq\omega_k$ into $\omega_j<\omega_k$ and $\omega_j>\omega_k$ terms. We can then collapse terms due to the symmetry with respect to the interchange $p_j \leftrightarrow p_k$:

$$\begin{aligned}
\mathcal{F}_{ll} &= 2\sum_{\omega_j\neq\omega_k}\beta^2 p_j\frac{(1-p_k/p_j)^2}{(1+p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2\\
&= 2\sum_{\omega_j>\omega_k}\beta^2 p_j\frac{(1-p_k/p_j)^2}{(1+p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2+2\sum_{\omega_j<\omega_k}\beta^2 p_j\frac{(1-p_k/p_j)^2}{(1+p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}\frac{p_j+p_k}{2}\left|\delta A^l_{jk}\right|^2\\
&= 4\sum_{\omega_j<\omega_k}\beta^2 p_j\frac{(1-p_k/p_j)^2}{(1+p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2. \qquad \text{(A24)}
\end{aligned}$$

Assume that the energies ω_j are in ascending order, such that $x_{\min}\le x:=p_k/p_j\le 1$, for $j<k$. The first term in (A24) contains a factor of the form $\frac{(1-x)^2}{(1+x)\ln^2(x)}$, which obeys the upper bound $\frac{(1-x)^2}{(1+x)\ln^2(x)}\le c_2\sqrt{x}$ for $0\le x\le 1$, for some c_2. The minimum value of x, at an inverse temperature β, is $x_{\min}:=\min_{\{j,k\}}p_k/p_j=\min_{\{j,k\}}e^{-\beta(\omega_k-\omega_j)}=e^{-\beta\|H\|_s}$. Therefore,

$$c_2 := \frac{1}{\beta^2\|H\|_s^2}e^{\frac{1}{2}\beta\|H\|_s}\frac{\left(1-e^{-\beta\|H\|_s}\right)^2}{1+e^{-\beta\|H\|_s}}=\frac{2\sinh(\beta\|H\|_{\rm s}/2)\tanh(\beta\|H\|_{\rm s}/2)}{\beta^2\|H\|_{\rm s}^2}\ge 0.42. \qquad \text{(A25)}$$

The inequality holds because $2\sinh(x/2)\tanh(x/2)/x^2\ge 0.42$ (as one can check using, e.g., Mathematica). The limiting values of c_2, as a function of temperature, are

$$c_2\approx\begin{cases} e^{\frac{1}{2}\beta\|H\|_s}/(\beta^2\|H\|_s^2)\,, & \text{for}\quad \beta\|H\|_s\gg 1,\\ 1/2\,, & \text{for}\quad \beta\|H\|_s\ll 1.\end{cases} \qquad \text{(A26)}$$

Let us apply Eq. (A25), with the general bound above, to Eq. (A24):

$$\begin{aligned}
\mathcal{F}_{ll} &= 4\sum_{\omega_j<\omega_k}\beta^2 p_j\frac{(1-p_k/p_j)^2}{(1+p_k/p_j)\ln^2(p_k/p_j)}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2\\
&\le 4c_2\sum_{\omega_j<\omega_k}\beta^2 p_j\sqrt{\frac{p_k}{p_j}}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2\\
&= 2c_2\beta^2\sum_{\omega_j<\omega_k}\sqrt{p_j}\sqrt{p_k}\left|\delta A^l_{jk}\right|^2+2c_2\beta^2\sum_{\omega_j>\omega_k}\sqrt{p_j}\sqrt{p_k}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2\\
&\le 2c_2\beta^2\sum_{\omega_j<\omega_k}\sqrt{p_j}\sqrt{p_k}\left|\delta A^l_{jk}\right|^2+2c_2\beta^2\sum_{\omega_j>\omega_k}\sqrt{p_j}\sqrt{p_k}\left|\delta A^l_{jk}\right|^2+\frac{c_2}{0.42}\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2 \qquad \text{(A27a)}\\
&\le 2.4c_2\beta^2\sum_{\omega_j<\omega_k}\sqrt{p_j}\sqrt{p_k}\left|\delta A^l_{jk}\right|^2+2.4c_2\beta^2\sum_{\omega_j>\omega_k}\sqrt{p_j}\sqrt{p_k}\left|\delta A^l_{jk}\right|^2+2.4c_2\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2 \qquad \text{(A27b)}\\
&= 2.4c_2\,\beta^2\,\mathrm{Tr}\left(\sqrt{\rho}\delta A_l\sqrt{\rho}\delta A_l\right). \qquad \text{(A27c)}
\end{aligned}$$

In Eqs. (A27a) and (A27b), we invoked $1\le c_2/0.42\le 2.4c_2$. Since $\mathrm{Tr}\left(\sqrt{\rho}\delta A_l\sqrt{\rho}\delta A_l\right)=(\Delta A_l)^2-\frac{1}{2}\left\|\left[\sqrt{\rho},A_l\right]\right\|_2^2$, we have proved the second upper bound on $\mathcal{F}_{ll}$, Eq. (7) in the main text.

Lower bound in terms of $(\Delta A_l)^2-\frac{1}{2}\|[\sqrt{\rho},A_l]\|_2^2$

Our general expression $\frac{(1-x)^2}{(1+x)\ln^2(x)}$ obeys the upper bound $\sqrt{x}/2.5\le\frac{(1-x)^2}{(1+x)\ln^2(x)}$. Applying this bound to Eq. (A17) yields

$$\begin{aligned}
\mathcal{F}_{ll} &\ge \frac{2}{2.5}\sum_{\omega_j\neq\omega_k}\beta^2 p_j\sqrt{\frac{p_k}{p_j}}\left|\delta A^l_{jk}\right|^2+\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2\\
&\ge 0.8\,\beta^2\sum_{\omega_j\neq\omega_k}\sqrt{p_j}\sqrt{p_k}\left|\delta A^l_{jk}\right|^2+0.8\,\beta^2\sum_{\omega_j=\omega_k}p_j\left|\delta A^l_{jk}\right|^2\\
&= 0.8\,\beta^2\,\mathrm{Tr}\left(\sqrt{\rho}\delta A_l\sqrt{\rho}\delta A_l\right). \qquad \text{(A28)}
\end{aligned}$$

This result completes the proof of Eq. (7) in the main text.

III. NONCOMMUTATIVITY AND PARAMETER ESTIMATION

In this section, we discuss the role of noncommutativity in parameter estimation. In Eq. (10) of the main text, we presented an upper and a lower bound on the optimal relative estimation error $\sqrt{\mathrm{var}_{\mathrm{opt}}(\hat{\mu}_l)}/|\mu_l|$ with which a parameter μ_l can be estimated from $\mathcal{N}$ copies of a thermal state. We reproduce the bound here for convenience:

$$\frac{1}{\sqrt{2.4c_2}\beta\sqrt{\mathcal{N}}\left((\Delta H_l)^2 - \frac{1}{2}\left\|\left[\sqrt{\rho}, H_l\right]\right\|_2^2\right)^{1/2}} \leq \frac{\sqrt{\mathrm{var}_{\mathrm{opt}}(\hat{\mu}_l)}}{|\mu_l|} \leq \frac{1}{\sqrt{0.8}\beta\sqrt{\mathcal{N}}\left((\Delta H_l)^2 - \frac{1}{2}\left\|\left[\sqrt{\rho}, H_l\right]\right\|_2^2\right)^{1/2}}. \tag{A29a}$$

Recall that H_l is the Hamiltonian term that contains the parameter μ_l. Due to the $\|[\sqrt{\rho}, H_l]\|_2$, noncommutativity between the state and H_l negatively impacts one's ability to estimate μ_l. Here, we elaborate on the role of noncommutativity in estimating a parameter from Hamiltonian evolution (as opposed to from a thermal state).

In the Hamiltonian-evolution setting, we estimate μ_l by evolving a probe state under a Hamiltonian

$$H = H_l + H' \equiv \mu_l A_l + H' \tag{A30}$$

for some time t. A_l is a Hermitian matrix (the generator of translations associated with μ_l). H' contains all the (possibly time-dependent) terms independent of μ_l. In our setting, $H' = \sum_{j\neq l} H_j$. The time-evolved state $\varrho(t)$ depends on μ_l. One can estimate μ_l from properly chosen measurements of copies of $\varrho(t)$. The minimum achievable variance is bounded in the single-parameter quantum Cramér-Rao bound, Eq. (5) in the main text.

The minimal variance can be achieved with a pure probe state $\varrho = |\psi\rangle\langle\psi|$. We have defined $|\psi\rangle = (|\lambda_{\mathrm{max}}\rangle + |\lambda_{\mathrm{min}}\rangle)/\sqrt{2}$, $|\lambda_{\mathrm{max}}\rangle$ and $|\lambda_{\mathrm{min}}\rangle$ denoting the eigenstates associated with the maximum and minimum A_l eigenvalues, λ_{max} and λ_{min} [2]. Suppose that $|\lambda_{\mathrm{max}}\rangle$ and $|\lambda_{\mathrm{min}}\rangle$ are H' eigenstates associated with unit eigenvalues:

$$H'\,|\lambda_{\mathrm{max}}\rangle = |\lambda_{\mathrm{max}}\rangle\,, \quad \text{and} \quad H'\,|\lambda_{\mathrm{min}}\rangle = |\lambda_{\mathrm{min}}\rangle\,. \tag{A31}$$

Evolution under H yields a final state $|\psi(t)\rangle = (|\lambda_{\mathrm{max}}\rangle + e^{(\lambda_{\mathrm{max}}-\lambda_{\mathrm{min}})\mu_l t}\,|\lambda_{\mathrm{min}}\rangle)/\sqrt{2}$, from which μ_l can be extracted with a variance $\sim [t(\lambda_{\mathrm{max}} - \lambda_{\mathrm{min}})]^{-2}$, which is optimal [3].

The conditions (A31), under which this optimal scheme works, can be replaced with the weaker condition $[H', H_l] = [H, H_l] = 0$. From here, we see the connection to Eq. (A29a): for Gibbs states ρ, if $[H, H_l] = 0$, then $[\sqrt{\rho}, H_l] = 0$. Consequently, we see a direct formal connection between the fact that noncommutativity of H_l with H negatively impacts the estimation of μ_l through Hamiltonian evolution and the fact that $[\sqrt{\rho}, H_l] \neq 0$ negatively impacts Hamiltonian learning from Gibbs states.

IV. QUANTUM FISHER INFORMATION OF A TWO-LEVEL SYSTEM

In this section, we calculate the quantum Fisher information about a parameter in a single-qubit Hamiltonian. Consider the Hamiltonian

$$H = \Omega_x\sigma_x + \Omega_z\sigma_z + \mu\sigma_z =: \vec{v}\cdot\vec{\sigma}. \tag{A32}$$

We have defined the vector $\vec{v} = (\Omega_x, \Omega_z + \mu)$ with the norm $v := \sqrt{\Omega_x^2 + (\Omega_z+\mu)^2}$, and $\vec{\sigma} = (\sigma_x, \sigma_z)$ is a vector of Pauli matrices. We aim to estimate μ, so $A = \sigma_z$. The thermal state is

$$\rho = \frac{e^{-\beta v\,\vec{v}\cdot\vec{\sigma}/v}}{Z} = \frac{\cosh(\beta v)\mathbb{1} - \sinh(\beta v)\vec{v}\cdot\vec{\sigma}/v}{Z}, \tag{A33}$$

where $Z = 2\cosh(\beta v)$. The Hamiltonian has a seminorm $\|H\|_s = 2v$.

We directly calculate the Wigner-Yanase skew information, using $\sqrt{\rho} = e^{-\beta H/2}/\sqrt{Z}$, $A = \sigma_z$, and Eq. (A33):

$$\begin{aligned}
\tfrac{1}{2}\left\|[\sqrt{\rho}, A]\right\|_2^2 &= \frac{1}{2Z}\left\|\sinh(\beta v/2)/v[\vec{v}\cdot\vec{\sigma}, \sigma_z]\right\|_2^2 = \frac{\sinh^2(\beta v/2)}{2v^2 Z}\left\|\Omega_x[\sigma_x, \sigma_z]\right\|_2^2 \\
&= \frac{\sinh^2(\beta v/2)}{2v^2 Z}\left\| -2i\Omega_x\sigma_y\right\|_2^2 = 2\frac{\sinh^2(\beta v/2)}{v^2 Z}\,\mathrm{Tr}\left(\left[-i\Omega_x\sigma_y\right]\left[i\Omega_x\sigma_y\right]\right) \\
&= 4\frac{\sinh^2(\beta v/2)}{v^2 Z}\Omega_x^2 = 4\frac{\frac{1}{2}(\cosh(\beta v) - 1)}{v^2 Z}\Omega_x^2 = 2\frac{\cosh(\beta v) - 1}{2v^2\cosh(\beta v)}\Omega_x^2 \\
&= \frac{1 - \mathrm{sech}(\beta v)}{v^2}\Omega_x^2\,.
\end{aligned} \tag{A34}$$

The thermal variance in σ_z is

$$(\Delta A)^2 = \mathrm{Tr}\,(\rho) - [\mathrm{Tr}\,(\rho\sigma_z)]^2 = 1 - \left[-\frac{1}{vZ}\sinh(\beta v)\,\mathrm{Tr}\,(\vec{v}\cdot\vec{\sigma}\sigma_z)\right]^2 = 1 - \frac{4\sinh^2(\beta v)}{v^2Z^2}(\Omega_z+\mu)^2$$
$$= 1 - \frac{\tanh^2(\beta v)}{v^2}(\Omega_z+\mu)^2\,. \tag{A35}$$

Subtracting Eq. (A34) from (A35) yields

$$(\Delta A)^2 - \tfrac{1}{2}\big\|[\sqrt{\rho},A]\big\|_2^2 = 1 - \frac{\tanh^2(\beta v)}{v^2}(\Omega_z+\mu)^2 - \frac{1-\mathrm{sech}(\beta v)}{v^2}\Omega_x^2\,. \tag{A36}$$

We can approximate this expression at high and low temperatures. If the temperature is high ($\beta v \ll 1$), then $\mathrm{sech}(\beta v) \approx 1-(\beta v)^2/2$, and $\tanh(\beta v)\approx \beta v$. Therefore,

$$(\Delta A)^2 - \tfrac{1}{2}\big\|[\sqrt{\rho},A]\big\|_2^2 \approx 1 - \beta^2(\Omega_z+\mu)^2 - \frac{\beta^2}{2}\Omega_x^2 = 1 - \frac{\beta^2}{2}(\Omega_z+\mu)^2 - \frac{\beta^2}{2}v^2\,. \tag{A37}$$

If the temperature is small, ($\beta v \gg 1$), then $\mathrm{sech}(\beta v)\approx 2e^{-\beta v}$, and $\tanh(\beta v)\approx 1$. Therefore,

$$(\Delta A)^2 - \tfrac{1}{2}\big\|[\sqrt{\rho},A]\big\|_2^2 \approx 1 - \frac{1}{v^2}(\Omega_z+\mu)^2 - \frac{1}{v^2}\Omega_x^2 + \frac{2e^{-\beta v}}{v^2}\Omega_x^2 = \frac{2e^{-\beta v}}{v^2}\Omega_x^2. \tag{A38}$$

Meanwhile, $c_2 \approx 1/2$ at high temperature ($\beta v \ll 1$), whereas $c_2 \approx e^{\beta\|H\|_s/2}/(\beta\|H\|_s)^2 = e^{\beta v}/(2\beta v)^2$ for $\beta v \gg 1$.

By Eq. (A37), at high temperatures ($\beta\|H\|_s \ll 1$) the bounds (7) in the main text become

$$\mathcal{F} \le 2.4c_2\,\beta^2\,\mathrm{Tr}\,(\sqrt{\rho}\delta A\sqrt{\rho}\delta A) \approx \frac{2.4}{2}\beta^2\left(1-\frac{\beta^2}{2}(\Omega_z+\mu)^2-\frac{\beta^2}{2}v^2\right) \approx 1.2\beta^2, \tag{A39a}$$

$$\mathcal{F} \ge 0.8\,\beta^2\,\mathrm{Tr}\,(\sqrt{\rho}\delta A\sqrt{\rho}\delta A) \approx 0.8\,\beta^2\left(1-\frac{\beta^2}{2}(\Omega_z+\mu)^2-\frac{\beta^2}{2}v^2\right) \approx 0.8\,\beta^2. \tag{A39b}$$

By Eq. (A38), at low temperature ($\beta\|H\|_s \gg 1$) the bounds (7) become

$$\mathcal{F} \le 2.4c_2\,\beta^2\,\mathrm{Tr}\,(\sqrt{\rho}\delta A\sqrt{\rho}\delta A) \approx 2.4\frac{e^{\beta v}}{4v^2}\left(\frac{2e^{-\beta v}}{v^2}\Omega_x^2\right) = 1.2\frac{\Omega_x^2}{v^4}, \tag{A40a}$$

$$\mathcal{F} \ge 0.8\,\beta^2\,\mathrm{Tr}\,(\sqrt{\rho}\delta A\sqrt{\rho}\delta A) \approx 0.8\,\beta^2\left(\frac{2e^{-\beta v}}{v^2}\Omega_x^2\right) = 1.6\,\frac{\beta^2e^{-\beta v}\Omega_x^2}{v^2}\,. \tag{A40b}$$

We want to compare these bounds with the values of the quantum Fisher information. Define the σ_z eigenstates such that $\sigma_z\,|1\rangle = |1\rangle$ and $\sigma_z\,|0\rangle = -\,|0\rangle$. $H = \vec{v}.\vec{\sigma}$ has the eigenvectors

$$|+\rangle = \frac{1}{\sqrt{2v\big(\Omega_z+\mu+v\big)}}\Big((\Omega_z+\mu+v)\,|1\rangle + \Omega_x\,|0\rangle\Big) \quad \text{and} \tag{A41}$$

$$|-\rangle = \frac{1}{\sqrt{2v\big(\Omega_z+\mu+v\big)}}\Big(-\Omega_x\,|1\rangle + (\Omega_z+\mu+v)\,|0\rangle\Big), \tag{A42}$$

corresponding to eigenvalues $\pm v$. By the expression (A17) for the quantum Fisher information, for a qubit,

$$\mathcal{F} = 2\sum_{\omega_j\neq\omega_k}\frac{(p_j-p_k)^2}{(p_j+p_k)(\omega_j-\omega_k)^2}\big|\delta A^l_{jk}\big|^2 + \sum_{\omega_j=\omega_k}\beta^2p_j\big|\langle j|\delta A_l|k\rangle\big|^2$$
$$= \frac{4}{Z}\left(\frac{(e^{-\beta v}-e^{\beta v})^2}{(e^{-\beta v}+e^{\beta v})(2v)^2}\right)\big|\langle +|\,\delta X\,|-\rangle\big|^2 + \beta^2\frac{e^{-\beta v}}{Z}\big|\langle +|\delta X|+\rangle\big|^2 + \beta^2\frac{e^{\beta v}}{Z}\big|\langle -|\delta X|-\rangle\big|^2. \tag{A43}$$

We evaluate this expression using

$$\langle +|\delta A|-\rangle = \frac{1}{2v\big(\Omega_z+\mu+v\big)}\Big(-\Omega_x(\Omega_z+\mu+v) - \Omega_x(\Omega_z+\mu+v)\Big) = -\frac{\Omega_x}{v}\,, \tag{A44a}$$

$$\langle +|\delta A|+\rangle = \frac{1}{2v\big(\Omega_z+\mu+v\big)}\Big((\Omega_z+\mu+v)^2-\Omega_x^2\Big) - \langle\sigma_z\rangle \coloneqq a - \langle\sigma_z\rangle, \quad \text{and} \tag{A44b}$$

$$\langle -|\delta A|-\rangle = \frac{1}{2v\big(\Omega_z+\mu+v\big)}\Big(\Omega_x^2-(\Omega_z+\mu+v)^2\Big) - \langle\sigma_z\rangle \coloneqq -a - \langle\sigma_z\rangle. \tag{A44c}$$

We have defined $a := \frac{1}{2v\left(\Omega_z+\mu+v\right)}\left((\Omega_z+\mu+v)^2-\Omega_x^2\right) = \frac{(\Omega_z+\mu)}{v\left(\Omega_z+\mu+v\right)}\left(\Omega_z+\mu+v\right) = \frac{(\Omega_z+\mu)}{v}$. Next, we evaluate Eq. (A43) using Eq. (A44), $Z = 2\cosh(\beta v)$, and $\langle\sigma_z\rangle = -\frac{\tanh(\beta v)}{v}(\Omega_z+\mu)$ from Eq. (A35):

$$
\begin{aligned}
\mathcal{F} &= \frac{4}{Z}\left(\frac{(e^{-\beta v}-e^{\beta v})^2}{(e^{-\beta v}+e^{\beta v})(2v)^2}\right)\frac{\Omega_x^2}{v^2} + \left(\beta^2\frac{e^{-\beta v}}{Z}+\beta^2\frac{e^{\beta v}}{Z}\right)(a^2+\langle\sigma_z\rangle^2) + \left(\beta^2\frac{e^{-\beta v}}{Z}-\beta^2\frac{e^{\beta v}}{Z}\right)(-2a\langle\sigma_z\rangle) \\
&= \frac{(-e^{-\beta v}+e^{\beta v})}{2\cosh(\beta v)}\tanh(\beta v)\frac{\Omega_x^2}{v^4} + \beta^2\frac{e^{-\beta v}+e^{\beta v}}{2\cosh(\beta v)}(a^2+\langle\sigma_z\rangle^2) - \beta^2\frac{e^{-\beta v}-e^{\beta v}}{2\cosh(\beta v)}2a\langle\sigma_z\rangle \\
&= \tanh^2(\beta v)\frac{\Omega_x^2}{v^4} + \beta^2(a^2+\langle\sigma_z\rangle^2) + 2\beta^2\tanh(\beta v)a\langle\sigma_z\rangle \\
&= \tanh^2(\beta v)\frac{\Omega_x^2}{v^4} + \beta^2\left(\frac{(\Omega_z+\mu)^2}{v^2}+\tanh^2(\beta v)\frac{(\Omega_z+\mu)^2}{v^2}\right) - 2\beta^2\tanh^2(\beta v)\frac{(\Omega_z+\mu)^2}{v^2} \\
&= \tanh^2(\beta v)\frac{\Omega_x^2}{v^4} + \beta^2\frac{(\Omega_z+\mu)^2}{v^2} - \beta^2\tanh^2(\beta v)\frac{(\Omega_z+\mu)^2}{v^2} \\
&= \tanh^2(\beta v)\frac{\Omega_x^2}{v^4} + \beta^2\frac{(\Omega_z+\mu)^2}{v^2}\left(1-\tanh^2(\beta v)\right).
\end{aligned} \tag{A45}
$$

Using that $\tanh(x)\approx x$ for $x << 1$ and that $\tanh(x)\approx 1$ for $x >> 1$ yields

$$
\mathcal{F} \approx \frac{\beta^2\Omega_x^2}{v^2} + \frac{\beta^2(\Omega_z+\mu)^2}{v^2}(1-\beta^2v^2) \approx \beta^2, \quad \text{for } \beta\|H\|_s \ll 1, \quad \text{and} \tag{A46a}
$$

$$
\mathcal{F} \approx \frac{\Omega_x^2}{v^4}, \quad \text{for } \beta\|H\|_s \gg 1. \tag{A46b}
$$

Let us compare the high-temperature upper bound (A39a) with the approximate value (A46a), as well as the low-temperature upper bound (A40a) with the approximate value (A46b). The main-text upper bound (7) is saturable, to within a constant multiplicative factor, in both temperature regimes. Together with the Cramér-Rao bound, our bounds imply that

$$
\text{var}_{\text{opt}}(\hat{\mu}_l) \approx \begin{cases} \frac{1}{\mathcal{N}\beta^2}, & \text{for } \beta\|H\|_s \ll 1, \\ \frac{\|H\|_s^4}{16\mathcal{N}\Omega_x^2}, & \text{for } \beta\|H\|_s \gg 1. \end{cases} \tag{A47}
$$

V. COMPARISONS OF BOUNDS ON THE QUANTUM FISHER INFORMATION

In this section, we calculate quantum Fisher information in a spin-chain example. We compare the exact value with our bounds, Eqs. (6) and (7) in the main text. We reproduce the bounds here for convenience:

$$
\mathcal{F}_{ll} \le \beta^2\,(\Delta A_l)^2, \tag{A48a}
$$

$$
\mathcal{F}_{ll} \ge 4\beta^2 c_1\,(\Delta A_l)^2, \tag{A48b}
$$

and

$$
\mathcal{F}_{ll} \le 2.4\,c_2\,\beta^2\left((\Delta A_l)^2 - \tfrac{1}{2}\left\|[\sqrt{\rho},A_l]\right\|_2^2\right), \tag{A49a}
$$

$$
\mathcal{F}_{ll} \ge 0.8\,\beta^2\left((\Delta A_l)^2 - \tfrac{1}{2}\left\|[\sqrt{\rho},A_l]\right\|_2^2\right). \tag{A49b}
$$

We also compare the bounds to the one derived in Ref. [4]:

$$
\mathcal{F}_{ll} \le \beta^2\int_0^1 \text{Tr}\left(\rho^a\delta A_l\rho^{1-a}\delta A_l\right)da, \tag{A50}
$$

with $\delta A_l := A_l - \langle A_l\rangle$.

Consider estimating the parameter μ from the thermal state of a spin chain. We consider a one-dimensional chain composed of n spin-1/2 systems, with the Hamiltonian

$$
H = \mu\sum_{j=1}^{n}\sigma_z^j + \lambda\sum_{j=1}^{n-1}\sigma_x^j\otimes\sigma_x^{j+1} =: \mu A_\mu + \lambda A_\lambda. \tag{A51}
$$

A_μ and A_λ are the operators that multiply the parameters μ and λ.

Figure 1 compares the quantum Fisher information about μ with the upper and lower bounds in Eqs. (6) and (7), and with the upper bound in Ref. [4]. We simulate $n = 5$ spins. The figure shows that the bounds are distinct and that none of them is tighter than another in all regimes: in each subfigure, the two blue curves (upper bounds derived in this Letter) and the orange star plot (bound in Ref. [4]) cross, as do the two red curves (lower bounds derived in this Letter). However, the bounds are always obeyed: the black curve (exactly calculated quantum Fisher information) always lies below the blue curves and orange star plot (upper bounds) and above the red curves (lower bounds).

(a) Quantum Fisher information vs. β ($\lambda/\mu = 5$) (b) Quantum Fisher information vs. λ ($\beta\mu = 0.1$)

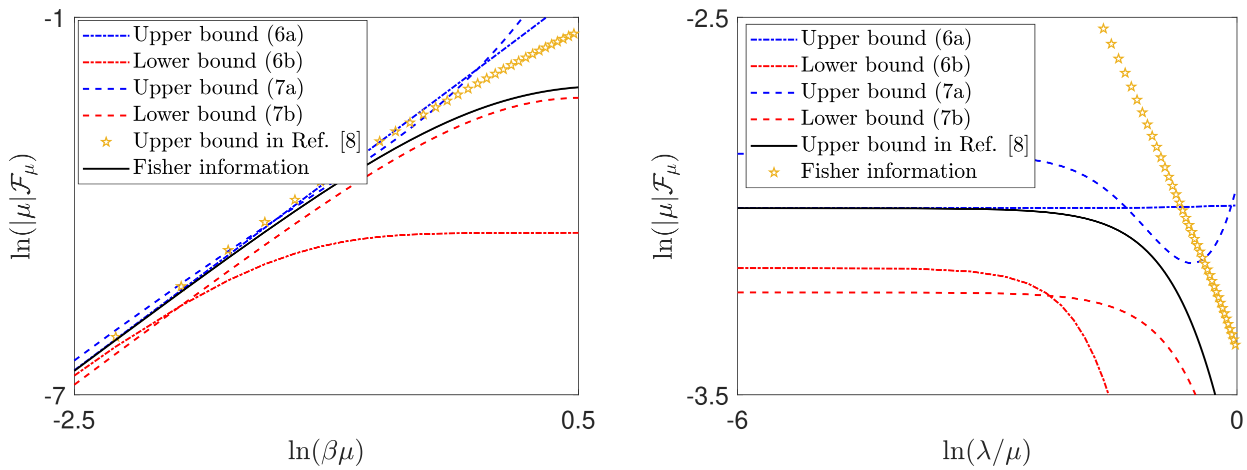

FIG. 1. **Comparisons of bounds on the quantum Fisher information.** The figure shows log-log plots of the quantum Fisher information $\mathcal{F}_\mu$ about parameter μ, as a function of the inverse temperature (left) and as a function of the parameter Ω (right). The 5-qubit system has the Hamiltonian $H = \mu \sum_{j=1}^{n} \sigma_z^j + \lambda \sum_{j=1}^{n-1} \sigma_x^j \otimes \sigma_x^{j+1} := \mu A_\mu + \lambda A_\lambda$. The plots also depict the upper and lower bounds in Eqs. (6) and (7), and the upper bound derived in Ref. [4]. Each plot illustrates (where a red line crosses a red line or a blue line or orange stars cross) how different bounds can be tighter in different regimes.

VI. A MODEL THAT CAN BEAT THE STANDARD QUANTUM LIMIT

Here, we prove that the Hamiltonian

$$H = \mu \sum_{j=1}^{n} \left(\sigma_z^j + 1\right) - \lambda \bigotimes_{j=1}^{n} n\sigma_x^j \equiv H_\mu + H_\lambda, \tag{A52}$$

considered in the main text has the GHZ state as its unique ground state when $\mu \ll \lambda$. We also prove that $(\Delta H_\mu)^2 \approx \mu^2 n^2$ for $\beta\lambda n \gg 1$.

For convenience, we shift H_μ by an irrelevant factor of $(\mu n)\mathbb{1}$ so that we consider the new Hamiltonian

$$\widetilde{H} = \mu \sum_{j=1}^{n} \left(\sigma_z^j + 1 + n\right) - \lambda \bigotimes_{j=1}^{n} n\sigma_x^j \equiv \widetilde{H}_\mu + H_\lambda. \tag{A53}$$

In the computational basis, labeled by bit strings $s \in \{0,1\}^n$, this Hamiltonian is block diagonal, with 2^{n-1} blocks of dimension two spanned by pairs of computational basis states $\{|s\rangle, |\overline{s}\rangle\}$. Here, $\overline{s}$ denotes the complement of s—i.e. $\overline{s}_j = s_j + 1 \pmod 2$. These blocks, each labeled by a bitstring s, take the form

$$\widetilde{H}_s = \begin{bmatrix} \mu z & -\lambda n \\ -\lambda n & -\mu z \end{bmatrix}, \tag{A54}$$

where $z \equiv 2|\overline{s}| - n$ and $|\overline{s}|$ denotes the Hamming weight (i.e. the number of ones) of the bitstring $\overline{s}$. Note that we have two distinct, but equivalent, choices of the bitstring s that labels each block. Also, $z \in [-n, n]$.

Each such block can be simply diagonalized, and, thus, so can $\widetilde{H} = \bigoplus_s \widetilde{H}_s$. The eigenvalues of Eq. (A54) are $\pm\sqrt{\mu^2 z^2 + \lambda^2 n^2}$. Consequently, the minimum eigenvalue of $\widetilde{H}$ occurs for the block where $|\bar{s}| = 0$ (under a different, but equivalent, choice of labeling this block, $|\bar{s}| = n$). The associated minimum eigenvalue is $-n\sqrt{\mu^2 + \lambda^2}$ and the corresponding eigenstate (the ground state of $\widetilde{H}$) is

$$|\mathrm{gs}\rangle \propto -\left(\frac{\mu - \sqrt{\mu^2 + \lambda^2}}{\lambda}\right)|s\rangle + |\bar{s}\rangle\,. \tag{A55}$$

Consequently, for $\mu/\lambda \ll 1$, it holds that $|\mathrm{gs}\rangle \propto |s\rangle + |\bar{s}\rangle$, which is precisely the GHZ state $|\Phi\rangle$.

For the GHZ state, it holds that $\Delta\widetilde{H}_\mu = \mu n$. Consequently, we might expect, at least at low temperatures, that the thermal states of this model might also exhibit estimation errors that decrease faster than the standard quantum limit. This expectation can be analytically validated. In particular, a general thermal state takes the form

$$\rho = Z_\beta^{-1} e^{-\beta\widetilde{H}} = Z_\beta^{-1} \bigoplus_s e^{-\beta\widetilde{H}_s}, \tag{A56}$$

where Z_β is the partition function. It is easy to evaluate

$$e^{-\beta\widetilde{H}_s} = \cosh(\beta\sqrt{\mu^2 z^2 + \lambda^2 n^2})\mathbb{1} - \sinh(\beta\sqrt{\mu^2 z^2 + \lambda^2 n^2})\frac{(\mu z)\sigma_z - (\lambda n)\sigma_x}{\sqrt{\mu^2 z^2 + \lambda^2 n^2}}. \tag{A57}$$

Consequently,

$$Z_\beta = 2\sum_s \cosh(\beta\sqrt{\mu^2 z^2 + \lambda^2 n^2}). \tag{A58}$$

We can evaluate the variance $(\Delta\widetilde{H}_\mu)^2$ as $\sum_s (\Delta\widetilde{H}_{\mu,s})^2$ where $H_{\mu,s} = \mu z \sigma_z$ is the block of H_μ labeled by the bit string s. In particular,

$$\begin{aligned}\left(\Delta\widetilde{H}_{\mu,s}\right)^2 &= \mathrm{Tr}\left[\rho_s \widetilde{H}_{\mu,s}^2\right] - \mathrm{Tr}\left[\rho_s \widetilde{H}_{\mu,s}\right]^2 \\ &= Z_\beta^{-1}\left[2\mu^2 z^2 \cosh(\beta\sqrt{\mu^2 z^2 + \lambda^2 n^2}) - \frac{4\mu^4 z^4 \sinh^2(\beta\sqrt{\mu^2 z^2 + \lambda^2 n^2})}{Z_\beta(\mu^2 z^2 + \lambda^2 n^2)}\right],\end{aligned} \tag{A59}$$

where we used that $\rho_s = Z_\beta^{-1} e^{-\beta\widetilde{H}_s}$ (note, $\rho = \bigoplus_s \rho_s$).

Asymptotically in $\beta\lambda n$, we only have to consider the $z = n$ block in Eqs. (A58)-(A59), as $\lim_{x\to\infty} \frac{\cosh((1+\epsilon)x)}{\cosh(x)} = \infty$ (also, $\lim_{x\to\infty} \frac{\sinh((1+\epsilon)x)}{\sinh(x)} = \infty$) for any $\epsilon > 0$. Consequently, asymptotically in $\beta\lambda n$,

$$\begin{aligned}\left(\Delta\widetilde{H}_\mu\right)^2 &\sim \frac{1}{2\cosh(\beta n\sqrt{\mu^2 + \lambda^2})}\left[2\mu^2 n^2 \cosh(\beta n\sqrt{\mu^2 + \lambda^2}) - \frac{2\mu^4 n^4 \sinh^2(\beta n\sqrt{\mu^2 + \lambda^2})}{n^2 \cosh(\beta n\sqrt{\mu^2 + \lambda^2})(\mu^2 + \lambda^2)}\right] \\ &= \mu^2 n^2 - \tanh^2(\beta n\sqrt{\mu^2 + \lambda^2})\frac{\mu^4 n^2}{\mu^2 + \lambda^2} \\ &\sim \mu^2 n^2\left(1 - \frac{\mu^2}{\mu^2 + \lambda^2}\right),\end{aligned} \tag{A60}$$

where, in the last line, we use that $\beta > 0$. Therefore, we find the scaling $\left(\Delta\widetilde{H}_\mu\right)^2 \sim \mu^2 n^2$ for $\beta\lambda n \gg 1$—up to subleading, constant factor contributions to the scaling that depend on λ. Since shifts by constants do not change the variance of an operator, this also implies that $(\Delta H_\mu)^2 \sim \mu^2 n^2$

VII. SATURABILITY OF THE MULTIPARAMETER CRAMÉR-RAO BOUND

In this section, we derive conditions under which the multiparameter Cramér-Rao bound is saturated. That is, we prove Eq. (12) in the main text. The multiparameter Cramér-Rao bound is saturable if and only if [1, 5]

$$\mathrm{Tr}\left(\rho[L_l, L_m]\right) = 0. \tag{A61}$$

To calculate this equation's left-hand side, we express the trace relative to the ρ eigenbasis. Relative to that eigenbasis, the symmetric logarithmic derivative is represented by a matrix with elements [1]

$$\langle j| L_l |k\rangle = 2\frac{\langle j| \partial_l \rho |k\rangle}{p_j + p_k} . \tag{A62}$$

After substituting into the trace, we invoke Eqs. (A14) and (A15):

$$\begin{aligned}
\mathrm{Tr}\left(\rho[L_l, L_m]\right) &= \sum_{jk} \Big(p_j \langle j| L_l |k\rangle \langle k| L_m |j\rangle - p_k \langle k| L_m |j\rangle \langle j| L_l |k\rangle \Big) = \sum_{jk} (p_j - p_k) \langle j| L_l |k\rangle \langle k| L_m |j\rangle \\
&= 4 \sum_{jk} \frac{(p_j - p_k)}{(p_j + p_k)^2} \langle j| \partial_l \rho |k\rangle \langle k| \partial_m \rho |j\rangle \\
&= 4 \sum_{\omega_j \neq \omega_k} \frac{(p_j - p_k)}{(p_j + p_k)^2} \langle j| \delta A_l |k\rangle \langle k| \delta A_m |j\rangle \frac{(p_j - p_k)^2}{(\omega_j - \omega_k)^2} \\
&+ 4\beta^2 \sum_{\omega_j = \omega_k} \frac{(p_j - p_k)}{(p_j + p_k)^2} \langle j| \delta A_l |k\rangle \langle k| \delta A_m |j\rangle \, p_j^2 \\
&= 4 \sum_{\omega_j \neq \omega_k} \frac{(p_j - p_k)^3}{(\omega_j - \omega_k)^2 (p_j + p_k)^2} \langle j| A_l |k\rangle \langle k| A_m |j\rangle .
\end{aligned} \tag{A63}$$

This expression and Eq. (A61) imply Eq. (12) in the main text.

In typical Hamiltonians, most parameters will not satisfy the rather stringent conditions (12) for saturation. They are satisfied, for example, when the operators A_l are diagonal relative to the energy eigenbasis. Hence the multiparameter Cramér-Rao bound is saturable when one is estimating the Hamiltonian eigenvalues ω_j.

The single parameter Cramér-Rao bound can be saturated with measurements in the eigenbasis of the symmetric logarithmic derivative L_l in Eq. (A64) [6]. Using Eqs. (A14) and (A15) into Eq. (A62), we find that

$$\begin{aligned}
L_l &= \sum_{\omega_j \neq \omega_k} 2\frac{\langle j| \partial_l \rho |k\rangle}{p_j + p_k} |j\rangle\langle k| + \sum_{\omega_j = \omega_k} 2\frac{\langle j| \partial_l \rho |k\rangle}{2p_j} |k\rangle\langle j| \\
&= 2 \sum_{\omega_j \neq \omega_k} \frac{(p_j - p_k)}{(p_j + p_k)(\omega_j - \omega_k)} \langle j| \delta A_l |k\rangle |j\rangle\langle k| - \beta \sum_{\omega_j = \omega_k} \langle j| \delta A_l |k\rangle |k\rangle\langle j| .
\end{aligned} \tag{A64}$$

Performing measurements on the eigenbasis of L_l would yield one protocol to saturate the Cramér-Rao bound.

VIII. COMPARISONS WITH THE HAMILTONIAN-LEARNING LITERATURE

In this section, we compare our bounds to earlier results concerning the Hamiltonian-learning problem. Two approaches to Hamiltonian learning are common: (i) the steady-state-based approach and (ii) the time-evolution-based approach. In the steady-state-based approach, one studies states ρ that are stationary with respect to evolution under the Hamiltonian H. These steady states satisfy the condition [7]

$$\partial_t \rho = -i\,[H, \rho] = 0.$$

Every Hamiltonian eigenstate is a steady state, as is the Gibbs state, $\frac{\exp(-\beta H)}{\mathrm{Tr}(\exp(-\beta H))}$. Several studies concern estimations of the Hamiltonian from eigenstates [7–14] or from Gibbs states [7, 15–18].

In the time-evolution-based approach, one analyzes the system's time evolution under the Hamiltonian. Several proposals concern learning the Hamiltonian from unitary dynamics [17, 19–22]. Experimental implementations [23, 24] of Hamiltonian-learning protocols have been carried out, too. In the Hamiltonian-learning problem, one aims to learn the Hamiltonian H from a physically relevant class of Hamiltonians, while minimizing the algorithm's run time and the number of copies of ρ. These two metrics are commonly known as sample complexity and time complexity, respectively.

In this work, we focus on learning about a Hamiltonian from Gibbs states. Our comparison of sample-complexity lower bounds with earlier works is presented in the context of the l_2 distance error, defined via $\epsilon = \left(\sum_{l=1}^{M} |\mu_l - \hat{\mu}_l|^2\right)^{\frac{1}{2}}$. Here, $\hat{\mu}_l$ denotes the estimate for μ_l. The rationale for this comparison criterion is due to our adoption of the related metric ϵ_{err}, defined via $\sum_{l=1}^{M} \mathrm{var}(\hat{\mu}_l) = \epsilon_{\mathrm{err}}^2$. We provide the following Lemma to relate the two error metrics.

Lemma 1. *For ϵ and ϵ_{err} defined as before, the following holds.*

1. *Prob* $\left(\epsilon^2 \geq a\right) \leq \frac{\epsilon_{err}^2}{a}$ *for all* $a > 0$.
2. *Prob* $\left(\left|\epsilon^2 - \epsilon_{err}^2\right| \geq a\right) \leq \frac{Var\left(\epsilon^2\right)}{a}$ *for all* $a > 0$.
3. *For any two real numbers b and a such that $b \geq a$, let Prob* $\left(a \leq \epsilon^2 \leq b\right) = 1$. *Then* $a \leq \epsilon_{err}^2 \leq b$.

Proof. (Proof of part 1) Note that $\epsilon_{\text{err}}^2 = \sum_{l=1}^{M} \text{var}\left(\hat{\mu}_l\right)$ and $\epsilon^2 = \sum_{l=1}^{M} \left(\hat{\mu}_l - \mu_l\right)^2$. Since $\hat{\mu}_l$ is an unbiased estimator for μ_l, we have $\mathbb{E}\left(\hat{\mu}_l\right) = \mu_l$ for $l \in \{1, 2, \cdots, M\}$.

Thus,

$$\text{var}\left(\hat{\mu}_l\right) = \mathbb{E}\left[\left(\hat{\mu}_l - \mathbb{E}\left(\hat{\mu}_l\right)\right)^2\right] = \mathbb{E}\left[\left(\hat{\mu}_l - \mu_l\right)^2\right]. \tag{A65}$$

Let us define a new random variable, $V_l = \left(\hat{\mu}_l - \mu_l\right)^2$. Thus, using Eq. (A65), we get

$$\epsilon_{\text{err}}^2 = \sum_{l=1}^{M} \mathbb{E}\left[V_l\right] \tag{A66}$$

and

$$\epsilon^2 = \sum_{l=1}^{M} V_l. \tag{A67}$$

Since V_l is a non-negative random variable, using Markov's inequality with Eqs. (A66) and (A67), we get

$$\text{Prob}\left(\epsilon^2 \geq a\right) \leq \frac{\epsilon_{\text{err}}^2}{a}$$

for $a > 0$. This completes the proof of part 1.

(Proof of part 2) If Y is a random variable with $\mathbb{E}\left(Y\right) = \alpha$ and $\text{Var}\left(Y\right) = \beta$, Chebyshev's inequality says

$$\text{Prob}\left(\left|Y - \alpha\right| \geq a\right) \leq \frac{\beta}{a} \quad \forall a > 0.$$

Applying Chebyshev's inequality to $Y = \epsilon^2 = \sum_{l=1}^{M} V_l$, we get

$$\text{Prob}\left(\left|\epsilon^2 - \alpha\right| \geq a\right) \leq \frac{\text{Var}\left(\epsilon^2\right)}{a} \quad \forall\ a > 0. \tag{A68}$$

Since expectation is linear, we have

$$\alpha = \mathbb{E}\left(\epsilon^2\right) = \epsilon_{\text{err}}^2. \tag{A69}$$

Using Eqs. (A68) and (A69), we get the desired result.

(Proof of part 3) For any random variable Z and two real numbers a, b such that $b \geq a$, the following holds:

$$\text{Prob}\left(a \leq Z \leq b\right) = 1 \implies a \leq \mathbb{E}\left(Z\right) \leq b \tag{A70}$$

Substituting $Z = \epsilon^2$ in Eq. (A70) and using Eq. (A69) for the expectation value of Z, we get the desired result. □

Distinctly from prior findings, our sample-complexity lower bound is defined by the commutativity of the Gibbs state with the terms in the Hamiltonian. Our approach relies on no assumptions about the Hamiltonian's structure. In contrast, earlier studies focused on low-interaction Hamiltonians: each term in the Hamiltonian is supported on a constant number of qubits. For a synopsis, refer to Table I.

Reference	Sample-complexity lower bound	Key technique
Bairey *et al.* [7]	?	NA
Anshu *et al.* [15]	$\Omega\left(\frac{\sqrt{M}+\log(1-\delta)}{\beta\epsilon}\right)$	Quantum state discrimination
Sbahi *et al.* [18]	?	NA
Haah *et al.* [16]	$\Omega\left(\frac{\exp(\beta)M}{\beta^2\epsilon^2}\right)$	Coding theory
Gu *et al.* [17]	?	NA
This work	$\Omega\left(\frac{M}{\beta^2\,\epsilon_{\text{err}}^2}\max\left\{\min_l \frac{1}{(\Delta A_l)^2}\,,\,\min_l \frac{c_2^{-1}/2}{(\Delta A_l)^2-\frac{1}{2}\left\lVert\left[\sqrt{\rho},A_l\right]\right\rVert_2^2}\right\}\right)$	Quantum Cramér-Rao bound

TABLE I. **Complexity of learning Hamiltonians via Gibbs states.** The error ϵ is the l_2-distance error in the estimate of the Hamiltonian parameters. We use the related quantity ϵ_{err}, defined via $\sum_{l=1}^{M}\text{var}(\hat{\mu}_l) = \epsilon_{\text{err}}^2$. Our sample-complexity lower bound, uniquely among the approaches, (i) is based on the commutativity of the Hamiltonian's terms with the Gibbs state and (ii) requires no assumptions about the Hamiltonian's structure. In contrast, previous studies were conducted for low-interaction Hamiltonians (each term in the Hamiltonian is supported on a constant number of qubits). The question marks (?) indicate that no value has been reported or is available. Among the five prior studies, three provide no lower bounds on sample complexity. Therefore, the "key technique" is listed as NA ("not applicable") for these studies.

[1] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, Quantum Fisher information matrix and multiparameter estimation, J. Phys. A: Math. Theor. **53**, 023001 (2019).
[2] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, Phys. Rev. Lett. **96**, 010401 (2006).
[3] S. Boixo, A. Datta, S. T. Flammia, A. Shaji, E. Bagan, and C. M. Caves, Quantum-limited metrology with product states, Phys. Rev. A **77**, 012317 (2008).
[4] H. J. Miller and J. Anders, Energy-temperature uncertainty relation in quantum thermodynamics, Nat. Commun. **9**, 2203 (2018).
[5] K. Matsumoto, A new approach to the Cramér-Rao-type bound of the pure-state model, J. Phys. A Math. Gen. **35**, 3111 (2002).
[6] M. G. Paris, Quantum estimation for quantum technology, Int. J. Quantum Inf. **7**, 125 (2009).
[7] E. Bairey, I. Arad, and N. H. Lindner, Learning a local Hamiltonian from local measurements, Phys. Rev. Lett. **122**, 020504 (2019).
[8] J. R. Garrison and T. Grover, Does a single eigenstate encode the full Hamiltonian?, Phys. Rev. X **8**, 021026 (2018).
[9] X.-L. Qi and D. Ranard, Determining a local Hamiltonian from a single eigenstate, Quantum **3**, 159 (2019).
[10] M. Greiter, V. Schnells, and R. Thomale, Method to identify parent Hamiltonians for trial states, Phys. Rev. B **98**, 081113 (2018).
[11] E. Chertkov and B. K. Clark, Computational inverse method for constructing spaces of quantum models from wave functions, Phys. Rev. X **8**, 031029 (2018).
[12] W. Zhu, Z. Huang, and Y.-C. He, Reconstructing entanglement Hamiltonian via entanglement eigenstates, Phys. Rev. B **99**, 235109 (2019).
[13] X. Turkeshi, T. Mendes-Santos, G. Giudici, and M. Dalmonte, Entanglement-guided search for parent Hamiltonians, Phys. Rev. Lett. **122**, 150606 (2019).
[14] J. Zhou and D. L. Zhou, Recovery of a generic local Hamiltonian from a steady state, Phys. Rev. A **105**, 012615 (2022).
[15] A. Anshu, S. Arunachalam, T. Kuwahara, and M. Soleimanifar, Sample-efficient learning of interacting quantum systems, Nat. Phys. , 1 (2021).
[16] J. Haah, R. Kothari, and E. Tang, Optimal learning of quantum Hamiltonians from high-temperature Gibbs states, in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, 2022) pp. 135–146.
[17] A. Gu, L. Cincio, and P. J. Coles, Practical hamiltonian learning with unitary dynamics and Gibbs states, Nat. Commun. **15**, 312 (2024).
[18] F. M. Sbahi, A. J. Martinez, S. Patel, D. Saberi, J. H. Yoo, G. Roeder, and G. Verdon, Provably efficient variational generative modeling of quantum many-body systems via quantum-probabilistic information geometry, arXiv preprint arXiv:2206.04663 10.48550/arXiv.2206.04663 (2022).
[19] D. Hangleiter, I. Roth, J. Fuksa, J. Eisert, and P. Roushan, Robustly learning the hamiltonian dynamics of a superconducting quantum processor, arXiv (2024), arXiv:2108.08319 [quant-ph].
[20] W. Yu, J. Sun, Z. Han, and X. Yuan, Robust and efficient Hamiltonian learning, Quantum **7**, 1045 (2023).
[21] Efficient and robust estimation of many-qubit hamiltonians, Nat. Commun. **15**, 311 (2024).
[22] H.-Y. Huang, Y. Tong, D. Fang, and Y. Su, Learning many-body Hamiltonians with Heisenberg-limited scaling, Phys. Rev. Lett. **130**, 200403 (2023).
[23] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. O'brien, J. G. Rarity, A. Laing, *et al.*, Experimental quantum Hamiltonian learning, Nat. Phys. **13**, 551 (2017).
[24] C. Senko, J. Smith, P. Richerme, A. Lee, W. Campbell, and C. Monroe, Coherent imaging spectroscopy of a quantum many-body spin system, Science **345**, 430 (2014).