

Robust estimation of the Quantum Fisher Information on a quantum processor

Vittorio Vitale,¹ Aniket Rath,¹ Petar Jurcevic,² Andreas Elben,^{3,4} Cyril Branciard,⁵ and Benoît Vermersch^{1,6,7}

¹Univ. Grenoble Alpes, CNRS, LPMAC, 38000 Grenoble, France

²IBM Quantum, Yorktown Heights, NY 10598 USA

³Institute for Quantum Information and Matter, Caltech, Pasadena, CA, USA

⁴Walter Burke Institute for Theoretical Physics, Caltech, Pasadena, CA, USA

⁵Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France

⁶Institute for Theoretical Physics, University of Innsbruck, Innsbruck, Austria

⁷Institute for Quantum Optics and Quantum Information of the Austrian Academy of Sciences, Innsbruck A-6020, Austria

(Dated: 11th July 2024)

We present the experimental measurement, on a quantum processor, of a series of polynomial lower bounds that *converge* to the quantum Fisher information (QFI), a fundamental quantity for certifying multipartite entanglement that is useful for metrological applications. We combine advanced methods of the randomized measurement toolbox to obtain estimators that are robust against drifting errors caused uniquely during the randomized measurement protocol. We estimate the QFI for Greenberger–Horne–Zeilinger states, observing genuine multipartite entanglement. Then, we prepare the ground state of the transverse field Ising model at the critical point using a variational circuit. We estimate its QFI and investigate the interplay between state optimization and noise induced by increasing the circuit depth.

The quantum Fisher information is defined with respect to a Hermitian operator A and a quantum state ρ , and can be expressed in terms of the spectral decomposition $\rho = \sum_{\mu} \lambda_{\mu} |\mu\rangle\langle\mu|$ of the state under consideration as [1]

$$F_Q = 2 \sum_{(\mu,\nu), \lambda_{\mu} + \lambda_{\nu} > 0} \frac{(\lambda_{\mu} - \lambda_{\nu})^2}{\lambda_{\mu} + \lambda_{\nu}} |\langle\mu| A |\nu\rangle|^2. \quad (1)$$

It plays a crucial role in various quantum phenomena, including quantum phase transitions [2, 3] and quantum Zeno dynamics [4] and exhibits profound connections with multipartite entanglement [5–8]. For N qubits, with a collective spin operator $A = \frac{1}{2} \sum_{j=1}^N \sigma_j^{(\tau)}$ [93], multipartite entanglement can be certified via QFI in terms of k -producibility of the state ρ , i.e. a decomposition into a statistical mixture of tensor products of k -particle states [5, 6]. The QFI also holds vast applications in resource theory [9], many-body physics [10, 11] and quantum metrology [12]. In particular, for quantum parameter estimation problems, the inverse of the QFI limits the estimation accuracy, as given by the quantum Cramér-Rao bound [13]. Therefore, it is fundamental for identifying states that provide sensitivities beyond the standard quantum limit [12].

In recent years, we have seen an important effort to try to measure the QFI in various experiments (which we will detail in Sec. I below). This is of interest to test that a quantum device is able to generate nontrivial multipartite entanglement, but also to benchmark the potential of a quantum state for quantum metrology. In this context, our work provides a measurement of the QFI in a large (up to 13 qubits) quantum processor. Note that, while not specifically designed for performing quantum metrology, quantum processors offer, at the moment, unique capabilities for estimating quantum state properties via fast, high-fidelity projective measure-

ments. The measurement of the QFI is particularly useful for understanding how metrologically relevant quantum states are affected in the presence of unavoidable experimental noise (or what states are more resilient), which in turn can inspire protocols for generating more robust quantum states in actual quantum sensors. Importantly, in contrast to previous related approaches accessing lower bounds with finite distance to the QFI, our work provides converging estimations of the QFI. This is particularly relevant for the type of mixed quantum states that are nowadays accessible with current quantum technology, as noisy quantum channels can unpredictably alter the QFI and the previously measured lower bounds.

Our work is an experimental demonstration of the randomized measurements (RM) protocol presented in Ref. [14], which proposes a systematic and state-agnostic way to estimate the QFI by measuring a converging series of monotonically increasing lower bounds. Several practical limitations of this RM protocol have so far prevented the experimental measurement of the QFI: (i) a prohibitive classical-postprocessing time for reconstructing the QFI from data, (ii) gate and readout errors affecting the RM protocol, and (iii) a prohibitive required number of measurements to overcome statistical errors. The present work takes advantage of three recently developed methods to adapt Ref. [14] with respect to the issues (i), (ii), and (iii), and experimentally realize high-fidelity measurements of the QFI. First, we use the “batch shadows” formalism [15] to decrease the required post-processing time by several orders of magnitude [issue (i)]. Second, we suppress the role of errors occurring in the RM protocol based on the experimental demonstration of robust classical shadows [16–18] [issue (ii)]. Finally, we apply the formalism of common randomized measurements [19] to significantly reduce statistical errors compared to the standard RM approach [issue (iii)]. This is particularly crucial to obtain a converging value of the QFI with ro-

bust classical shadows that typically require more measurements compared to previous approaches [16–18]. We show, in particular, for the largest system size attainable, that the error mitigation of the RM protocol becomes essential, providing estimates that are compatible with theoretical predictions and up to three times larger than non-mitigated results. Note that, beyond the measurement of the QFI, we believe that experimentally demonstrating such practical upgrades of the RM toolbox will be useful to measure more faithfully and efficiently other important physical properties for characterizing quantum processors, such as entanglement entropies [20–26], negativities [27–29] and state overlaps [30–32].

The manuscript starts with a discussion that puts our results within the framework of previous literature about QFI estimations (Sec. I). Our approach is discussed in Sec. II, where we elaborate on the robust randomized measurement protocol that we implement. There, we also discuss the methods introduced above to post-process efficiently the collected experimental RM data [15, 19]. Finally, in Sec. III, we discuss the experimental results where our protocol is applied to measure the converging lower bounds of the QFI for quantum states with up to 13 qubits, prepared on the IBM superconducting device ‘ibm_prague’ [33]. In particular, we consider two prototypical examples of states: (i) the Greenberger–Horne–Zeilinger (GHZ) state and (ii) the ground state of the transverse field Ising model (TFIM) at the critical point [34]. Additionally, we provide more details on our work in the Appendices, organized as follows. In App. A and B we provide the analytical expressions of the lower bounds of the QFI, and analyze their behavior in the presence of global depolarizing errors. In App. C we give more analytical details of our post-processing protocol and an experimental analysis of the noise mitigation parameters we employ. In App. D we investigate the noise in the platform. In App. E we introduce an estimator to verify the locality of the noise and measure it for our experimental setup. In App. F we provide more results for both experiments discussed in the main text. In App. G, we perform a numerical investigation on the statistical error of our estimators to justify the choice of the parameters used in the experiment.

I. PREVIOUS WORKS ON MEASURING QUANTUM FISHER INFORMATION AND OUR CONTRIBUTION

In this section we discuss previous works that estimated the QFI and lower bounds to it. In light of these, we describe how our work goes beyond previous contributions. In what follows, we find it useful to distinguish measurement protocols devised in the context of quantum metrology from the tomographic and randomized measurement methods usually associated with quantum processors. The distinction between these two classes of quantum devices is, however, not 100% sharp,

as one can, for instance, consider performing state tomography in a small quantum metrological device based on using local basis transformations.

A. Quantum metrology methods

The QFI was first introduced in the context of quantum metrology for quantifying how accurate the estimation of an unknown parameter θ could be, and it was readily used to show that the precision of the measurement could go beyond the shot-noise limit (or standard quantum limit) [12, 13]. In quantum metrology, one typically considers realizing the transformation $\rho \mapsto \rho(\theta) = e^{-iA\theta}\rho e^{iA\theta}$ for some Hermitian operator A . The phase shift θ is then determined through projective measurements in a given measurement setting, with measurement outcomes μ coming with probabilities $P(\mu|\theta)$ [12]. The corresponding uncertainty $\Delta\theta$ is bounded by the classical Fisher information [35, 36] (CFI)

$$F(\theta) = \sum_{\mu} \frac{1}{P(\mu|\theta)} \left(\frac{\partial P(\mu|\theta)}{\partial \theta} \right)^2, \quad (2)$$

according to the Cramér-Rao bound [1] $\Delta\theta \geq 1/\sqrt{F(\theta)}$. The QFI, as defined in Eq. (1), is then an upper bound to the CFI and is obtained by maximizing over all possible quantum mechanical measurements [1].

In quantum metrological devices, one can thus estimate a lower bound to the QFI by implementing physically the state $\rho(\theta)$, performing projective measurements to estimate $P(\mu|\theta)$ and accessing the CFI $F(\theta) \leq F_Q$ using the above relation [94]. We emphasize here that only when the optimal measurement setting is chosen, the CFI and QFI coincide. The CFI has been measured in Refs. [37, 38] and has been employed to show genuine multipartite entanglement in GHZ states, for up to 6 qubits [39]. In App. G 1 we provide a numerical comparison of the CFI and QFI for noisy GHZ states. We observe that in the presence of noise, the CFI for some fixed measurement setting provides values that decrease much faster than the QFI as a function of the noise strength, and thus does not represent the optimal metrological content of the prepared state.

Apart from the CFI, various other bounds to the QFI have been proposed and experimentally measured in the quantum metrology context [40–45]. This includes in particular spin-squeezing inequalities that can be directly extracted from measuring expectation values of simple quantum observables. Again, in the presence of unavoidable experimental noise, the distance between a given bound and the QFI is a priori unknown, and it is thus desirable to develop complementary methods to access the QFI directly.

B. Tomographic and randomized measurements methods

Quantum processors such as quantum computers are perhaps not the most natural platform for performing parameter estimation. However, they can be used to experimentally study the generation of quantum states relevant to quantum metrology but also to verify the presence of multipartite entanglement. In this context, one can extract an estimate of the QFI via quantum state tomography (QST), and randomized measurement (RM) methods. Note that these methods are also in principle available in quantum metrological platforms, such as cold atoms [37], but quantum processors, like the superconducting platform employed here, are at the moment more suited for these tasks as they allow for high-fidelity and fast measurement (at the kHz rate).

QST allows us to reconstruct the state ρ from projective measurements and, therefore, to estimate the QFI by evaluating Eq. (1). For an unknown quantum state, described by density matrices of rank χ , the required number of measurements scales as $\mathcal{O}(\chi^2 2^N)$ [46–48] and thus is prohibitively demanding for a large number of qubits. This method has been used to estimate the QFI for small system sizes, up to 4 qubits only [49].

One can go beyond QST to access QFI in a more scalable way. Under the assumption of thermal states, one can measure the QFI using dynamical susceptibilities [2]. For generic quantum states, one can also rewrite or approximate the QFI in a form that is more suitable for measurements in a quantum processor, i.e., that does not require QST. In recent years, many works proposed nonlinear quantities as a function of the density matrix that lower bounds the QFI [50–55]. A notable experimental implementation of this approach was presented in Ref. [56] where the authors provide an estimate of a lower bound to the QFI, named sub-QFI [55], for GHZ states up to 4 qubits using randomized measurements. As we will explain below, RM protocols require projective measurements in multiple measurements basis, as in QST, but the data is processed to access directly functions of ρ , in that case the sub-QFI. In Ref. [56], RMs were performed on two states $\rho(\theta)$ and $\rho(\theta + d\theta)$, giving access to $G(\rho(\theta), \rho(\theta + d\theta))/d\theta^2$ where G is a generalized overlap between quantum states [55]. The protocol suffers from two important limitations: (i) it only estimates the sub-QFI in the limit of small values of $d\theta$ ($d\theta \rightarrow 0$), which therefore tends to amplify statistical errors of RMs, (ii) the distance between the QFI and the sub-QFI can be significant in the presence of noise.

The protocol presented in Ref. [14], which we experimentally implement here, addressed these two challenges. To be specific, we use RMs to access a converging series of lower bounds of the QFI, i.e., not only one bound to the QFI. In addition, our estimators do not rely on measuring asymptotically a relative overlaps between two states $\rho(\theta)$ and $\rho(\theta + d\theta)$. This allows us to reach a number of qubits of 13, i.e., more than three times the system size

that was achieved so far with QST [49]. In App. F 1, we compare our experimental estimations of the QFI with the estimator in Ref. [56], which we show we can extract faithfully from a single RM experiment.

II. ESTIMATION OF MULTI-COPY OBSERVABLES FROM NOISY RANDOMIZED MEASUREMENT DATA

In this section we describe all the building blocks for the estimation of multi-copy observables employing RM data. In particular, we present robust estimators of the QFI, the purity and the fidelity of a quantum state, and detail the needed post-processing protocol to mitigate the noise and reduce statistical errors.

A. Data acquisition with randomized measurements

Our approach, illustrated in Fig. 1, is comprised of several repetitions of two building blocks: (i) Calibration of randomized measurements and (ii) randomized measurements on the state of interest ρ . The calibration step is employed to learn and mitigate the gate and readout errors that affect the measurements, as described in Refs. [16–18]. This relies on the ability to prepare on the experimental platform a specific state with high fidelity. In this work, we fix the calibration state to be $|0\rangle \equiv |0\rangle^{\otimes N}$, which is producible with high efficiency on our quantum processor. The data collected in step (ii) are then used for estimating the observables we are interested in. We call each run of (i) and (ii) an ‘iteration’ of the experiment. Performing consecutive iterations allows us to account for the temporal variations in gate and readout errors. Assuming that the temporal fluctuations of the errors affecting the randomized measurements protocol for each iteration are negligible, each calibration step captures the specific error profile of a distinct time window within the overall experimental run.

Let us start by recalling the randomized measurement protocol in the absence of noise. We begin by preparing the N -qubit quantum state ρ . Then we apply local random unitaries $U = U_1 \otimes \dots \otimes U_N$ where the local (single-qubit) unitaries U_j ($j = 1, \dots, N$) are sampled from the circular unitary ensemble [57]. The rotated state $U\rho U^\dagger$ is then projected onto a computational basis state $|s\rangle = |s_1, \dots, s_N\rangle$, where $s_j \in \{0, 1\}$ for $j = 1, \dots, N$, by performing a measurement. To make the protocol robust against the noise occurring in the quantum device, we apply the measurement sequence described above on the states $|0\rangle$, ρ in steps (i), (ii) of Fig. 1, respectively. As mentioned before, the data collected from (i) is used to mitigate the errors induced by the noisy measurement protocol in step (ii) [16, 17].

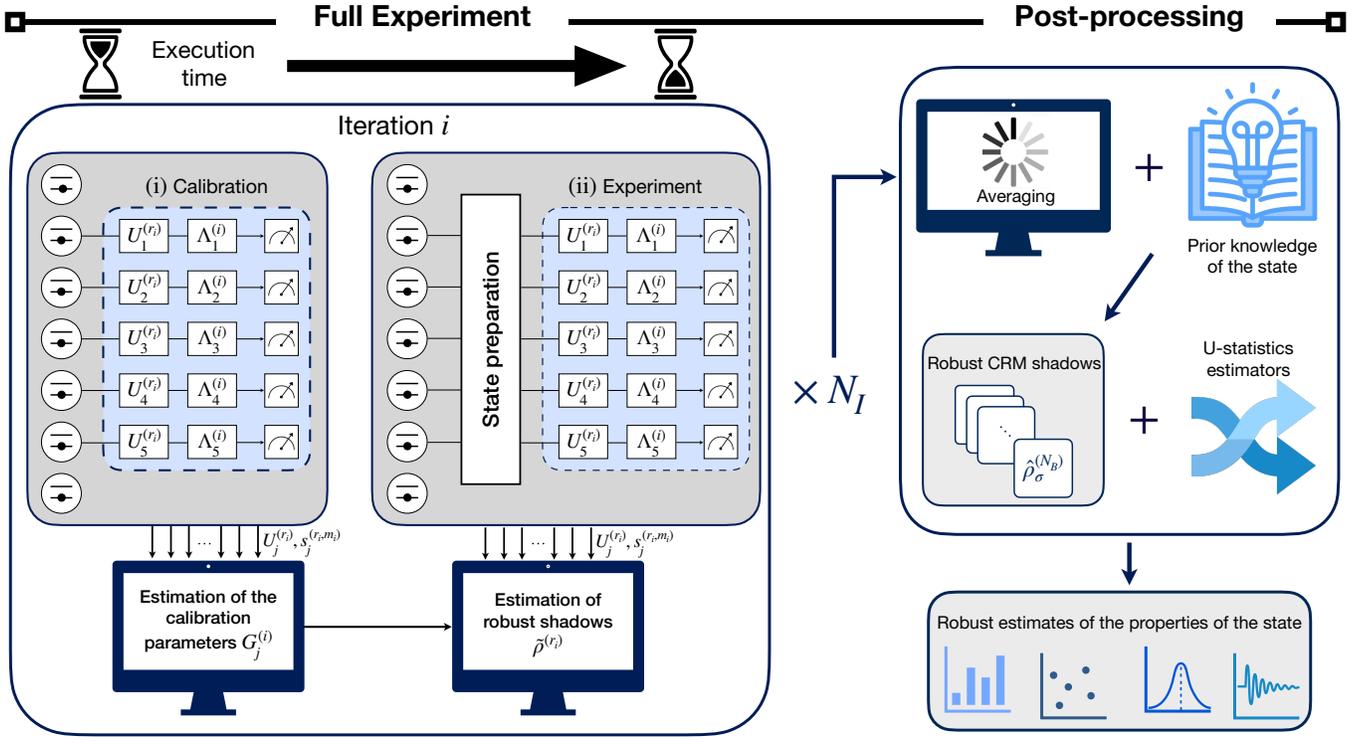


Figure 1: Overview of experimental protocol and post-processing. The experimental protocol is comprised of several ‘iterations’ $i = 1, \dots, N_I$. Each of them consists of a calibration step (i) and the experiment on the state of interest ρ (ii). From the unitaries and bit-strings recorded in (i) we estimate the noise parameter $G_j^{(i)}$ (Eq. (4)), that is used for the construction of the robust shadows $\tilde{\rho}^{(r_i)}$ (Eq. (3)) together with the data from step (ii). Integrating an approximation σ of the state ρ , all robust shadows are then collected and averaged in batches to obtain $b = 1, \dots, N_B$ robust common randomized batch shadows $\hat{\rho}_\sigma^{(b)}$ (Eq. (12)). These are used for computing the multi-copy observables of interest through the U-statistics estimator (Eq. (15)). Our experiments are performed on the IBM superconducting qubit device ‘ibm_prague’ [33].

B. Assumptions on the noise affecting the randomized measurements

The basic assumptions on the noise model for our post-processing protocol are as follows. As in Ref. [17], we consider a gate-independent noise channel Λ , applied after the random unitaries: that is, for each chosen U the state ρ transforms as $\Lambda(U\rho U^\dagger)$. We assume that the noise channel Λ is constant during each iteration i – we label it as $\Lambda^{(i)}$ – and may change between each iteration. We provide experimental evidence of the variation of the noise over different iterations – that is remarkably captured by our protocol – in App. D1. Additionally, we assume the noise to be local for each qubit, so that $\Lambda^{(i)} = \Lambda_1^{(i)} \otimes \dots \otimes \Lambda_N^{(i)}$. In App. E we provide and implement a method to verify the assumption of locality of the noise, based on the calibration data. Additionally, in App. F3, we show that tracking the variation of the noise over the different iterations is essential to provide faithful estimations of the QFI.

C. Robust classical shadows

The first step towards the measurement of non-local observables of interest is to construct estimators of the density matrix ρ from the noisy measurements. This object, called a ‘robust shadow’ [17] (see also App. C1), can be defined as

$$\begin{aligned} \tilde{\rho}^{(r_i)} &= \sum_{\mathbf{s}} \hat{P}(\mathbf{s}|U^{(r_i)}) \bigotimes_{j=1}^N \left(\alpha_j^{(i)} U_j^{(r_i)\dagger} |s_j\rangle\langle s_j| U_j^{(r_i)} - \beta_j^{(i)} \mathbb{1} \right), \end{aligned} \quad (3)$$

where $\alpha_j^{(i)} = \frac{3}{2G_j^{(i)} - 1}$ and $\beta_j^{(i)} = \frac{G_j^{(i)} - 2}{2G_j^{(i)} - 1}$. Here r_i labels a unitary in iteration i and $\hat{P}(\mathbf{s}|U^{(r_i)}) = \sum_{m_i=1}^{N_M} \frac{\delta_{\mathbf{s}, \mathbf{s}^{(r_i, m_i)}}}{N_M}$ is the estimated (noisy) Born probability, where m_i labels an individual measurement performed after the application of $U_j^{(r_i)}$, whose outcome is the bit-string denoted $\mathbf{s}^{(r_i, m_i)}$. The quantity in Eq. (3) satisfies $\mathbb{E}[\tilde{\rho}^{(r_i)}] = \rho$, where the average is taken over the applied unitaries. This equality is necessary in particular to derive the un-

biased estimators of the lower bounds of the QFI [14].

The quantity $G_j^{(i)}$ introduced above contains the relevant information about the noise on qubit j in iteration i of the measurement protocol. It is defined as

$$G_j^{(i)} = \frac{1}{2} \sum_{s_j=0,1} \langle s_j | \Lambda_j^{(i)} (|s_j\rangle\langle s_j|) |s_j\rangle \quad (4)$$

and can be interpreted as the average ‘survival probability’ of the two basis states of qubit j . In the absence of noise, $G_j^{(i)} = 1$ ($\forall j = 1, \dots, N$), and one recovers the standard ‘classical shadow’ [58]. In the opposite limit of fully depolarising noise, $G_j^{(i)} = 1/2$, the coefficient $\alpha_j^{(i)}, \beta_j^{(i)}$ diverge and the estimators suffer from large statistical errors [17]. In our work $G_j^{(i)} \sim 0.98$ (See App. C3).

For each iteration i and each qubit j , $G_j^{(i)}$ is computed through the experimental data collected during the calibration step. As detailed in App. C2, we can define the unbiased estimator

$$\hat{G}_j^{(i)} = \frac{3}{N_U} \sum_{r_i, s_j} \hat{P}(s_j | U_j^{(r_i)}) P(s_j | U_j^{(r_i)}) - 1. \quad (5)$$

Here, $\hat{P}(s_j | U_j^{(r_i)}) = \sum_{m_i=1}^{N_M} \frac{\delta_{s_j, s_j^{(r_i, m_i)}}}{N_M}$ is the estimated (noisy) Born probability and $P(s_j | U_j^{(r_i)}) = |\langle s_j | U_j^{(r_i)} | 0 \rangle|^2$ is the theoretical (noiseless) one for a single qubit j . The information on the noise is contained in $\hat{P}(s_j | U_j^{(r_i)})$, that approaches the theoretical noisy Born probability $P_\Lambda(s_j | U_j^{(r_i)}) = \langle \mathbf{s} | \Lambda(U^{(r_i)} \rho U^{(r_i)\dagger}) | \mathbf{s} \rangle$ in the limit $N_M \rightarrow \infty$. We remark here that all our results are compatible with the ones in Ref. [17], where a slightly different formalism has been employed.

D. Common randomized shadows and noise estimator

The statistical errors can be significantly reduced by using *common randomized measurements* [19] to define an estimator with smaller variance. The central idea is to construct the robust shadows integrating an approximation of the state ρ in the form of some classical representation σ . In practice, we consider σ to be the ideal pure state that we intend to realise in our experiment. We build ‘common randomized’ (CRM) shadows as

$$\tilde{\rho}_\sigma^{(r_i)} = \tilde{\rho}^{(r_i)} - \sigma^{(r_i)} + \sigma, \quad (6)$$

where the term $\sigma^{(r_i)}$ is constructed from σ as

$$\sigma^{(r)} = \sum_{\mathbf{s}} P_\sigma(\mathbf{s} | U^{(r)}) \bigotimes_{j=1}^N \left(3 U_j^{(r)\dagger} |s_j\rangle\langle s_j| U_j^{(r)} - \mathbb{1} \right) \quad (7)$$

with $P_\sigma(\mathbf{s} | U^{(r)}) = \langle \mathbf{s} | U^{(r)} \sigma U^{(r)\dagger} | \mathbf{s} \rangle$. The latter is a fictitious probability distribution obtained from computational basis measurements on the σ rotated by the same

unitaries applied in the experiment $U^{(r_i)}$, done on a classical device. One may notice that $\mathbb{E}[\sigma^{(r_i)}] = \sigma$ [19]. Thus, $\hat{\rho}_\sigma^{(r_i)}$ is an unbiased estimator of ρ , as

$$\mathbb{E}[\hat{\rho}_\sigma^{(r_i)}] = \rho - \sigma + \sigma = \rho, \quad (8)$$

irrespective of the choice of σ . Crucially, this procedure enters entirely during post-processing thus leaving the data acquisition of the experiment independent from it.

The same reasoning of common randomized numbers [19] can be used to improve the statistical accuracy of the estimator of $G_j^{(i)}$. Let us introduce for that the quantity

$$B_j^{(i)} = \frac{3}{N_U} \sum_{r_i, s_j} P(s_j | U_j^{(r_i)})^2. \quad (9)$$

With this, we then define

$$\hat{\mathcal{G}}_j^{(i)} = \hat{G}_j^{(i)} - B_j^{(i)} + \mathbb{E}[B_j^{(i)}]. \quad (10)$$

Here, $\hat{\mathcal{G}}_j^{(i)}$ and $\hat{G}_j^{(i)}$ have the same expectation value, but the variance of $\hat{\mathcal{G}}_j^{(i)}$ is smaller because $\hat{G}_j^{(i)}$ and $B_j^{(i)}$ are positively correlated. Observing in particular that $\mathbb{E}[B_j^{(i)}] = 2$, see App. C2, we can then write the new estimator as

$$\hat{\mathcal{G}}_j^{(i)} = \frac{3}{N_U} \sum_{r_i, s_j} \widehat{\Delta P}(s_j | U_j^{(r_i)}) P(s_j | U_j^{(r_i)}) + 1, \quad (11)$$

where $\widehat{\Delta P}(s_j | U_j^{(r_i)}) = \hat{P}(s_j | U_j^{(r_i)}) - P(s_j | U_j^{(r_i)})$ is the difference between the experimentally estimated (noisy) Born probability and the theoretical (noiseless) one. The fact that $\hat{\mathcal{G}}_j^{(i)}$ is a more efficient estimator of the noise term $G_j^{(i)}$ is shown explicitly in App. C3, employing our experimental data.

E. Data compression and U-statistics estimators

The last step of our protocol consists in compressing the CRM shadows in order to minimize the post-processing time. To do so, we compress the N_I estimators $\tilde{\rho}^{(r_i)}$ into N_B ‘robust CRM batch shadows’ [15] $\hat{\rho}_\sigma^{(b)}$ as (we assume N_B divides N_I):

$$\hat{\rho}_\sigma^{(b)} = \frac{N_B}{N_I} \sum_{i=(b-1)N_I/N_B+1}^{bN_I/N_B} \left(\sum_{r_i} \frac{\tilde{\rho}_\sigma^{(r_i)}}{N_U} \right) \quad (12)$$

for $b = 1, \dots, N_B$. The latter can be used to estimate any multi-copy observable of interest, i.e. functions f_n in the form $f_n = \text{tr}(O^{(n)} \rho^{\otimes n})$. Given the N_B robust CRM batch shadows $\hat{\rho}_\sigma^{(b)}$, one can provide an unbiased estimator of the function f_n using U-statistics [59]. This is achieved by replacing each copy of the density matrix in the multi-copy function f_n with a different robust CRM

batch shadow and computing the average over all possible such choices. This is explicitly expressed as

$$\hat{f}_n = \frac{1}{n! \binom{N_B}{n}} \sum_{b_1 \neq \dots \neq b_n} \text{Tr} \left(O^{(n)} \hat{\rho}_\sigma^{(b_1)} \otimes \dots \otimes \hat{\rho}_\sigma^{(b_n)} \right), \quad (13)$$

where $\rho^{\otimes n}$ from $f_n = \text{Tr}(O^{(n)} \rho^{\otimes n})$ is replaced by an average over $\hat{\rho}_\sigma^{(b_1)} \otimes \dots \otimes \hat{\rho}_\sigma^{(b_n)}$ with $b_1 \neq \dots \neq b_n$. Such estimators can be evaluated with a classical post-processing which scales with the number of elements to be evaluated in the sum. Then, it is clear that the compression of the data from N_I to N_B shadows allows for faster post-processing time, that changes from $\mathcal{O}(N_I^n)$ to $\mathcal{O}(N_B^n)$. This comes at the expense of storing large dense matrices, albeit not compromising the statistical performances [15]. Finally, as thoroughly explained in Ref. [15], one has to consider that while the statistical accuracy increases with N_B , so does also the post-processing time. Then one has to find an N_B that provides a good balance between good statistical performances and reasonable post-processing cost.

F. Estimators of the QFI as a converging series of polynomials

Let us now discuss in detail the estimators of the QFI F_Q that we use. As shown in Ref. [14], while the QFI cannot be accessed directly by randomized measurements, as written in Eq. (1), it can be alternatively expressed and estimated in terms of a converging series of monotonically increasing lower bounds F_n . For the first three orders $n = 0, 1, 2$, one can write explicitly

$$\begin{aligned} F_0 &= 4 \text{Tr}(\rho[\rho, A]A), \\ F_1 &= 2F_0 - 4 \text{Tr}(\rho^2[\rho, A]A), \\ F_2 &= 3(F_1 - F_0) + 4 \text{Tr}(\rho^2[\rho^2, A]A), \end{aligned} \quad (14)$$

where $[\cdot, \cdot]$ is the commutator. We provide the general expression for F_n in App. A. Each function F_n is a polynomial function of the density matrix (of order $n+2$); such functions can now be accessed via randomized measurements, as it has been shown for entropies [20–25], negativities [27–29], state overlaps [30–32], scrambling [60] and topological invariants [61, 62].

We define unbiased estimators \hat{F}_n for the lower bounds F_n according to the rules of U-statistics [58, 59] by summing over all possible disjoint indices, as in Eq. (13) above. In practice [14], for $n = 0, 1, 2$ one can thus

write (assuming $N_B > n+2$)

$$\begin{aligned} \hat{F}_0 &= \frac{4(N_B-2)!}{N_B!} \sum_{b_1 \neq b_2} \text{Tr} \left(\hat{\rho}_\sigma^{(b_1)} [\hat{\rho}_\sigma^{(b_2)}, A]A \right), \\ \hat{F}_1 &= 2\hat{F}_0 - \frac{4(N_B-3)!}{N_B!} \sum_{b_1 \neq \dots \neq b_3} \text{Tr} \left(\hat{\rho}_\sigma^{(b_1)} \hat{\rho}_\sigma^{(b_2)} [\hat{\rho}_\sigma^{(b_3)}, A]A \right), \\ \hat{F}_2 &= 3(\hat{F}_0 - \hat{F}_1) + \frac{4(N_B-4)!}{N_B!} \sum_{b_1 \neq \dots \neq b_4} \text{Tr} \left(\hat{\rho}_\sigma^{(b_1)} \hat{\rho}_\sigma^{(b_2)} [\hat{\rho}_\sigma^{(b_3)} \hat{\rho}_\sigma^{(b_4)}, A]A \right). \end{aligned} \quad (15)$$

where $[\cdot, \cdot]$ denotes the commutator and we choose $N_B = 10$. The estimators \hat{F}_n suffer from statistical errors arising due to the finite number of unitaries and measurements performed. Even though \hat{F}_n exponentially converges to the true value of QFI as a function of the order n of the bound, the statistical error on the estimator increases with n for a fixed measurement budget [14]. In App. G3 we show the scaling of the required number of measurements for a given value of statistical error \mathcal{E} as a function of the system size N . Accurate variance bounds for \hat{F}_n have been discussed in Ref. [14].

G. Estimators of fidelity and purity

As anticipated, the protocol presented in this manuscript is not restricted to the estimation of the lower bounds F_n . For example, we can extract two other important quantities: fidelity with respect to an ideal state, and purity. We write here the estimators in terms of the CRM batch shadows $\hat{\rho}_\sigma^{(b)}$. Following Eq. (13) once again, the estimator of the purity can be expressed as

$$\widehat{\text{Tr}(\rho^2)} = \frac{1}{N_B(N_B-1)} \sum_{b_1 \neq b_2} \text{Tr} \left(\hat{\rho}_\sigma^{(b_1)} \hat{\rho}_\sigma^{(b_2)} \right). \quad (16)$$

On the other hand, assuming that an ideal state is described by the density matrix σ' , one can estimate the overlap of the latter with the prepared state ρ as

$$\widehat{\text{Tr}(\rho\sigma')} = \frac{1}{N_B} \sum_b \text{Tr} \left(\hat{\rho}_\sigma^{(b)} \sigma' \right). \quad (17)$$

In the following, since we are interested in the quality of the state preparation on the hardware, we will testify it by measuring these two quantities.

III. EXPERIMENTAL RESULTS

In this section we describe the experimental results that were obtained on IBM superconducting processors. As mentioned before, we will consider two states: the GHZ state in Sec. III B and the ground state of the transverse field Ising model (TFIM) at the critical point in Sec. III C. In our work, the observable under consideration is taken to be $A = \frac{1}{2} \sum_j \sigma_j^z$, where σ_j^z is the Pauli- z operator acting on qubit j .

A. Measurement budget

The full experiment is divided into a total of N_I iterations (labeled by $i = 1, \dots, N_I$). For steps (i) and (ii) in each iteration i , we apply the same $N_U = 200$ local random unitaries $U^{(r_i)} = U_1^{(r_i)} \otimes \dots \otimes U_N^{(r_i)}$, with $r_i = 1, \dots, N_U$ [95], and (for each unitary) record $N_M = 1000$ measurement outcome bit-strings $\mathbf{s}^{(r_i, m_i)} = (s_1^{(r_i, m_i)}, \dots, s_N^{(r_i, m_i)})$ with $m_i = 1, \dots, N_M$.

The total measurement budget ($N_I N_U N_M$) that is required to reach a given accuracy for an estimator depends on the size of the system N [63]. In particular, for our experiments, we employ a total of $N_U^{\text{tot}} = N_I N_U = 300 \cdot 2^{0.5N}$ unitaries to obtain an estimation error of $\sim 10\%$ on the highest-order estimated lower bound of the QFI (without exploiting any prior knowledge of the prepared quantum state). Note that the higher the order, the more measurements are needed to overcome statistical fluctuations. Numerical investigations on the measurement budget are detailed in App. G3.

B. GHZ states

The Greenberger-Horne-Zeilinger (GHZ) state is a fundamental resource for various quantum information processing tasks, including quantum teleportation [64, 65], quantum error correction [66, 67], and quantum cryptography [68]. It can be written as

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle^{\otimes N} + |1\rangle^{\otimes N} \right). \quad (18)$$

Remarkably, GHZ states are ideal candidates for quantum metrology as they saturate the value of the QFI ($F_Q = N^2$) and, thus, can be used to reach higher sensitivities in parameter estimation that scale as $\sim N^{-1}$ (known as the *Heisenberg limit*), and is beyond the standard shot-noise limit $\sim N^{-1/2}$ [5, 69, 70].

By implementing randomized measurements, we experimentally estimate the QFI as a function of different system sizes N and witness the presence of multipartite entanglement [5, 6, 8]. An N -qubit pure GHZ state is genuinely multipartite entangled (GME), i.e., it cannot be decomposed into a statistical mixture of tensor products of $(N-1)$ -particle states. In general, one can use the inequality $F_Q > \Gamma(N, k)$, with $\Gamma(N, k) = \lfloor \frac{N}{k} \rfloor k^2 + (N - \lfloor \frac{N}{k} \rfloor k)^2$, to certify that a state is not k -producible, i.e., that it has an ‘entanglement depth’ of at least $k+1$ [5, 6]. In this case, it is said to be $(k+1)$ -partite entangled. The inequality is particularly relevant in the presence of noise, where a perfect pure state is not produced. Until now, fidelity measurements have allowed validating GME in GHZ states prepared on superconducting qubits [71, 72], 14 trapped ions [40], 18 photonic qubits [73] and other multipartite entangled

states [74–77]. Additionally, GME states can also be verified via multiple coherences for GHZ states [78, 79].

We show our experimental results in Fig. 2. As we mentioned earlier, with the RM framework, we can access many interesting quantum properties from the same experimental dataset. First, to test the quality of the state preparation on our quantum hardware, we extract two important quantities, namely the purity of the final state ($\text{Tr}(\rho^2)$) and the fidelity ($\text{Tr}(\rho\sigma)$) with respect to a pure GHZ state $\sigma = |\text{GHZ}\rangle\langle\text{GHZ}|$. We plot these results in Fig. 2(a) and in Fig. 2(b) respectively. For each panel, the blue points denote the experimental results in the case error mitigation is performed, while the orange ones correspond to the case when it is not, i.e. we take $G_j^{(i)} = 1$, $\alpha_j^{(i)} = 3$ and $\beta_j^{(i)} = -1$ in Eq. (3). We observe clearly that the robust protocol mitigates the errors occurring during the measurement phase as indicated by higher values of fidelities (between 0.85–0.97) and purities (between 0.8–0.95) compared to the unmitigated results for the whole qubit range. In both cases, we observe the decreasing trend of the fidelity and the purity as a function of the system-size N . This signature clearly indicates that noise is induced during state preparation of the GHZ states as the two-qubit gate count increases with N .

Let us analyze now the convergence of the lower bounds to the QFI, in Fig. 2(c). Here we plot F_n ($n = 0, 1, 2, 3$) for the GHZ state prepared on a system of $N = 8$ (green) and $N = 10$ (violet) qubits. In the absence of noise, the theoretical value of the QFI for a GHZ state is $F_Q = N^2$. This is plotted as a thick black line for both $N = 8$ and $N = 10$. From the experimental results, it is clear that the lower bounds do not converge to N^2 , but to a lower value. We can understand it considering that the state preparation is affected by noise, as it is suggested by the measurements of purity and overlap. One simple way for modeling this is by assuming that the system is affected by global depolarization, turning the pure GHZ state into $\rho = (1-p_D) |\text{GHZ}\rangle\langle\text{GHZ}| + p_D \mathbb{1}/2^N$. The noise parameter p_D can be extracted from the experimental values of the purity according to the following relation

$$\text{Tr}(\rho^2) = (1-p_D)^2 + \frac{2p_D - p_D^2}{2^N}. \quad (19)$$

With this specific noise model, one can better understand the convergence of the lower bounds F_n to a finite value of QFI as a function of the bound order n . While in the noiseless scenario one has $F_Q = N^2$, the theoretical value of the QFI of the noisy GHZ state is found to be [6]

$$F_Q(p_D) = N^2 \frac{(1-p_D)^2}{1-p_D + 2p_D/2^N} \quad (20)$$

For $N = 8$ and $N = 10$ qubits, they correspond to the values of $F_Q(p_D) = 60.3 \pm 0.45$ and $F_Q(p_D) = 92.7 \pm 1.83$. The respective values of p_D are 0.056 ± 0.007 and 0.072 ± 0.018 , which are extracted from the mitigated values of

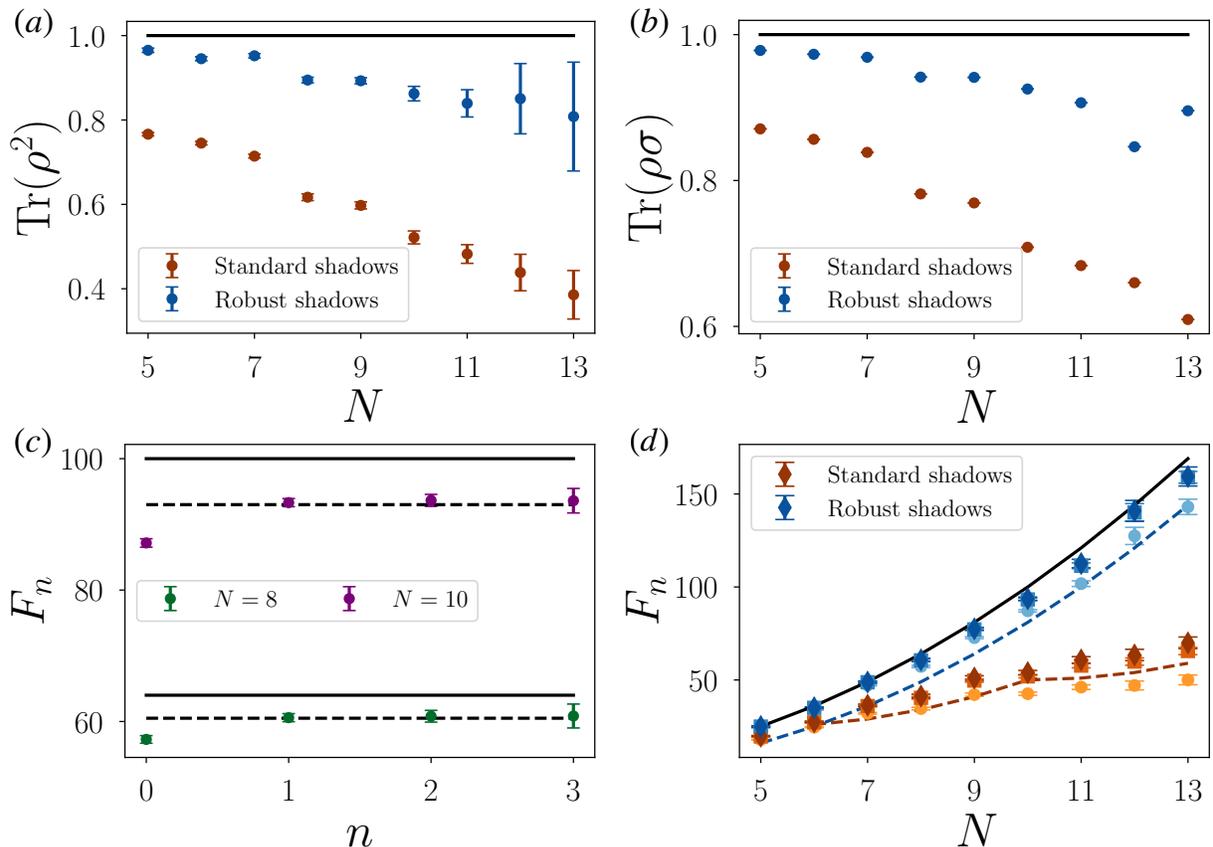


Figure 2: Experimental results for the pure GHZ state. (a) Purity of the prepared state. (b) Fidelity with respect to the GHZ state $\sigma = |\text{GHZ}\rangle\langle\text{GHZ}|$. (c) Convergence of F_n as a function of n for $N = 8, 10$ qubits (green and violet respectively) to the true value of QFI. Solid lines represent the theoretical value for pure GHZ states, dashed lines denote the case of noise affecting the system in the form of global depolarization. (d) F_0, F_1, F_2 (light to dark with circle, square, and diamond respectively) as a function of the number of qubits N . The solid line is the exact value of the QFI $F_Q = N^2$ for pure GHZ states. The dashed blue line corresponds to the entanglement witness $\Gamma(N, k = N - 1) = (N - 1)^2$ above which the state is genuinely multipartite entangled. The dashed orange line corresponds to the entanglement witness $\Gamma(N, k = 5)$ above which we detect a state to be at least 6-partite entangled. The details of the experimental protocol are described in Sec. III A.

purities, i.e. 0.89 ± 0.0066 and 0.86 ± 0.017 (See App. B). In Fig. 2(c) we draw F_Q and $F_Q(p_D)$ as solid and dashed black lines, respectively. We observe the convergence of F_n to the values of the QFI for the extracted value of p_D within error bars of the estimations, for both values of N . As mentioned in Sec. II E, we observe that, even though F_n exponentially converges to the QFI as a function of n , its statistical error increases at a fixed measurement budget. This is thoroughly discussed in Ref. [15].

In Fig. 2(d) we show the experimental measurements of F_0, F_1 , and F_2 (light to dark) on the prepared GHZ state as a function of N . The black thick line provides the ideal scaling of the QFI ($F_Q = N^2$) for pure GHZ states. The black dashed line, instead, denotes the entanglement witness $\Gamma(N, k = N - 1)$ that scales as $(N - 1)^2$ and above which we can consider our prepared states to be GME. The experimental points correspond to the measured values for two different cases: mitigated results through our calibration step in blue, and raw data without performing the calibration step in orange. We observe that the

mitigated data used to estimate F_n violates the necessary entanglement witness to be GME for any size N , hence all our prepared states have an entanglement depth of $k = N$. Thus, we demonstrate the presence of multipartite entanglement through the estimation of *converging* lower bounds to the QFI, whose convergence to the true value has been shown in Fig. 2(a).

Analyzing the raw data (orange points in Fig. 2(b)) that are prone to errors during the randomized measurement protocol gives us lower estimations of the bounds. They do not violate the GME threshold and do not follow the expected scaling seen for the mitigated data points. This shows that the error mitigation in the measurement protocol is decisive and useful for estimating the underlying properties of the prepared quantum states. In the case of the analysis of the raw data, we can assert from the witness bound $F_Q > \Gamma(N, k)$ [5, 6] that our prepared state contains an entanglement depth of $k = 6$ for $N \geq 6$. Importantly, in App. F 3, we show the estimation of the lower bounds in the case when the calibration (step (i))

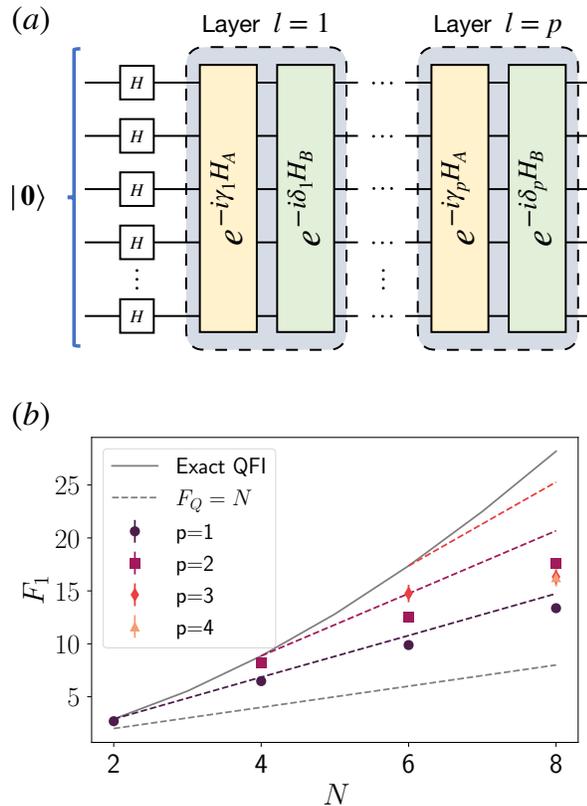


Figure 3: Experimental results for the lower bounds of the QFI for the ground state of the Ising model at the critical point. (a) Sketch of the circuit used to variationally prepare the ground state. (b) Results for F_1 estimated with the robust estimator as a function of the number of qubits N , for different circuit depth p . The solid grey line corresponds to the exact QFI value. The colored dashed lines correspond to the theoretical value for the given depth of the circuit p . The dashed grey line denotes the threshold $F_Q = N$, above which the state is entangled. The measurement budget employed is described in Sec. III A.

in Fig. 1) is done entirely at the beginning and is followed by the experiment (step (ii) in Fig. 1). We observe clearly that performing the calibration in multiple iterations provides better results (closer to the theoretical values) for larger system sizes where the full experimental duration starts to increase.

C. Ground state of the critical TFIM

While the GHZ has a simple analytical wavefunction, we find it instructive to apply our protocol now on a quantum state with a more complex multipartite entanglement structure. We study here the behavior of the QFI at a critical point that presents a rich structure of multipartite entanglement [3, 4, 80, 81]. In particular, we consider the transverse field Ising model (TFIM)

Hamiltonian

$$H = -J \sum_j \sigma_j^z \sigma_{j+1}^z - h \sum_j \sigma_j^x, \quad (21)$$

where h is the transverse field and we set $J = 1$. It displays a quantum phase transition at $h = 1$ that manifests as a growth of multipartite entanglement that can be witnessed by the QFI [2, 81]. We employ classical numerical simulations for estimating variationally the ground state at the critical point, optimizing the parameters of a circuit as it is done for the quantum adiabatic optimization algorithm [34] (See Fig. 3(a)). Then we study the interplay between the depth p of the circuit realized and the approximation of the ground state in an actual experiment. Indeed, in recent times there has been significant interest in measuring the QFI in states prepared through variational circuits on current quantum processors [54, 82, 83].

The preparation of the state entails a series of unitary quantum evolutions under the non-commuting terms in Eq. (21), i.e. $H_A = -J \sum_j \sigma_j^z \sigma_{j+1}^z$ and $H_B = -h \sum_j \sigma_j^x$, that are applied to an initial quantum state $|\psi_0\rangle$ (Fig. 3(a)). The final state after p layers can be written as:

$$|\psi(\boldsymbol{\delta}, \boldsymbol{\gamma})\rangle = \prod_{l=1}^p e^{-i\delta_l H_B} e^{-i\gamma_l H_A} |\psi_0\rangle \quad (22)$$

where the ‘angles’ δ_l and γ_l are variational parameters used in the l -th layer to minimise the final energy $\langle \psi(\boldsymbol{\delta}, \boldsymbol{\gamma}) | H | \psi(\boldsymbol{\delta}, \boldsymbol{\gamma}) \rangle$. The optimal parameters are found by employing a suitable optimization algorithm. In the particular case of our target state, it has been shown that it could be prepared exactly in $p = N/2$ steps, where N is the total number of qubits [84].

In Fig. 3(b) we plot F_1 for different values of the depth p of the circuits as a function of the number of qubits N for the robust estimator. The solid grey line represents the exact numerical value of the QFI. Instead, the colored dashed lines correspond to the exact value of QFI for different depths $p = 1, 2, 3$ from bottom to top. Our first observation consists in remarking that we generate and certify the presence of entanglement in all our prepared states as $F_Q \geq F_1 \geq N$ [12] within error bars for all values of depths p and system size N . The corresponding threshold is shown as a dashed grey line in the plot.

We observe that a larger circuit depth is not tightly linked to a higher measured value of F_1 . Indeed, the increase in the circuit depth p , incorporates more noisy gates that reduce the fidelity of the prepared state compared to its true target state. This results in a decrease of the QFI estimation compromised by the noisy state preparation that is captured very well in Fig. 3(b). In the ideal scenario, increasing the number of layers should guarantee better convergence to the target state. The effect of noise is clearly shown in App. F, where we also provide the estimation of the fidelity of the state preparation and the purity of the prepared state.

Importantly, we establish the presence of multipartite entanglement via F_1 as we violate the entanglement witness $F_1 > \Gamma(N, k = 2)$ [5, 6]. This confirms the presence of an entanglement depth of $k = 3$ for all prepared states of system-size $N > 2$, as the experimental points are above the witness depicted by the dashed dark grey line in Fig. 3(b). Thus, our method allows us to quantify the true metrological power in the form of generating multipartite entanglement in our noisy prepared states. Additionally, we remark here that we have focused on the bound F_1 because it provides more reliable estimates with respect to the other estimators (F_0 and F_2) even if their qualitative behavior is the same. The reason behind this choice is that $F_2 > F_1 > F_0$, but the higher the order, the larger the statistical error at fixed measurement budget, as explained in Sec. II E. We provide the other experimental results in App. F 2.

IV. CONCLUSIONS

In this paper, we have provided an experimental estimation of the quantum Fisher information (QFI). This was achieved on a quantum processor with up to 13 qubits based on the measurements of a *converging* series of polynomial lower bounds. By combining advanced methods from the randomized measurement toolbox, we have been able to overcome drifting gate and readout errors and obtain a robust and unbiased estimator for the QFI.

We applied our method to two different states: GHZ states and the ground state of the TFIM at the critical point. For the former, our measurements are in perfect agreement with theoretical predictions and allow us to witness the presence of multipartite entanglement. With the error mitigation procedure that we introduce here, we observed that all our prepared GHZ states were GME. In the variational preparation of the ground state of the critical TFIM, we utilize the estimated QFI to observe an interesting trade-off. While the theoretical approximation accuracy of the ground state increases with the circuit depth and is optimal at depth $p = N/2$, the best estimation of the theoretically predicted ground state QFI is obtained with a smaller circuit depth. We attribute this effect to noise and decoherence increasing with circuit depth as well.

We have gone beyond previous work aiming at estimating a converging series of lower bounds on the QFI (rather than a single lower bound to it), employing larger system sizes, and obtaining a better convergence to

the true value of the QFI for the prepared state. This has been possible, by exploiting several state-of-the-art protocols under the umbrella of randomized measurements.

We stress that our method is well-suited for following the drifting errors in the hardware as experimentally shown in App. D 1. Performing a calibration at the beginning of the whole experiment is not sufficient for taming and understanding the errors in the randomized measurement protocol, of which we show evidence in App. F 3.

Furthermore, our approach is not limited to the measurement of the QFI. Our results extend easily to obtain robust and unbiased estimators for arbitrary nonlinear multi-copy functionals that can be expressed as observables acting on multiple copies of the quantum state. This extends the applicability of our methodology beyond the QFI and opens up possibilities for other quantum information processing tasks, such as exploring many-body entanglement phases by measuring partial transpose moments [29]. Additionally, as the robust calibration method is memory efficient, it can be performed to measure observables such as energy estimation of the ground state of quantum chemistry Hamiltonians prepared on large-scale quantum devices [85, 86] that can be further boosted by employing common randomized measurements techniques [19]. Finally, our method could be used in combination with machine-learning approaches to learn complex phases of matter with robust shadows [87, 88].

Acknowledgements

We thank S. Flammia for his valuable comments on our manuscript. Work in Grenoble is funded by the French National Research Agency via the JCJC project QRand (ANR-20-CE47-0005), and via the research programs EPIQ (ANR-22-PETQ-0007, Plan France 2030), and QUBITAF (ANR-22-PETQ-0004, Plan France 2030). B.V. acknowledges funding from the Austrian Science Foundation (FWF, P 32597 N). A.R. acknowledges support by Laboratoire d'excellence LANEF in Grenoble (ANR-10-LABX-51-01) and from the Grenoble Nanoscience Foundation. A.E. acknowledges funding by the German National Academy of Sciences Leopoldina under grant number LPDS 2021-02 and by the Walter Burke Institute for Theoretical Physics at Caltech. For some of our numerical simulations, we used the quantum toolbox QuTiP [89].

[1] S. L. Braunstein and C. M. Caves, *Physical Review Letters* **72**, 3439 (1994).
 [2] P. Hauke, M. Heyl, L. Tagliacozzo, and P. Zoller, *Nature Physics* **12**, 778 (2016).
 [3] S. Pappalardi, A. Russomanno, A. Silva, and R. Fazio, *Journal of Statistical Mechanics: Theory and Experi-*

ment **2017**, 053104 (2017).
 [4] A. Smerzi, *Physical Review Letters* **109**, 150410 (2012).
 [5] G. Tóth, *Physical Review A* **85**, 22322 (2012).
 [6] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, *Physical Review A* **85**, 22321 (2012).

- [7] L. Pezzè, M. Gabbriellini, L. Lepori, and A. Smerzi, *Physical Review Letters* **119**, 250401 (2017).
- [8] Z. Ren, W. Li, A. Smerzi, and M. Gessner, *Physical Review Letters* **126**, 80502 (2021).
- [9] V. Katariya and M. M. Wilde, *Quantum Information Processing* **20**, 78 (2021).
- [10] T.-L. Wang, L.-N. Wu, W. Yang, G.-R. Jin, N. Lambert, and F. Nori, *New Journal of Physics* **16**, 063039 (2014).
- [11] K. Macieszczak, M. Guță, I. Lesanovsky, and J. P. Garrahan, *Physical Review A* **93**, 022103 (2016).
- [12] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, *Reviews of Modern Physics* **90**, 35005 (2018).
- [13] S. L. Braunstein, C. M. Caves, and G. J. Milburn, *Annals of Physics* **247**, 135 (1996).
- [14] A. Rath, C. Branciard, A. Minguzzi, and B. Vermersch, *Physical Review Letters* **127**, 260501 (2021).
- [15] A. Rath, V. Vitale, S. Murciano, M. Votto, J. Dubail, R. Kueng, C. Branciard, P. Calabrese, and B. Vermersch, *PRX Quantum* **4**, 010318 (2023).
- [16] D. E. Koh and S. Grewal, *Quantum* **6**, 776 (2022).
- [17] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, *PRX Quantum* **2**, 030348 (2021).
- [18] E. van den Berg, Z. K. Mineev, and K. Temme, *Physical Review A* **105**, 032620 (2022).
- [19] B. Vermersch, A. Rath, B. Sundar, C. Branciard, J. Preskill, and A. Elben, *PRX Quantum* **5**, 010352 (2024).
- [20] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, *Science* **364**, 260 (2019).
- [21] A. Rath, R. van Bijnen, A. Elben, P. Zoller, and B. Vermersch, *Physical Review Letters* **127**, 200503 (2021).
- [22] K. J. Satzinger, Y.-J. Liu, A. Smith, C. Knapp, M. Newman, C. Jones, Z. Chen, C. Quintana, X. Mi, A. Dunsworth, C. Gidney, I. Aleiner, F. Arute, K. Arya, J. Atalaya, Babbush, *et al.*, *Science* **374**, 1237 (2021).
- [23] X.-D. Yu, S. Imai, and O. Gühne, *Physical Review Letters* **127**, 060504 (2021).
- [24] V. Vitale, A. Elben, R. Kueng, A. Neven, J. Carrasco, B. Kraus, P. Zoller, P. Calabrese, B. Vermersch, and M. Dalmonte, *SciPost Physics* **12**, 106 (2022).
- [25] *Nature* **622**, 481 (2023).
- [26] L. K. Joshi, J. Franke, A. Rath, F. Ares, S. Murciano, F. Kranzl, R. Blatt, P. Zoller, B. Vermersch, P. Calabrese, C. F. Roos, and M. K. Joshi, *Phys. Rev. Lett.* **133**, 010402 (2024).
- [27] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, *Physical Review Letters* **125**, 200501 (2020).
- [28] A. Neven, J. Carrasco, V. Vitale, C. Kokail, A. Elben, M. Dalmonte, P. Calabrese, P. Zoller, B. Vermersch, R. Kueng, and B. Kraus, *npj Quantum Information* **7**, 152 (2021).
- [29] J. Carrasco, M. Votto, V. Vitale, C. Kokail, A. Neven, P. Zoller, B. Vermersch, and B. Kraus, “Entanglement phase diagrams from partial transpose moments,” (2024).
- [30] A. Elben, B. Vermersch, R. Van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, *Physical Review Letters* **124**, 10504 (2020).
- [31] D. Zhu, Z. P. Cian, C. Noel, A. Risinger, D. Biswas, L. Egan, Y. Zhu, A. M. Green, C. H. Alderete, N. H. Nguyen, Q. Wang, A. Maksymov, Y. Nam, M. Cetina, N. M. Linke, M. Hafezi, and C. Monroe, *Nature Communications* **13**, 6620 (2022).
- [32] M. K. Joshi, C. Kokail, R. van Bijnen, F. Kranzl, T. V. Zache, R. Blatt, C. F. Roos, and P. Zoller, “Exploring large-scale entanglement in quantum simulation,” (2023), [arXiv:2306.00057](https://arxiv.org/abs/2306.00057).
- [33] IBM-Quantum, “<https://quantum-computing.ibm.com/>,” (2021).
- [34] E. Farhi, J. Goldstone, and S. Gutmann, “A quantum approximate optimization algorithm,” (2014), [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [35] R. A. Fisher, *Philosophical transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character* **222**, 309 (1922).
- [36] R. A. Fisher, in *Mathematical proceedings of the Cambridge philosophical society*, Vol. 22 (Cambridge University Press, 1925) pp. 700–725.
- [37] H. Strobel, W. Muessel, D. Linnemann, T. Zibold, D. B. Hume, L. Pezzè, A. Smerzi, and M. K. Oberthaler, *Science* **345**, 424 (2014).
- [38] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, and J. J. Bollinger, *Science* **352**, 1297 (2016).
- [39] L. Pezzè, Y. Li, W. Li, and A. Smerzi, *Proceedings of the National Academy of Sciences* **113**, 11459 (2016).
- [40] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, *Physical Review Letters* **106**, 130506 (2011).
- [41] G. Barontini, L. Hohmann, F. Haas, J. Estève, and J. Reichel, *Science* **349**, 1317 (2015).
- [42] R. Schmied, J.-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, and N. Sangouard, *Science* **352**, 441 (2016).
- [43] L. Pezzè and A. Smerzi, *Physical Review Letters* **102**, 100401 (2009).
- [44] M. Gessner, A. Smerzi, and L. Pezzè, *Physical Review Letters* **122**, 090503 (2019).
- [45] M. Yu, Y. Liu, P. Yang, M. Gong, Q. Cao, S. Zhang, H. Liu, M. Heyl, T. Ozawa, N. Goldman, *et al.*, *npj Quantum Information* **8**, 56 (2022).
- [46] R. O’Donnell and J. Wright, in *STOC’16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing* (ACM, New York, 2016) pp. 899–912.
- [47] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, *IEEE Transactions on Information Theory* **63**, 5628 (2017).
- [48] S. T. Flammia and R. O’Donnell, “Quantum chi-squared tomography and mutual information testing,” (2024).
- [49] R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzè, and A. Smerzi, *Physical Review Letters* **107**, 080504 (2011).
- [50] A. Rivas and A. Luis, *Physical Review A* **77**, 063813 (2008).
- [51] A. Rivas and A. Luis, *Physical Review Letters* **105**, 010403 (2010).
- [52] C. Zhang, B. Yadin, Z.-B. Hou, H. Cao, B.-H. Liu, Y.-F. Huang, R. Maity, V. Vedral, C.-F. Li, G.-C. Guo, and D. Girolami, *Physical Review A* **96**, 042327 (2017).
- [53] D. Girolami and B. Yadin, *Entropy* **19**, 124 (2017).
- [54] J. L. Beckey, M. Cerezo, A. Sone, and P. J. Coles, *Physical Review Research* **4**, 013083 (2022).
- [55] M. Cerezo, A. Sone, J. L. Beckey, and P. J. Coles, *Quantum Science and Technology* **6**, 035008 (2021).

- [56] M. Yu, D. Li, J. Wang, Y. Chu, P. Yang, M. Gong, N. Goldman, and J. Cai, *Physical Review Research* **3**, 043122 (2021).
- [57] F. Mezzadri, “How to generate random matrices from the classical compact groups,” (2007), [arXiv:math-ph/0609050](https://arxiv.org/abs/math-ph/0609050).
- [58] H.-Y. Huang, R. Kueng, and J. Preskill, *Nature Physics* **16**, 1050 (2020).
- [59] W. Hoeffding, in *Breakthroughs in Statistics* (Springer, 1992) pp. 308–334.
- [60] M. K. Joshi, A. Elben, B. Vermersch, T. Brydges, C. Maier, P. Zoller, R. Blatt, and C. F. Roos, *Physical Review Letters* **124**, 240505 (2020).
- [61] Z.-P. Cian, H. Dehghani, A. Elben, B. Vermersch, G. Zhu, M. Barkeshli, P. Zoller, and M. Hafezi, *Physical Review Letters* **126**, 050501 (2021).
- [62] A. Elben, J. Yu, G. Zhu, M. Hafezi, F. Pollmann, P. Zoller, and B. Vermersch, *Science Advances* **6**, eaaz3666 (2020).
- [63] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, *Nature Reviews Physics* **5**, 9 (2023).
- [64] D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu, *Physical Review Letters* **80**, 1121 (1998).
- [65] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, *Nature* **390**, 575 (1997).
- [66] D. Gottesman and I. L. Chuang, *Nature* **402**, 390 (1999).
- [67] E. Knill, R. Laflamme, and W. H. Zurek, *Science* **279**, 342 (1998).
- [68] C. H. Bennett, G. Brassard, and N. D. Mermin, *Physical Review Letters* **68**, 557 (1992).
- [69] V. Giovannetti, S. Lloyd, and L. Maccone, *Physical Review Letters* **96**, 010401 (2006).
- [70] G. Tóth and I. Apellaniz, *Journal of Physics A: Mathematical and Theoretical* **47**, 424006 (2014).
- [71] C. Song, K. Xu, W. Liu, C.-p. Yang, S.-B. Zheng, H. Deng, Q. Xie, K. Huang, Q. Guo, L. Zhang, P. Zhang, D. Xu, D. Zheng, X. Zhu, H. Wang, Y.-A. Chen, C.-Y. Lu, S. Han, and J.-W. Pan, *Physical Review Letters* **119**, 180511 (2017).
- [72] G. J. Mooney, G. A. L. White, C. D. Hill, and L. C. L. Hollenberg, *Journal of Physics Communications* **5**, 095004 (2021).
- [73] X.-L. Wang, Y.-H. Luo, H.-L. Huang, M.-C. Chen, Z.-E. Su, C. Liu, C. Chen, W. Li, Y.-Q. Fang, X. Jiang, J. Zhang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, *Physical Review Letters* **120**, 260502 (2018).
- [74] C. Song, K. Xu, H. Li, Y.-R. Zhang, X. Zhang, W. Liu, Q. Guo, Z. Wang, W. Ren, J. Hao, H. Feng, H. Fan, D. Zheng, D.-W. Wang, H. Wang, and S.-Y. Zhu, *Science* **365**, 574 (2019).
- [75] M. Gong, M.-C. Chen, Y. Zheng, S. Wang, C. Zha, H. Deng, Z. Yan, H. Rong, Y. Wu, S. Li, F. Chen, Y. Zhao, F. Liang, J. Lin, Y. Xu, C. Guo, L. Sun, A. D. Castellano, H. Wang, C. Peng, C.-Y. Lu, X. Zhu, and J.-W. Pan, *Physical Review Letters* **122**, 110501 (2019).
- [76] I. Pogorelov, T. Feldker, C. D. Marciniak, L. Postler, G. Jacob, O. Kriegelsteiner, V. Podlesnic, M. Meth, V. Negnevitsky, M. Stadler, B. Höfer, C. Wächter, K. Lakhmanskii, R. Blatt, P. Schindler, and T. Monz, *PRX Quantum* **2**, 020343 (2021).
- [77] S. Cao, B. Wu, F. Chen, M. Gong, Y. Wu, Y. Ye, C. Zha, H. Qian, C. Ying, S. Guo, Q. Zhu, H.-L. Huang, Y. Zhao, S. Li, S. Wang, J. Yu, D. Fan, D. Wu, H. Su, H. Deng, H. Rong, Y. Li, K. Zhang, T.-H. Chung, F. Liang, J. Lin, Y. Xu, L. Sun, C. Guo, N. Li, Y.-H. Huo, C.-Z. Peng, C.-Y. Lu, X. Yuan, X. Zhu, and J.-W. Pan, *Nature* **619**, 738 (2023).
- [78] A. Omran, H. Levine, A. Keesling, G. Semeghini, T. T. Wang, S. Ebadi, H. Bernien, A. S. Zibrov, H. Pichler, S. Choi, *et al.*, *Science* **365**, 570 (2019).
- [79] K. X. Wei, I. Lauer, S. Srinivasan, N. Sundaresan, D. T. McClure, D. Toyli, D. C. McKay, J. M. Gambetta, and S. Sheldon, *Physical Review A* **101**, 032343 (2020).
- [80] M. Gabbriellini, A. Smerzi, and L. Pezzè, *Scientific Reports* **8**, 15663 (2018).
- [81] I. Frérot and T. Roscilde, *Physical review letters* **121**, 020402 (2018).
- [82] B. Koczor, S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin, *New Journal of Physics* **22**, 083038 (2020).
- [83] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, *Nature Reviews Physics* **3**, 625 (2021).
- [84] W. W. Ho and T. H. Hsieh, *SciPost Physics* **6**, 029 (2019).
- [85] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, *Physical Review X* **8**, 031022 (2018).
- [86] H.-Y. Huang, R. Kueng, and J. Preskill, *Physical Review Letters* **127**, 030503 (2021).
- [87] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, *Science* **377**, eabk3333 (2022).
- [88] L. Lewis, H.-Y. Huang, V. T. Tran, S. Lehner, R. Kueng, and J. Preskill, *nature communications* **15**, 895 (2024).
- [89] J. Johansson, P. Nation, and F. Nori, *Computer Physics Communications* **184**, 1234 (2013).
- [90] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, *Physical Review A* **99**, 052323 (2019).
- [91] B. Baheri, Q. Guan, V. Chaudhary, and A. Li, in *2022 IEEE 28th International Symposium on On-Line Testing and Robust System Design (IOLTS)* (2022) pp. 1–5.
- [92] Y. Hirasaki, S. Daimon, T. Itoko, N. Kanazawa, and E. Saitoh, *Applied Physics Letters* **123** (2023).
- [93] $\sigma_j^{(\tau)}$ is the Pauli matrix in an arbitrary direction (τ) acting on the j^{th} spin (identity operators on the other qubits are implicit).
- [94] Note that F_Q is independent of θ [12].
- [95] One can also implement different unitaries for calibration and estimation. For the present experiment, we did not notice significant differences as we estimated different quantities.

Appendices

Contents

A. Converging series of lower bounds of the QFI	13
B. Quantum properties under global depolarization	13
C. Derivation of the robust shadow estimator with local noise	14
1. Robust shadow from randomized measurements	14
2. Calibration step	16
3. Experimental comparison of the estimators \hat{G}_j and $\hat{\mathcal{G}}_j$	17
D. Experimental results on the noise	18
1. Verification of the time dependence of the noise in ‘ibm_prague’	18
2. Check on the origin of the noise	19
E. Verification of the validity of the assumption of local noise	20
1. Derivation of the estimator of locality of noise	20
2. Application to a two-qubit readout error model	22
3. Experimental investigation on the employed platform	22
F. Further experimental results	24
1. Comparison with previous work in Ref. [56]	24
2. Ground state of the TFIM at the critical point	25
3. Comparison of error mitigation protocols	25
G. Numerical investigations	26
1. Classical and Quantum Fisher information comparison in noisy GHZ states	26
2. \mathcal{G}_j as a function of the number of unitaries and readout error	28
3. Scaling of the measurement budget for the lower bound F_2	28
4. Numerical simulation of the experiment for perfect GHZ states and readout errors	29

Appendix A: Converging series of lower bounds of the QFI

As shown in Ref. [14], the QFI can be expanded in terms of a Taylor series in the eigenvalues λ_μ of the density matrix $\rho = \sum_\mu \lambda_\mu |\mu\rangle\langle\mu|$. This reads as:

$$F_Q = 2 \sum_{\ell=0}^{\infty} \sum_{(\mu,\nu), \lambda_\mu + \lambda_\nu > 0} (\lambda_\mu - \lambda_\nu)^2 (1 - \lambda_\mu - \lambda_\nu)^\ell |\langle\mu| A |\nu\rangle|^2. \quad (\text{A1})$$

We note that each term in the infinite sum is positive. Truncating the summation at a finite value n , we thus obtain a converging series of polynomial lower bounds F_n that can be measured experimentally:

$$F_n = 2 \sum_{\ell=0}^n \sum_{(\mu,\nu), \lambda_\mu + \lambda_\nu > 0} (\lambda_\mu - \lambda_\nu)^2 (1 - \lambda_\mu - \lambda_\nu)^\ell |\langle\mu| A |\nu\rangle|^2 = 2 \sum_{q=0}^n \binom{n+1}{q+1} (-1)^q \sum_{m=0}^{q+2} C_m^{(q)} \text{Tr}(\rho^{q+2-m} A \rho^m A), \quad (\text{A2})$$

where we have introduced the coefficients $C_m^{(q)} = \binom{q}{m} - 2\binom{q}{m-1} + \binom{q}{m-2}$, with the binomial coefficients defined such that $\binom{q}{m'} = 0$ if $m' < 0$ or $m' > q$. The last equality can be proven by injecting the eigenvalue decomposition of ρ in the right-hand side and rearranging the sums [14].

Appendix B: Quantum properties under global depolarization

Let us consider a quantum state defined as $\rho(p_D) = (1 - p_D) |\psi\rangle\langle\psi| + p_D \mathbb{1}/2^N$, where $|\psi\rangle$ is a pure state and $\mathbb{1}/2^N$ is the fully mixed state. The state $\rho(p_D)$ is mixed with global depolarizing noise of strength p_D . The distinct eigenvalues

of $\rho(p_D)$ are $\lambda_1 = (1 - p_D) + p_D/2^N$ (with multiplicity 1) and $\lambda_2 = p_D/2^N$ (with multiplicity $2^N - 1$). As shown previously in [6], for $\rho(p_D)$, the QFI is given (replacing the eigenvalues in Eq. (1)) by

$$F_Q = 4 \left(\langle \psi | A^2 | \psi \rangle - \langle \psi | A | \psi \rangle^2 \right) \frac{(1 - p_D)^2}{1 - p_D + 2p_D/2^N}. \quad (\text{B1})$$

Similarly, we notice that for this specific state, all non-zero terms in Eq. (A2) are equivalent to $(1 - p_D)^2 (p_D - 2p_D/2^N)^\ell$. Thus we can provide an expression of the lower bounds F_n under global depolarization noise as:

$$F_n = 4 \left(\langle \psi | A^2 | \psi \rangle - \langle \psi | A | \psi \rangle^2 \right) (1 - p_D)^2 \sum_{\ell=0}^n (p_D - 2p_D/2^N)^\ell \quad (\text{B2})$$

Analogously, we can also express the analytical form of the purity for the state $\rho(p_D)$ as

$$\text{Tr}(\rho(p_D)^2) = (1 - p_D)^2 + \frac{2p_D - p_D^2}{2^N}. \quad (\text{B3})$$

Assuming this specific noise model, one can perform multiple interesting analyses with the above analytical expressions. One such investigation is to better understand the convergence of the lower bounds F_n to a finite value of QFI as a function of the bound order n . For this purpose, we could use the experimentally recorded values of purity to invert the above purity expression and extract values of p_D for a given state of interest. We perform this analysis for the 8 and 10 qubit GHZ states as shown in Fig. 2(a). The mitigated values of purities as estimated and shown in Fig. 2(a) are respectively 0.89 ± 0.0066 and 0.86 ± 0.017 which give the corresponding values of p_D as 0.056 ± 0.007 and 0.072 ± 0.018 for the 8 and 10 qubit GHZ states. One can estimate easily the theoretical value of the QFI using Eq. (B1) under this noise assumption. This presents an easy method to check the convergence of the lower bounds to a finite value of QFI as presented in Fig. 2 of the main text.

Appendix C: Derivation of the robust shadow estimator with local noise

In this section we construct the robust classical shadow estimator given in Eq. (3), equivalent to the one presented for the first time in Ref. [17]. We consider a situation where we have performed randomized measurements on a N -qubit state ρ , which are affected by noise. We assume that the noise is gate-independent, Markovian and stationary within each iteration, and that it occurs between the random unitaries and the measurements (not before the unitaries). This ensures that we can model noisy randomized measurements as $\mathcal{M} \circ \Lambda^{(i)} \circ \mathcal{U}^{(r)}$ where $\mathcal{U}^{(r)}$ is the ideal unitary channel describing the application of ideal random unitaries $U^{(r)}$, $\Lambda^{(i)}$ is the noise channel in iteration i , encapsulating gate noise and readout errors, and \mathcal{M} is the measurement channel, describing an ideal computational basis measurement [17]. In addition, we assume local noise, i.e. the noise channel decomposes as $\Lambda^{(i)} = \Lambda_1^{(i)} \otimes \dots \otimes \Lambda_N^{(i)}$ and local random unitaries, i.e. the ideal unitary channel is realized with local unitary transformations $U^{(r)} = U_1^{(r)} \otimes \dots \otimes U_N^{(r)}$. Here, the local unitaries $U_j^{(r)}$ are sampled independently and uniformly from the circular unitary ensemble, i.e. the Haar measure on the unitary group $U(2)$.

As described in the main text, we employ first a calibration protocol, equivalent to the one described in Ref. [17], to characterize the local noise channel $\Lambda^{(i)}$ in terms of N parameters $G_j^{(i)}$. To perform this calibration, we assume that the state $|0\rangle = |0\rangle^{\otimes N}$ can be prepared with a high fidelity in our experiment. The calibration results are then used to build an unbiased estimator $\hat{\rho}$ of the density matrix ρ – a robust classical shadow – from randomized measurements performed on ρ , that mitigates the noise errors induced by $\Lambda^{(i)}$.

In the remainder of this section, we drop the superscript i denoting the iteration of the experiment to simplify the notation.

1. Robust shadow from randomized measurements

Under the noise assumptions described above, noisy randomized measurements provide access to the probability distribution of the measured bit-strings $\mathbf{s} = (s_1, \dots, s_N)$, conditioned on the application of a local random unitary $U^{(r)} = U_1^{(r)} \otimes \dots \otimes U_N^{(r)}$:

$$P_\Lambda(\mathbf{s}|U^{(r)}) = \langle \mathbf{s} | \Lambda(U^{(r)} \rho U^{(r)\dagger}) | \mathbf{s} \rangle = \text{Tr} \left(\rho U^{(r)\dagger} \Lambda^*(|\mathbf{s}\rangle\langle\mathbf{s}|) U^{(r)} \right), \quad (\text{C1})$$

where Λ is the trace-preserving noise channel and Λ^* is its adjoint. We aim to construct an unbiased estimator of ρ – robust classical shadow – in terms of the statistics of $P_\Lambda(\mathbf{s}|U^{(r)})$. We choose an ansatz of the form

$$\begin{aligned}\tilde{\rho}^{(r)} &= \sum_{\mathbf{s}} P_\Lambda(\mathbf{s}|U^{(r)}) U^{(r)\dagger} O(\mathbf{s}) U^{(r)} \\ &= \sum_{\mathbf{s}} \text{Tr} \left(\rho U^{(r)\dagger} \Lambda^*(|\mathbf{s}\rangle\langle\mathbf{s}|) U^{(r)} \right) U^{(r)\dagger} O(\mathbf{s}) U^{(r)} \\ &= \sum_{\mathbf{s}} \text{Tr}_1 \left((\rho \otimes \mathbb{1}) (U^{(r)\dagger})^{\otimes 2} [\Lambda^*(|\mathbf{s}\rangle\langle\mathbf{s}|) \otimes O(\mathbf{s})] U^{(r)\otimes 2} \right)\end{aligned}\tag{C2}$$

with $O(\mathbf{s}) = \bigotimes_j O_j(s_j)$ being a local hermitian operator, which we take to be diagonal in the computational basis, and Tr_1 denoting the partial trace over the first copy of the N -qubit system. The ensemble average over the random unitaries $U^{(r)}$ yields

$$\mathbb{E}[\tilde{\rho}^{(r)}] = \text{Tr}_1 \left[(\rho \otimes \mathbb{1}) \Phi^{(2)} \left(\sum_{\mathbf{s}} \Lambda^*(|\mathbf{s}\rangle\langle\mathbf{s}|) \otimes O(\mathbf{s}) \right) \right] = \text{Tr}_1 \left[(\rho \otimes \mathbb{1}) \Phi^{(2)}(Q) \right]\tag{C3}$$

with

$$Q = \sum_{\mathbf{s}} \Lambda^*(|\mathbf{s}\rangle\langle\mathbf{s}|) \otimes O(\mathbf{s}) = \bigotimes_{j=1}^N \left[\sum_{s_j=0,1} \Lambda_j^*(|s_j\rangle\langle s_j|) \otimes O_j(s_j) \right] = \bigotimes_{j=1}^N Q_j.\tag{C4}$$

Here, we used the local noise assumption (noting that $\Lambda^* = (\bigotimes_j \Lambda_j)^* = \bigotimes_j \Lambda_j^*$) and $\Phi^{(2)}(\cdot) = \mathbb{E}[(U^{(r)\dagger})^{\otimes 2}(\cdot)U^{(r)\otimes 2}]$ denotes the two-copy local unitary ‘twirling channel’ [90]. It evaluates to

$$\Phi^{(2)}(Q) = \left(\frac{1}{3}\right)^N \bigotimes_{j=1}^N \left(\left(\text{Tr}(Q_j) - \frac{1}{2} \text{Tr}(\mathbb{S}_j Q_j) \right) \mathbb{1}_j^{(2)} + \left(\text{Tr}(\mathbb{S}_j Q_j) - \frac{1}{2} \text{Tr}(Q_j) \right) \mathbb{S}_j \right)\tag{C5}$$

with the swap operator $\mathbb{S}_j = \sum_{s_{j_1}, s_{j_2}} |s_{j_2}\rangle\langle s_{j_1}| \otimes |s_{j_1}\rangle\langle s_{j_2}|$ acting on two copies of qubit j and $\mathbb{1}_j^{(2)} = \mathbb{1}_j \otimes \mathbb{1}_j$ the identity.

The estimator $\tilde{\rho}^{(r)}$ is an unbiased estimator of ρ if the average over the Haar random unitaries yields the true density matrix of the quantum state, $\mathbb{E}[\tilde{\rho}^{(r)}] = \rho$. Observing that $\text{Tr}_1((\rho \otimes \mathbb{1}) \mathbb{S}) = \rho$ where $\mathbb{S} = \bigotimes_{j=1}^N \mathbb{S}_j$ is the swap operator between two copies of the entire system, we thus find, from Eq. (C3), that the estimator is unbiased for any state ρ if and only if $\Phi^{(2)}(Q) = \mathbb{S}$, or equivalently, using Eq. (C5),

$$\left(\text{Tr}(Q_j) - \frac{1}{2} \text{Tr}(\mathbb{S}_j Q_j) \right) \mathbb{1}_j^{(2)} + \left(\text{Tr}(\mathbb{S}_j Q_j) - \frac{1}{2} \text{Tr}(Q_j) \right) \mathbb{S}_j = 3\mathbb{S}_j \quad \forall j.\tag{C6}$$

On top of the assumption that $O_j(s_j)$ is diagonal in the computational basis, we further assume that it is of the form $O_j(s_j) = \alpha_j |s_j\rangle\langle s_j| + \beta_j \mathbb{1}$, with α_j, β_j real numbers that do not depend on s_j . With this, we can evaluate the terms appearing in Eq. (C6) above as follows:

$$\begin{aligned}\text{Tr}(Q_j) &= \sum_{s_j} \text{Tr}(\Lambda^*(|s_j\rangle\langle s_j|)) \text{Tr}(O_j(s_j)) = \sum_{s_j} \text{Tr}(\Lambda^*(|s_j\rangle\langle s_j|)) (\alpha_j + 2\beta_j) \\ &= \text{Tr}(\Lambda(\mathbb{1})) (\alpha_j + 2\beta_j) = 2\alpha_j + 4\beta_j,\end{aligned}\tag{C7}$$

$$\begin{aligned}\text{Tr}(\mathbb{S}_j Q_j) &= \sum_{s_j} \text{Tr}(\Lambda^*(|s_j\rangle\langle s_j|) O_j(s_j)) \\ &= \alpha_j \sum_j \text{Tr}(\Lambda^*(|s_j\rangle\langle s_j|) |s_j\rangle\langle s_j|) + 2\beta_j = 2\alpha_j G_j + 2\beta_j,\end{aligned}\tag{C8}$$

where we have used also that the noise channel is trace preserving and $\text{Tr}(\mathbb{S}A \otimes B) = \text{Tr}(AB)$. Here, we have introduced the quantity

$$G_j = \frac{1}{2} \sum_{s_j} \langle s_j | \Lambda_j(|s_j\rangle\langle s_j|) |s_j\rangle,\tag{C9}$$

that contains all the relevant information on the noise acting on qubit j during the randomized measurement protocol, and which we interpret as the average ‘survival probability’ of the two basis states of qubit j . Thus, to characterise the noise that affects the experimental protocol we only need to learn how it acts on the computational basis states $|s_j\rangle$. With the above expressions, inverting Eq. (C6) gives

$$\alpha_j = \frac{3}{2G_j - 1}, \quad \beta_j = \frac{G_j - 2}{2G_j - 1}. \quad (\text{C10})$$

Inserting this into Eq. (C2), we finally can write the estimator $\tilde{\rho}$ as

$$\begin{aligned} \tilde{\rho}^{(r)} &= \sum_{\mathbf{s}} P_{\Lambda}(\mathbf{s}|U^{(r)}) \bigotimes_{j=1}^N \left(\alpha_j U_j^{(r)\dagger} |s_j\rangle\langle s_j| U_j^{(r)} + \beta_j \mathbb{1} \right) \\ &= \sum_{\mathbf{s}} P_{\Lambda}(\mathbf{s}|U^{(r)}) \bigotimes_{j=1}^N \left(\frac{3}{2G_j - 1} U_j^{(r)\dagger} |s_j\rangle\langle s_j| U_j^{(r)} + \frac{G_j - 2}{2G_j - 1} \mathbb{1} \right). \end{aligned} \quad (\text{C11})$$

In the absence of noise $G_j = 1, \forall j$, so that the usual formula for the estimator of the density matrix from randomized measurements is recovered: $O_j(s_j) = 3 |s_j\rangle\langle s_j| - \mathbb{1}$ [58, 62]. For a fully depolarising channel, on the other hand, one gets $G_j = 1/2$, in which case we are not able to extract any information by measuring the state as the coefficients in Eq. (C10) diverge.

2. Calibration step

The parameters in Eq. (C10) rely on the estimation of G_j . It is based on the calibration procedure described in the main text (Sec. II A). In a nutshell, the system is prepared in a state with high fidelity, namely $|\mathbf{0}\rangle \equiv |0\rangle^{\otimes N}$, and the randomized measurement protocol is applied. We show here that G_j can be directly linked to the random unitaries $U^{(r)} = U_1^{(r)} \otimes \dots \otimes U_N^{(r)}$, with $r = 1, \dots, N_U$, and the bit-strings of measurement outcomes $\mathbf{s}^{(r,m)} = (s_1^{(r,m)}, \dots, s_N^{(r,m)})$ with $m = 1, \dots, N_M$.

Let us introduce the following quantity

$$C_j = \sum_{s_j=0,1} \mathbb{E} \left[\langle s_j | \Lambda_j(U_j^{(r)}) |0\rangle\langle 0| U_j^{(r)\dagger} |s_j\rangle\langle s_j| U_j^{(r)} |0\rangle\langle 0| U_j^{(r)\dagger} |s_j\rangle \right], \quad (\text{C12})$$

where $|0\rangle$ represents the calibration state of the single qubit j and $\mathbb{E}[\cdot]$ is the average over the circular unitary ensemble. We can define an estimator as:

$$\hat{C}_j = \frac{1}{N_U} \sum_r \sum_{s_j=0,1} \hat{P}(s_j|U_j^{(r)}) P(s_j|U_j^{(r)}), \quad (\text{C13})$$

where $\hat{P}(s_j|U_j^{(r)}) = \sum_{m=1}^{N_M} \frac{\delta_{s_j, s_j^{(r,m)}}}{N_M}$ is the estimated (noisy) Born probability and $P(s_j|U_j^{(r)}) = |\langle s_j | U_j^{(r)} |0\rangle|^2$ is the theoretical (noiseless) one. The information on the noise is contained in $\hat{P}(s_j|U_j^{(r)})$, that approaches the theoretical noisy Born probability $P_{\Lambda}(s_j|U_j^{(r)})$ in the limit $N_M \rightarrow \infty$. Thus, since with our noise model $\mathbb{E}_{\text{QM}}[\hat{P}(s_j|U^{(r)})] = \langle s_j | \Lambda_j(U_j^{(r)}) |0\rangle\langle 0| U_j^{(r)\dagger} |s_j\rangle$, we have $\mathbb{E}[\mathbb{E}_{\text{QM}}[\hat{C}_j]] = C_j$, i.e., \hat{C}_j is an unbiased estimator for C_j . Here $\mathbb{E}_{\text{QM}}[\cdot]$ is the quantum mechanical average over the Born probabilities. We note that under the (idealized) assumption of strictly gate-independent noise (same noise channel Λ_j for any $U_j^{(r)}$, including the idle gate $\mathbb{1}_j$), we could measure G_j directly from its definition, Eq. (C9). In practice, we expect that C_j (and its estimator \hat{C}_j) captures the actual noise acting during the measurement stage more faithfully, as exactly the same experimental resources are employed and any weakly gate dependent noise is averaged (twirled) to yield approximately the same gate-independent average noise channel, $\mathbb{E}[\Lambda_U(\rho)] \approx \Lambda(\mathbb{E}[U\rho U^\dagger])$. We refer to more details on gate-dependent noise in Ref. [17].

Let us now link C_j and the quantity G_j in Eq. (C9). We observe that C_j can be written as

$$C_j = \mathbb{E} \left[\langle 0 |^{\otimes 2} U_j^{(r)\dagger \otimes 2} \left(\sum_{s_j=0,1} \Lambda_j^*(|s_j\rangle\langle s_j|) \otimes |s_j\rangle\langle s_j| \right) U_j^{(r) \otimes 2} |0\rangle^{\otimes 2} \right], \quad (\text{C14})$$

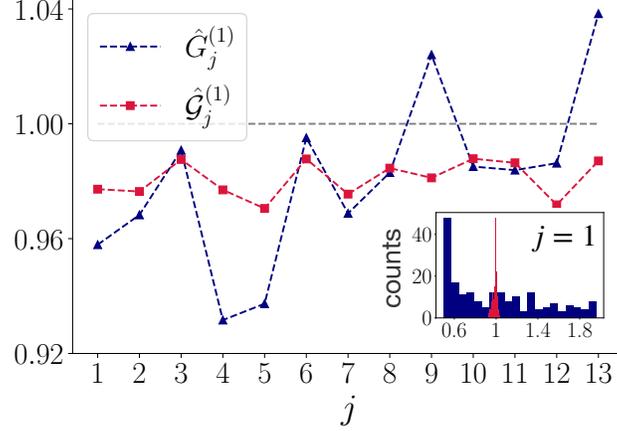


Figure 4: Comparison of G_j using the enhanced estimator \hat{G}_j and the previous estimator \hat{G}_j for a 13-qubit state on ‘ibm_prague’. The quantities are measured according to the calibration protocol described in Sec. II A and depicted in Fig. 1. $N_U^{\text{tot}} = N_I N_U = 27000$ (number of unitaries in the randomized measurement protocol) and $N_M = 1000$ (number of measurements per unitary). In the inset we compare the estimators for iteration $i = 1$, i.e. $\hat{G}_j^{(1)}$ (blue) and $\hat{G}_j^{(1)}$ (red), for the first qubit ($j = 1$) by plotting a histogram where each occurrence corresponds to an element of the sum over r in Eq. (C13) for $\hat{G}_j^{(1)}$, and of the analogous sum for $\hat{G}_j^{(1)}$.

where we have used the property $\text{Tr}(\Lambda(A)B) = \text{Tr}(A\Lambda^*(B))$. As in the previous section, we can express the average in C_j over unitaries in terms of a twirling channel $\Phi_j^{(2)}(Q_j)$ (a single-qubit version of the 2-copy channel Φ introduced before). In this case we write the two-copy operator $Q = \bigotimes_{j=1}^N Q_j$ with $Q_j = \sum_{s_j=0,1} \Lambda_j^*(|s_j\rangle\langle s_j|) \otimes |s_j\rangle\langle s_j|$. Using again the twirling formula in Eq. (C5) (now for two copies of a single qubit) one obtains

$$C_j = \langle 0|^{\otimes 2} \Phi_j^{(2)} \left(\sum_{s_j} \Lambda_j^*(|s_j\rangle\langle s_j|) \otimes |s_j\rangle\langle s_j| \right) |0\rangle^{\otimes 2} = \frac{1 + G_j}{3}. \quad (\text{C15})$$

The link between C_j and G_j being clear, one can define an estimator for G_j in terms of the one for C_j in Eq. (C13):

$$\hat{G}_j = 3 \hat{C}_j - 1. \quad (\text{C16})$$

In the absence of noise one can check that $G_j = 1$ and $C_j = \frac{2}{3}$, $\forall j = 1, \dots, N$ [58]. We remark here that all our results are compatible with the ones in Ref. [17], where a slightly different formalism has been employed.

We have discussed in Sec. II D that the statistical accuracy of the estimator of G_j can be improved by making use of *common randomized numbers* [19] to define an estimator with smaller variance with respect to \hat{G}_j .

The enhanced estimator \hat{G}_j is connected as well to the quantity C_j that we access in the experiments, as in Eq. (10). Explicitly, we write:

$$\hat{G}_j = 3 \hat{C}_j - 1 - B_j + \mathbb{E}[B_j], \quad (\text{C17})$$

where B_j has been defined in Eq. (9). Therefore the common randomized measurements procedure enters solely into the post-processing. The variance of estimators obtained from such practice has been studied analytically in Ref. [19], where it was shown that involving positively correlated random variables, as above, indeed allows one to significantly reduce the variance upper bounds. In the following section we compare based on our experimental data this estimator with the one introduced in Eq. (C16).

3. Experimental comparison of the estimators \hat{G}_j and \hat{G}_j

Let us consider the $N = 13$ qubit experiment that has been performed on the ‘ibm_prague’ processor. We have performed a calibration of the device as described in Sec. II A and depicted in the step (i) of Fig. 1. For each iteration

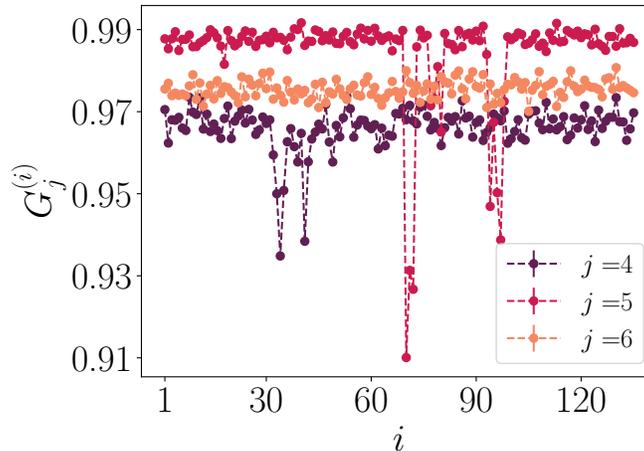


Figure 5: $G_j^{(i)}$ as a function of the iteration i for a 13-qubit state on ‘ibm_prague’. The quantity is estimated from the calibration data of the protocol depicted in Fig. 1, using the estimator $\hat{G}_j^{(i)}$ of Eq. (11). Here $N_U^{\text{tot}} = N_I N_U = 27000$ (number of unitaries in the randomized measurement protocol) and $N_M = 1000$ (number of measurements per unitary). We present the result for qubits $j = 4, 5, 6$.

$i = 1, \dots, N_I$ and for each applied unitary $U^{(r_i)}$ ($r_i = 1, \dots, N_U$), we collect $N_M = 1000$ bit-strings of measurement outcomes. From the unitaries and the bit-strings we compute the quantities $\hat{G}_j^{(i)}$ and $\hat{G}_j^{(i)}$ as defined above, which contain the information about the local errors in the measurement protocol within each iteration i .

In Fig. 4 we show the comparison of the two estimators, for iteration $i = 1$ and all the qubits ($j = 1, \dots, 13$). In the inset we show a comparison of the histograms of the values that build up $\hat{G}_j^{(1)}$ and $\hat{G}_j^{(1)}$ for the first qubit ($j = 1$), where each point corresponds to an element of the sum over r in Eq. (C13) for $\hat{G}_1^{(1)}$, and of the analogous sum for $\hat{G}_1^{(1)}$. Remarkably, we observe that the contributions to $\hat{G}_1^{(1)}$ are much less spread than those of $\hat{G}_1^{(1)}$; in particular the contributions to $\hat{G}_1^{(1)}$ range in $\sim (0.6, 1.8)$, while the $\hat{G}_1^{(1)}$ counts are sharply peaked around ~ 1 . We argue this is due to the trick of common random numbers [19] employed to define $\hat{G}_j^{(i)}$, which in general allows to reduce the variance of the estimator. The same holds for any qubit j .

Appendix D: Experimental results on the noise

In this section we perform an experimental analysis on the noise in the quantum platform we employ. We investigate the time dependence of the noise, noticing huge fluctuations in the quantities we use to estimate the errors, and we observe that the most important contribution to the single qubit error can be identified to be caused due to readout errors.

1. Verification of the time dependence of the noise in ‘ibm_prague’

Let us consider again the $N = 13$ qubit experiment that has been performed on the ‘ibm_prague’ processor. In Fig. 5 we study the behaviour of $G_j^{(i)}$ (estimated through $\hat{G}_j^{(i)}$ of Eq. (11)) as a function of the iterations i . The error in the quantum device fluctuates in time, we want to verify that it is important to perform consecutive iterations of experiments to account for the temporal variations in gate and readout errors instead of performing a single calibration in advance. We plot $G_j^{(i)}$ as a function of i for three different qubits, labeled by j . For $j = 4, 5$, we observe fluctuating events given $G_j^{(i)}$ as a function of the iterations i , hinting that it is important to follow the temporal fluctuations of the noise to provide reliable and robust estimations. This is not the case for all the qubits; e.g., we do not see such fluctuations for $j = 6$ in the plot. Similar effects have been observed in other type of error mitigation protocols with superconducting qubits [91, 92].

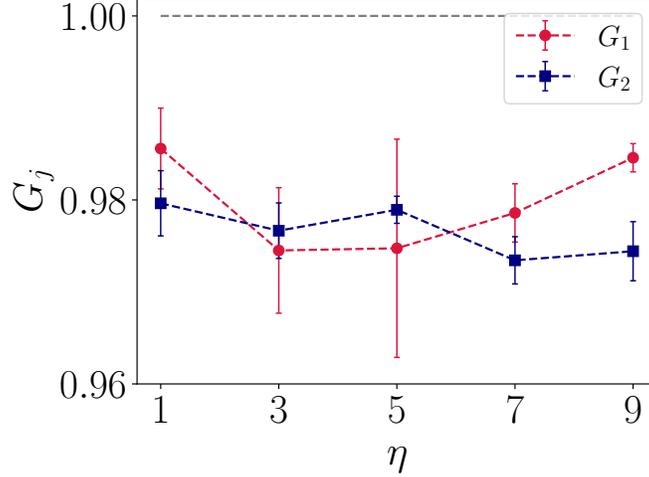


Figure 6: G_j for a two-qubit system realized on ‘ibm_lagos’, as a function of the number of layers of unitaries η applied to the initial state. We use the estimator \hat{G}_j of Eq. (11). A value compatible with 1 means that the noise can be neglected. Here $N_U = 800$ (number of unitaries in the randomized measurement protocol) and $N_M = 1000$ (number of measurements per unitary).

2. Check on the origin of the noise

Our aim here is to study what is the most important source of errors in the randomized measurement protocol. In Fig. 6 we consider a two-qubit system realized on the ‘ibm_lagos’ processor. We employ the calibration method described in Sec. II A and depicted in Fig. 1. In order to discriminate between the various sources of noises, instead of applying a single unitary $U = \bigotimes_{j=1}^N U_j$, we employ several layers of unitaries, given by the number η that are sampled independently and uniformly from the circular unitary ensemble. We measure the quantity G_j as a function of the parameter η using the enhanced estimator \hat{G}_j of Eq. (11). The rationale behind this approach is the following. We can write the noise channel Λ acting during the measurement protocol as two separate contributions: one due to errors on the unitaries Λ_U and one due to the readout Λ_{meas} with $\Lambda = \Lambda_{\text{meas}} \circ \Lambda_U$. By applying η layers of unitaries one would get $\Lambda(\eta) = \Lambda_{\text{meas}} \circ (\Lambda_U)^\eta$. Following the effect of the noise as a function of η , we may be able to discriminate the contributions of Λ_U and Λ_{meas} . This idea can be formalized based on a simple noise model defined by

$$\begin{aligned}\Lambda_U(\rho_j) &= (1 - p_U)\rho_j + \frac{p_U}{3} \sum_{\alpha} \sigma_j^\alpha \rho_j \sigma_j^\alpha, \\ \Lambda_{\text{meas}}(\rho_j) &= (1 - p_{\text{meas}})\rho_j + p_{\text{meas}} \sigma_j^x \rho_j \sigma_j^x.\end{aligned}\tag{D1}$$

Here $\sigma^\alpha = \sigma^x, \sigma^y, \sigma^z$ are single qubit Pauli matrices with ρ_j being a single qubit density matrix. The action of the unitary gates is modeled as a depolarizing noise channel Λ_U with parameter p_U , while the readout errors are described by bit flips that happen with probability p_{meas} . The full channel $\Lambda(\eta)(\rho_j)$ applied on a single qubit state ρ_j gives

$$\Lambda(\eta)(\rho_j) = (1 - \eta p_U - p_{\text{meas}})\rho_j + \frac{p_U \eta}{3} \sum_{\alpha} \sigma_j^\alpha \rho_j \sigma_j^\alpha + p_{\text{meas}} \sigma_j^x \rho_j \sigma_j^x.\tag{D2}$$

We can compute explicitly the behavior of G_j at first order in $p_U, p_{\text{meas}} \ll 1$ and obtain

$$G_j(\eta) = 1 - \frac{2p_U}{3}\eta - p_{\text{meas}}.\tag{D3}$$

We observe that the unitary contribution would monotonically decrease G_j as a function of η while the readout error yields a fixed shift by p_{meas} . From Fig. 6, we observe that G_j remains essentially constant within error bars for different values of η , hence increasing the number of unitaries does not induce more noise (in terms of the parameter G_j) in the system. This suggests that the most relevant contribution to the noise in the randomized measurement protocol is due to readout errors.

Appendix E: Verification of the validity of the assumption of local noise

In this section we propose a method to test the assumption of a local noise channel, i.e. $\Lambda = \bigotimes_{j=1}^N \Lambda_j$, that is based on analysing the statistical correlations among qubit pairs. We employ the calibration data that has been used for the mitigation of the QFI results on the prepared GHZ states. The section is structure as follows: At first we drop the assumption of locality, i.e. we consider a general noise channel Λ , and introduce a quantity \tilde{R} that can be used for testing its locality; then, we provide an illustrative analytical example in the case of cross-talk errors for two qubits; finally, we show an experimental indication of the validity of the assumption of locality.

1. Derivation of the estimator of locality of noise

Let us start by extending Eq. (C12) to measurements that act on the whole device, writing

$$\begin{aligned} \tilde{C}_j &= \mathbb{E} \left[\sum_{s_j} \text{Tr} \left[\langle s_j | \Lambda(U^{(r)} | \mathbf{0} \rangle \langle \mathbf{0} | U^{(r)\dagger}) | s_j \rangle \right] P(s_j | U_j^{(r)}) \right] \\ &= \mathbb{E} \left[\sum_{s_j} \text{Tr} \left[|s_j\rangle \langle s_j| \text{Tr}_{k \neq j} (\Lambda(U^{(r)} | \mathbf{0} \rangle \langle \mathbf{0} | U^{(r)\dagger})) \right] P(s_j | U_j^{(r)}) \right], \end{aligned} \quad (\text{E1})$$

where \mathbb{E} denotes the average over all local unitaries $U_k^{(r)}$'s and again $P(s_j | U_j^{(r)}) = |\langle s_j | U_j^{(r)} | 0 \rangle|^2$. The latter corresponds to the C_j introduced in Eq. (C12) if $\Lambda = \bigotimes_{j=1}^N \Lambda_j$ and can be estimated from the calibration data as explained in Sec. C2, according to Eq. (C13). If we perform an average over all local random unitaries $U_k^{(r)}$ with $k \neq j$ (denoted as $\mathbb{E}_{\{k \neq j\}}$), we can exploit the twirling identity for a single-qubit operator O_j , $\Phi_j^{(1)}(O_j) = \mathbb{E}[U_j^{(r)} O_j U_j^{(r)\dagger}] = \frac{1}{2} \text{Tr}(O_j)$, such that

$$\begin{aligned} \mathbb{E}_{\{k \neq j\}} [U^{(r)} | \mathbf{0} \rangle \langle \mathbf{0} | U^{(r)\dagger}] &= \mathbb{E} \left[U_1^{(r)} | 0 \rangle \langle 0 | U_1^{(r)\dagger} \right] \otimes \dots \otimes U_j^{(r)} | 0 \rangle \langle 0 | U_j^{(r)\dagger} \otimes \dots \otimes \mathbb{E} \left[U_N^{(r)} | 0 \rangle \langle 0 | U_N^{(r)\dagger} \right] \\ &= \Phi_1^{(1)}(|0\rangle\langle 0|) \otimes \dots \otimes U_j^{(r)} | 0 \rangle \langle 0 | U_j^{(r)\dagger} \otimes \dots \otimes \Phi_N^{(1)}(|0\rangle\langle 0|) \\ &= \mathbb{1}/2 \otimes \dots \otimes \mathbb{1}/2 \otimes U_j^{(r)} | 0 \rangle \langle 0 | U_j^{(r)\dagger} \otimes \mathbb{1}/2 \otimes \dots \otimes \mathbb{1}/2. \end{aligned} \quad (\text{E2})$$

and write

$$\tilde{C}_j = \mathbb{E} \left[\sum_{s_j=0,1} \langle s_j | \tilde{\Lambda}_j(U_j^{(r)} | 0 \rangle \langle 0 | U_j^{(r)\dagger}) | s_j \rangle P(s_j | U_j^{(r)}) \right], \quad (\text{E3})$$

where we have defined the ‘marginal channel’ $\tilde{\Lambda}_j(\rho_j) = \text{Tr}_{k \neq j} (\Lambda(\mathbb{1}/2 \otimes \dots \otimes \mathbb{1}/2 \otimes \rho_j \otimes \mathbb{1}/2 \otimes \dots \otimes \mathbb{1}/2))$. Note that if $\Lambda = \bigotimes_{j=1}^N \Lambda_j$, we obtain $\tilde{\Lambda}_j = \Lambda_j$.

Employing the same reasoning as in Eq. (C15), we can average over the unitaries and use known results about two-copy twirling channels to find an expression for \tilde{C}_j :

$$\tilde{C}_j = \sum_{s_j=0,1} \langle 0 |^{\otimes 2} \Phi_j^{(2)} \left(\tilde{\Lambda}_j^*(|s_j\rangle\langle s_j|) \otimes |s_j\rangle\langle s_j| \right) | 0 \rangle^{\otimes 2} = \frac{1}{6} \sum_{s_j=0,1} (\langle s_j | \tilde{\Lambda}_j^*(|s_j\rangle\langle s_j|) | s_j \rangle + \text{Tr}[\Lambda_j^*(|s_j\rangle\langle s_j|)]) = \frac{1 + \tilde{G}_j}{3}. \quad (\text{E4})$$

Here $\tilde{G}_j = \frac{1}{2} \sum_{s_j} \langle s_j | \tilde{\Lambda}_j(|s_j\rangle\langle s_j|) | s_j \rangle$ contains the information of the single qubit noise in term of a marginal channel, i.e. without the assumption of locality of the noise, and coincides with the one in Eq. (C9) in the case $\Lambda = \bigotimes_{j=1}^N \Lambda_j$.

Let us proceed in a similar way for each pair of qubits (j, j') of an N -qubit system in order to derive a quantity

that also contains information about cross-talk errors. In analogy with Eqs. (E1) and (E3), for two qubits we define

$$\begin{aligned}
\tilde{D}_{j,j'} &= \mathbb{E} \left[\sum_{s_j, s_{j'}} \text{Tr} \left[\langle s_j, s_{j'} | \Lambda(U^{(r)} | \mathbf{0} \rangle \langle \mathbf{0} | U^{(r)\dagger}) | s_j, s_{j'} \rangle \right] P(s_j | U_j^{(r)}) P(s_{j'} | U_{j'}^{(r)}) \right] \\
&= \mathbb{E} \left[\sum_{s_j, s_{j'}} \langle s_j, s_{j'} | \tilde{\Lambda}_{j,j'} \left(U_j^{(r)} \otimes U_{j'}^{(r)} | 00 \rangle \langle 00 | U_j^{(r)\dagger} \otimes U_{j'}^{(r)\dagger} \right) | s_j, s_{j'} \rangle P(s_j | U_j^{(r)}) P(s_{j'} | U_{j'}^{(r)}) \right] \\
&= \mathbb{E} \left[\sum_{s_j, s_{j'}} \langle 0 |^{\otimes 4} U_j^{(r)\dagger \otimes 2} U_{j'}^{(r)\dagger \otimes 2} \left(\tilde{\Lambda}_{j,j'}^* (|s_j, s_{j'} \rangle \langle s_j, s_{j'} |) \otimes |s_j, s_{j'} \rangle \langle s_j, s_{j'} | \right) U_j^{(r)\otimes 2} U_{j'}^{(r)\otimes 2} | 0 \rangle^{\otimes 4} \right], \quad (\text{E5})
\end{aligned}$$

where we have made use of the definition of the ‘marginal channel’ $\tilde{\Lambda}_{j,j'}(\rho_j \otimes \rho_{j'})$ defined as

$$\tilde{\Lambda}_{j,j'}(\rho_j \otimes \rho_{j'}) = \text{Tr}_{k \neq j, j'} (\Lambda(\mathbb{1}/2 \otimes \dots \otimes \mathbb{1}/2 \otimes \rho_j \otimes \mathbb{1}/2 \dots \mathbb{1}/2 \otimes \rho_{j'} \otimes \mathbb{1}/2 \dots \otimes \mathbb{1}/2)) \quad (\text{E6})$$

that exploits the same reasoning as Eq. (E2). This quantity can be estimated from the calibration data \tilde{C}_j by extending the estimators in Eqs (C13)-(10) to two-qubit measurements.

As previously done for the single qubit quantity \tilde{C}_j we can now perform explicitly the average over the unitaries on the pair of qubits (j, j') exploiting the appropriate twirling channel identities. In particular, we can write

$$\tilde{D}_{j,j'} = \sum_{s_j, s_{j'}=0,1} \langle 0 |^{\otimes 4} \Phi_{j,j'}^{(2)} \left(\tilde{\Lambda}_{j,j'}^* (|s_j, s_{j'} \rangle \langle s_j, s_{j'} |) \otimes |s_j, s_{j'} \rangle \langle s_j, s_{j'} | \right) | 0 \rangle^{\otimes 4} = \langle 0 |^{\otimes 4} \Phi_{j,j'}^{(2)} (Q_{j,j'}) | 0 \rangle^{\otimes 4}, \quad (\text{E7})$$

where we have defined $Q_{j,j'} = \sum_{s_j, s_{j'}} \tilde{\Lambda}_{j,j'}^* (|s_j, s_{j'} \rangle \langle s_j, s_{j'} |) \otimes |s_j, s_{j'} \rangle \langle s_j, s_{j'} |$. Here we also introduced $\Phi_{j,j'}^{(2)}$, such that $\Phi_{j,j'}^{(2)}(Q_j \otimes Q_{j'}) = \mathbb{E}[U_j^{(r)\dagger \otimes 2} U_{j'}^{(r)\dagger \otimes 2} (Q_j \otimes Q_{j'}) U_j^{(r)\otimes 2} U_{j'}^{(r)\otimes 2}] = \Phi_j^{(2)}(Q_j) \otimes \Phi_{j'}^{(2)}(Q_{j'})$, that can be extended linearly to non-product observables $Q_{j,j'}$. Using the twirling formula in Eq. (C5) and working out the analytics, one obtains (with implicit identity operators)

$$\tilde{D}_{j,j'} = \langle 0 |^{\otimes 4} \Phi_{j,j'}^{(2)} (Q_{j,j'}) | 0 \rangle^{\otimes 4} = \frac{1}{36} [\text{Tr}(Q_{j,j'}) + \text{Tr}(\mathbb{S}_j Q_{j,j'}) + \text{Tr}(\mathbb{S}_{j'} Q_{j,j'}) + \text{Tr}(\mathbb{S}_j \mathbb{S}_{j'} Q_{j,j'})]. \quad (\text{E8})$$

We can then compute

$$\begin{aligned}
\text{Tr}(Q_{j,j'}) &= 4, \\
\text{Tr}(\mathbb{S}_j Q_{j,j'}) &= \sum_{s_j, s_{j'}} \langle s_j | \text{Tr}_{j'} [\tilde{\Lambda}_{j,j'}^* (|s_j, s_{j'} \rangle \langle s_j, s_{j'} |)] | s_j \rangle = 2 \sum_{s_j} \langle s_j | \tilde{\Lambda}_j (|s_j \rangle \langle s_j |) | s_j \rangle = 4\tilde{G}_j, \\
\text{Tr}(\mathbb{S}_{j'} Q_{j,j'}) &= 4\tilde{G}_{j'}, \\
\text{Tr}(\mathbb{S}_j \mathbb{S}_{j'} Q_{j,j'}) &= \sum_{s_j, s_{j'}} \langle s_j s_{j'} | \tilde{\Lambda}_{j,j'} (|s_j, s_{j'} \rangle \langle s_j, s_{j'} |) | s_j s_{j'} \rangle \equiv 4\tilde{G}_{j,j'},
\end{aligned} \quad (\text{E9})$$

where

$$\tilde{G}_{j,j'} = \frac{1}{4} \sum_{s_j, s_{j'}} \langle s_j, s_{j'} | \tilde{\Lambda}_{j,j'} (|s_j, s_{j'} \rangle \langle s_j, s_{j'} |) | s_j, s_{j'} \rangle. \quad (\text{E10})$$

Eventually, we arrive at the following expression for $\tilde{D}_{j,j'}$

$$\tilde{D}_{j,j'} = \frac{1}{9} \left(1 + \tilde{G}_j + \tilde{G}_{j'} + \tilde{G}_{j,j'} \right). \quad (\text{E11})$$

Estimating \tilde{C}_j , $\tilde{D}_{j,j'}$, we have thus experimental access to the terms $\tilde{G}_j, \tilde{G}_{j,j'}$ that contain information about the noise channel Λ . Both of them are equal to 1 in the absence of noise ($\tilde{G}_j = \tilde{G}_{j,j'} = 1$) and if $\tilde{G}_{j,j'} \neq \tilde{G}_j \tilde{G}_{j'}$ then $\Lambda \neq \bigotimes_{j=1}^N \Lambda_j$, i.e. the error is not local. Thus, we introduce the following quantity

$$\tilde{R} = \tilde{G}_{j,j'} - \tilde{G}_j \tilde{G}_{j'} \quad (\text{E12})$$

as a proxy of cross-talk effects. In particular, $\tilde{R} \neq 0$ witnesses the presence of cross-talk in the system according to the previous reasoning, namely $\tilde{R} \neq 0$ implies that Λ is not factorized. Let us remark here that $\tilde{R} = 0$ cannot exclude the presence of cross-talk. In fact, there exist noise channels Λ that introduce cross-talk effects but satisfy the condition $\tilde{R} = 0$. In the following, we provide an example of a noise channel that could model measurement errors in simple cases and show that \tilde{R} is able to detect cross-talk noise contributions in this case. Furthermore, employing this noise model, we observe that such contributions are negligible compared to the local ones.

2. Application to a two-qubit readout error model

Since the analysis presented in Fig. 6 suggests that the error in the platform is mostly due to readout, in this section we focus on a simple readout error model. We consider the following noise channel $\Lambda_{j,j'} = \Lambda_{j,j'}^{(2)} \circ (\Lambda_j^{(1)} \otimes \Lambda_{j'}^{(1)})$ for two qubits (j, j') where:

$$\begin{aligned}\Lambda_{j,j'}^{(2)}(\rho_{j,j'}) &= (1 - p_{\text{NL}})\rho_{j,j'} + p_{\text{NL}}\sigma_j^x\sigma_{j'}^x\rho_{j,j'}\sigma_j^x\sigma_{j'}^x; \\ \Lambda_j^{(1)}(\rho_j) &= (1 - p_L^{(j)})\rho_j + p_L^{(j)}\sigma_j^x\rho_j\sigma_j^x.\end{aligned}\tag{E13}$$

This model contains cross-talk errors with probability p_{NL} – namely, correlated bit-flips (which could model measurement errors) for qubits j and j' – and single qubit bit-flips with probability $p_L^{(k)}$, which could a priori be different for each qubit $k = j, j'$. In the low noise limit $p_{\text{NL}}, p_L^{(k)} \ll 1$, at first order, one can write the noise channel $\Lambda_{j,j'}$ as

$$\Lambda_{j,j'}(\rho_{j,j'}) \simeq (1 - p_{\text{NL}} - p_L^{(j)} - p_L^{(j')})\rho_{j,j'} + p_{\text{NL}}\sigma_j^x\sigma_{j'}^x\rho_{j,j'}\sigma_j^x\sigma_{j'}^x + p_L^{(j)}\sigma_j^x\rho_{j,j'}\sigma_j^x + p_L^{(j')}\sigma_{j'}^x\rho_{j,j'}\sigma_{j'}^x.\tag{E14}$$

Employing the definitions of Sec. E 1 one gets

$$\tilde{G}_{j,j'} \simeq 1 - p_{\text{NL}} - p_L^{(j)} - p_L^{(j')},\tag{E15}$$

$$\tilde{G}_j \simeq 1 - p_{\text{NL}} - p_L^{(j)},\tag{E16}$$

$$\tilde{G}_{j'} \simeq 1 - p_{\text{NL}} - p_L^{(j')},\tag{E17}$$

and that gives

$$\tilde{R} \simeq p_{\text{NL}}.\tag{E18}$$

For any small values of p_L , \tilde{R} is uniquely related to the cross-talk probability p_{NL} . Furthermore, $\tilde{R} \neq 0$ when the nonlocal term p_{NL} is different from zero and can be used to detect the cross-talk noise according to noise model employed. Hence, in the following we will adopt this noise model to investigate the strength of the cross-talk error in the quantum platform we have used in this work.

3. Experimental investigation on the employed platform

Here, we study the locality of the noise on the platform ‘ibm_prague’ that we have used to prepare the states of interest and measure the QFI. We employ the quantity \tilde{R} and estimate it using the calibration data for the 13-qubit state collected according to the indications in Sec. II A. In Fig. 7 we show \tilde{R} for neighbouring qubits, averaged over the iterations $i = 1, \dots, N_I$, namely $\tilde{R} = \frac{1}{N_I} \sum_{i=1}^{N_I} \tilde{R}^{(i)}$. The error bars are estimated as the standard deviation of the mean of the different estimates. A value not compatible with 0 (horizontal grey line) witnesses the presence of cross-talk, namely $\tilde{R} \neq 0 \Rightarrow \Lambda \neq \bigotimes_{i=1}^N \Lambda_i$. We observe that it is the case for the pairs of qubits (4, 5), (10, 11), (11, 12), (12, 13).

To estimate the strength of the cross-talk with respect to the local noise in the system we employ the noise model introduced in the previous section, Eq. (E13) to compute $p_L^{(k)}$ and p_{NL} from the measured values of $\tilde{G}_{j,j'}$ and \tilde{G}_j . At first order in $p_{\text{NL}}, p_L^{(k)}$ – in the limit $p_{\text{NL}}, p_L^{(k)} \ll 1$ – one obtains

$$p_L^{(j)} \simeq \tilde{G}_{j'} - \tilde{G}_{j,j'},\tag{E19}$$

$$p_L^{(j')} \simeq \tilde{G}_j - \tilde{G}_{j,j'},\tag{E20}$$

$$p_{\text{NL}} \simeq \tilde{R}.\tag{E21}$$

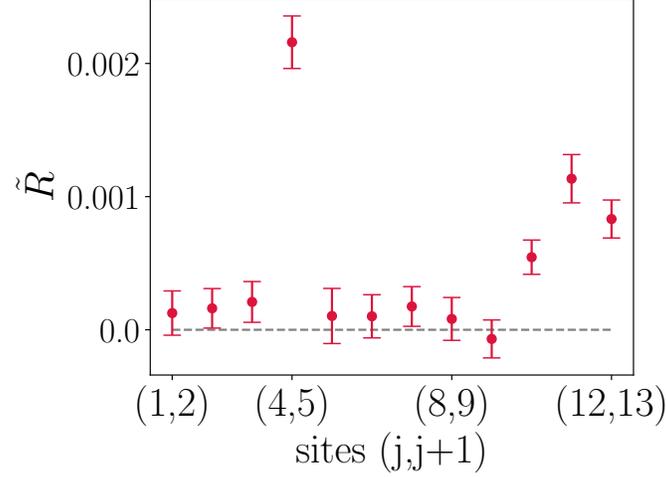


Figure 7: We estimate \tilde{R} for neighbouring qubits (j, j') as $\hat{R} = \frac{1}{N_I} \sum_{i=1}^{N_I} \hat{R}^{(i)}$, where $\hat{R}^{(i)} = \hat{\mathcal{G}}_{j,j'}^{(i)} - \hat{\mathcal{G}}_j^{(i)} \hat{\mathcal{G}}_{j'}^{(i)}$. The error bars are estimated as the standard deviation of the mean. A value not compatible within errorbars with 0 (horizontal grey line) means that cross-talk between the qubits is present. The measurement budget is the one described in Sec. III A for the calibration experiment.

pair (j,j')	\tilde{G}_j	$\tilde{G}_{j'}$	$\tilde{G}_{j,j'}$	$p_L^{(j)}$	$p_L^{(j')}$	$p_{NL} \simeq \tilde{R}$
(1, 2)	0.9775(3)	0.9783(2)	0.9565(4)	0.0218(5)	0.0210(3)	0.0001(1)
(2, 3)	0.9783(2)	0.9873(1)	0.9661(3)	0.0212(4)	0.0122(3)	0.0002(1)
(3, 4)	0.9873(1)	0.9756(2)	0.9634(3)	0.0121(4)	0.0238(4)	0.0002(1)
(4, 5)	0.9756(2)	0.9661(4)	0.9447(5)	0.0214(5)	0.0308(6)	0.0022(1)
(5, 6)	0.9661(4)	0.9847(10)	0.9515(10)	0.0332(11)	0.0146(14)	0.0001(1)
(6, 7)	0.9847(10)	0.9754(2)	0.9606(10)	0.0148(14)	0.0240(10)	0.0001(1)
(7, 8)	0.9754(2)	0.9843(2)	0.9604(2)	0.0239(3)	0.0150(3)	0.0002(1)
(8, 9)	0.9843(2)	0.9815(2)	0.9662(3)	0.0152(3)	0.0181(3)	0.0002(1)
(9, 10)	0.9815(2)	0.9843(2)	0.9669(3)	0.0182(4)	0.0154(3)	-0.0001(1)
(10, 11)	0.9844(2)	0.9863(2)	0.9713(3)	0.0149(4)	0.0129(5)	0.0005(1)
(11, 12)	0.9863(2)	0.9727(3)	0.9595(3)	0.0121(4)	0.0267(5)	0.0030(1)
(12, 13)	0.9717(3)	0.9892(1)	0.9621(3)	0.0271(4)	0.0096(4)	0.0008(1)

Table I: Table containing the experimentally measured values of \tilde{G}_j , $\tilde{G}_{j'}$, $\tilde{G}_{j,j'}$ and p_{NL} , $p_L^{(k)}$ ($k = j, j'$) calculated according to Eqs. (E15)–(E18) and using the estimators $\hat{\mathcal{G}}_j$, $\hat{\mathcal{G}}_{j,j'}$. The number in parentheses is the numerical value of the statistical error referred to the corresponding last digits of the result.

by inverting Eqs. (E15)–(E18). Plugging in these equations the experimental values of $\tilde{G}_{j,j'}$ and \tilde{G}_j it is possible to compute the probability ratio $p_{NL}/p_L^{(k)}$ that is informative of the relative strength of nonlocal noise. Such as for \tilde{R} , the measured values of $\tilde{G}_{j,j'}$ and \tilde{G}_j are an average over the estimates of the different iterations $i = 1, \dots, N_I$ and their error bars are calculated as the standard deviation of the mean. We employ the estimators $\hat{\mathcal{G}}_j$ and $\hat{\mathcal{G}}_{j,j'}$ discussed in the main text and Sec. E10.

We give the experimental results for any pair of neighbouring qubits in Tab. I. We observe that in the illustrative case of qubits (4, 5) – where $\tilde{R} = 0.002$ signals the presence of nonlocal noise – we obtain $p_{NL}/p_L^{(k)} \simeq 10^{-1}$ for both $k = j, j'$. More in general, for pairs (4, 5), (11, 12) and (12, 13), p_{NL} is not compatible with zero within errors. However, given that $p_{NL}/p_L^{(k)} \lesssim 10^{-1}$ we can conclude that the amount of cross-talk error in our platform would not harm the robust shadow protocol that we employ, as investigated numerically in Ref. [17]. The dominant source of

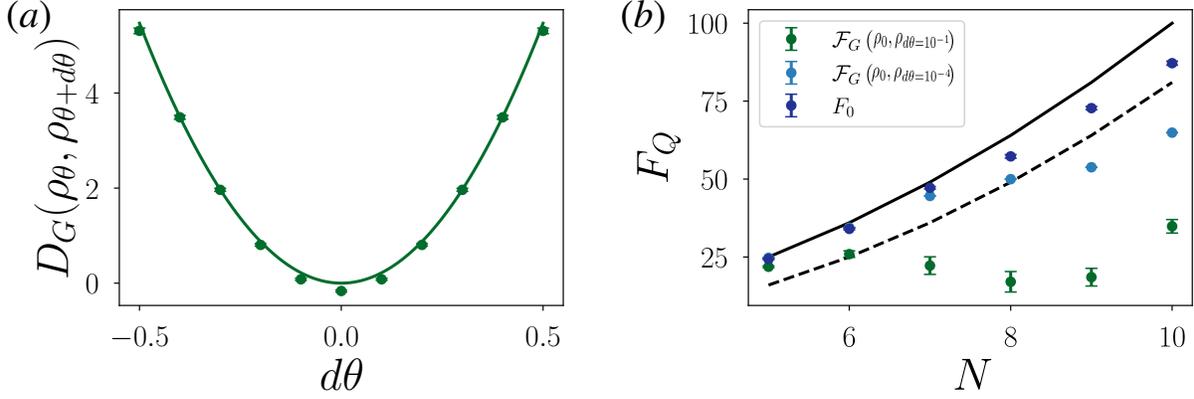


Figure 8: (a) Modified Bures distance D_G (and quadratic fit) for a GHZ state comprised of $N = 5$ qubits. (b) Comparison between \mathcal{F}_G (for different values of $d\theta$) and the lower bound F_0 as a function of the number of qubits N . The solid black line corresponds to $F_Q = N^2$, the dashed black line to the threshold above which the state can be considered GME.

error, under the assumptions of our noise model, corresponds to local measurement errors, which can be corrected faithfully via local robust shadows.

Appendix F: Further experimental results

In this appendix we study the estimator in Ref. [56] in comparison with our bounds F_n . Then, we provide more experimental results on the QAOA protocol and a comparison of two different methods of calibration, namely calibrating the circuit only once at the beginning of the experiment or repeating the calibration in each iteration to follow the time fluctuations.

1. Comparison with previous work in Ref. [56]

In this section, we post-process our recorded robust RM measurement data to obtain another lower bound to the QFI that has been previously estimated using standard RM formalism. The lower bound of interest is defined in Refs. [55, 56] as a function of the quantum states ρ_θ and $\rho_{\theta+d\theta}$, as

$$\mathcal{F}_G(\rho_{d\theta}) \equiv \frac{D_G(\rho_\theta, \rho_{\theta+d\theta})}{d\theta^2} = \frac{8 \left[1 - \text{Tr}(\rho_\theta \rho_{\theta+d\theta}) + \sqrt{(1 - \text{Tr}(\rho_\theta^2))(1 - \text{Tr}(\rho_{\theta+d\theta}^2))} \right]}{d\theta^2}. \quad (\text{F1})$$

It is important to note for the above lower bound that $\lim_{d\theta \rightarrow 0} \mathcal{F}_G(\rho_{d\theta}) = F_0$ [55] and estimating it requires one to be able to distinguish between a state ρ_θ and its neighbor $\rho_{\theta+d\theta}$ that encode an unknown parameter θ .

In our analysis, we estimate unbiased estimators for each of the terms in RHS of Eq. (F1) according to U-statistics, as detailed in the main text (Sec. II E). The θ parametrized state is defined as $\rho_\theta = e^{-i\theta A} \rho e^{i\theta A}$ with $A = \frac{1}{2} \sum_j \sigma_z^{(j)}$. We consider here $\theta = 0$ for simplicity as done in Ref. [56]. Additionally, compared to Ref. [56] that prepared experimentally the two parametrized states ρ_θ and $\rho_{\theta+d\theta}$, with the robust classical shadow formalism, we can estimate $D_G(\rho_\theta, \rho_{\theta+d\theta})$ by performing this step classically during the post-processing stage. We remark that robust classical shadows defined in Eq. (3) satisfy $\rho_\theta = \mathbb{E}[e^{-i\theta A} \tilde{\rho}^{(r_i)} e^{i\theta A}]$ with the average taken over the applied random unitaries and measurements.

Fig. 8 summarizes the analysis for our experimental data. We consider the RM data taken after preparing a N -qubit GHZ states. Firstly, as shown in Fig. 8(a), we estimate the modified Bures distance $D_G(\rho_{\theta_1}, \rho_{\theta_2})$ as a function of $d\theta = \theta_2 - \theta_1$ for a 5-qubit GHZ state [56, Fig. 2(b) for 4-qubit GHZ state]. We then perform a polynomial fit as shown by the solid line in Fig. 8(a). The coefficient of the quadratic term of this fit provides an estimation of $\mathcal{F}_G(\rho_{d\theta})$. This is extracted for qubit sizes ranging from 5 to 10 in Fig. 8(b) for two different values of $d\theta$. We observe, that $\mathcal{F}_G(\rho_{d\theta=10^{-1}}) \leq \mathcal{F}_G(\rho_{d\theta=10^{-3}}) \leq F_0$ for all qubit sizes. It is important to note that in order to obtain values of \mathcal{F}_G that are comparable to F_0 , one has to perform the procedure employed in Ref. [56] by encoding very small parameter shifts that become extremely challenging experimentally.

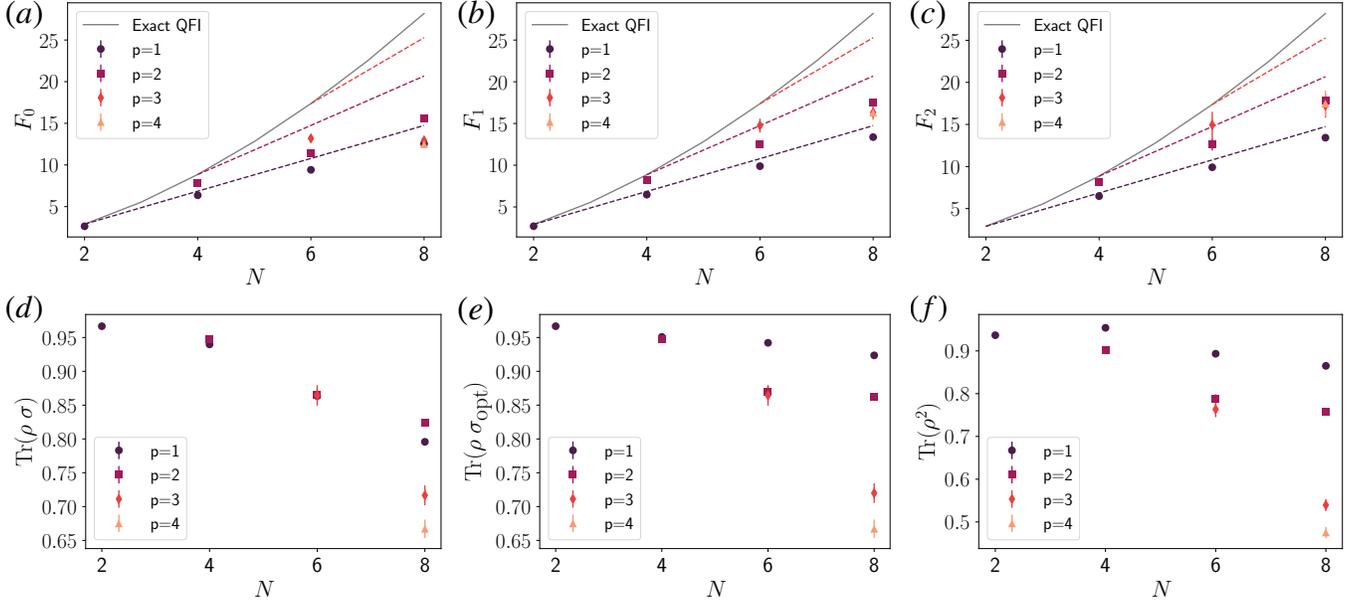


Figure 9: More experimental results on the TFIM at the critical point. In the upper panels we present results for (a) F_0 , (b) F_1 and (c) F_2 . In the lower panels we show (d) the fidelity with respect to the real ground state σ , (e) the fidelity with respect to the state σ_{opt} obtained by classical optimization and experimentally prepared through QAOA at fixed depth p , and (f) the purity $\text{Tr}(\rho^2)$ of the prepared state. The results for F_1 have already been presented in the main text. The measurement protocol details are described in Sec. II A. In all the plots, p is the number of layers in the circuit, i.e. the depth of the circuit. The solid black lines correspond to the exact value of the QFI F_Q . The colored dashed line denote the theoretical value of the QFI at fixed depth.

2. Ground state of the TFIM at the critical point

In this section we give additional experimental results concerning the ground state of the TFIM at $h = 1$, prepared with the variational circuit described in the main text. We have showed explicitly the value of the bound F_1 in the main text. In the upper plots of Fig. 9 we show F_0 , F_1 and F_2 . The lines in (a), (b) and (c) denote the theoretical QFI for a fixed circuit depth p . In the lower plots we show the fidelities of state preparation and purity in the lower ones. In particular, in (d) we show the fidelity with respect to the real ground state of the TFIM at the critical point, computed with exact diagonalization. We observe that increasing the depth p the fidelity generally drops, because the more layers the higher the noise in the system. The only exception is the case of depth $p = 2$ for $N = 8$, the rationale behind it is that the approximation of the ground state for $p = 1$ ad $N = 8$ is very poor, and thus, even in the presence of noise, for $p = 2$ the prepared state is closer to the true one. The worsening of state preparation with p is also evident in (e) where we plot the fidelity with respect to the state prepared in the case of a noiseless QAOA. The fidelity is always better for $p = 1$, but we remind the reader that the prepared state is not faithful to the ground state of the TFIM for larger system sizes. Finally in (f) we show the purity of the prepared states. Again we observe that it drops with increasing p due to the presence of noise. The final state should be a pure state in the ideal scenario, i.e. $\text{Tr}(\rho^2) = 1$. Here we observe that increasing the number of layers tends to decrease the purity of the prepared state, e.g. for $p = 4$ and $N = 8$ one has $\text{Tr}(\rho^2) \sim 0.5$.

3. Comparison of error mitigation protocols

In this section, we show experimental evidence that calibrating in iterations is more efficient than calibrating it once at the beginning of the experiment. In Fig. 10 we compare the estimation of the bound of the QFI when the calibration of the device is performed at the beginning of the whole experimental procedure or according to our prescriptions. We present the error mitigated experimental estimations of F_0 and F_1 (light to dark). In Fig. 10(a) the calibration is performed at the beginning. We observe that the robust estimation for larger system sizes is not compatible with the $\sim N^2$ scaling predicted by the theory, and that it does not violate the witness of $(N - 1)^2$ that validates GME. In Fig. 10(b) we present the same experimental results of the main text for F_0 and F_1 . As already commented, we

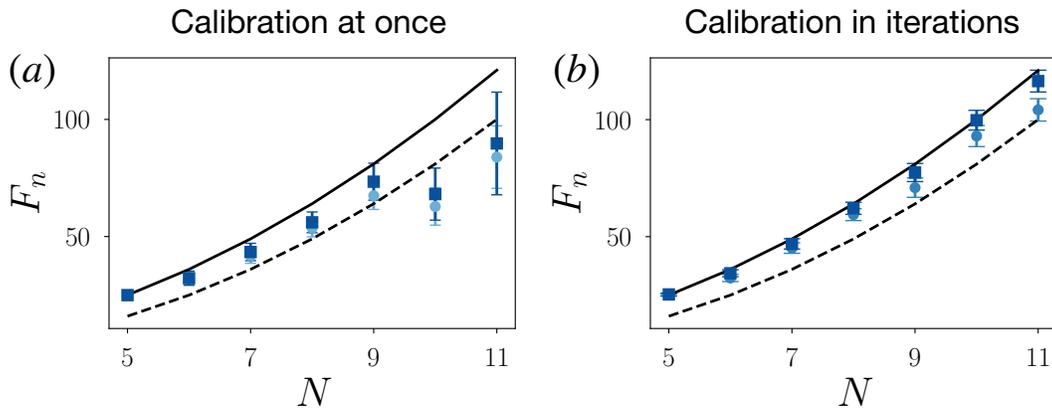


Figure 10: Comparison of different methods of calibration. F_0 and F_1 (light to dark) when (a) the calibration of the device is performed at the beginning of the whole experimental run, (b) the calibration is performed in each iteration according to the experimental protocol described in the main text. The measurement budget of both experiments is the same and is detailed in Sec. III A. The solid line is the exact value of the QFI ($F_Q = N^2$) for pure GHZ states. The dashed black line corresponds to the entanglement witness $\Gamma(N, k = N - 1) = (N - 1)^2$ above which the state can be claimed to be GME.

observe the $\sim N^2$ scaling of QFI and witness GME. The discrepancy is due to the fluctuating gate and readout errors in the quantum processors that affect the reliability of the results when the experimental run takes long times, that is for larger N .

Appendix G: Numerical investigations

In this section we study the behavior of the classical Fisher information in comparison with the quantum Fisher information. We investigate the scaling of \hat{G}_j as a function of the number of unitaries N_U for different values of the readout error and the scaling of the required number of measurements to achieve a given level of statistical errors on our highest measured bound F_2 . Lastly, we present classical numerical experiment for GHZ states prepared without any state preparation errors.

1. Classical and Quantum Fisher information comparison in noisy GHZ states

The Fisher information is a fundamental concept in statistics and information theory [35, 36]. It measures the amount of information that a random variable carries about an unknown parameter when sampled from a given probability distribution and plays a crucial role in the field of metrology. In particular, in the context of estimating an unknown parameter θ encoded in a quantum state ρ , it has been used to show that the precision of the measurement could go beyond the shot-noise limit [12].

The choice of the measurement setting significantly influences the accuracy of the estimation process. Optimal choices are characterized by measurement results exhibiting a statistical distribution that is highly sensitive to variations in θ . Indicating as $P(\mu|\theta)$ the probability of a measurement result μ given that the parameter has the value θ , the classical Fisher information can be written as:

$$F(\theta) = \sum_{\mu} \frac{1}{P(\mu|\theta)} \left(\frac{\partial P(\mu|\theta)}{\partial \theta} \right)^2. \quad (\text{G1})$$

An upper bound to the Fisher information is obtained by maximizing the previous equation over all possible generalized measurements settings [1] and corresponds to the quantum Fisher information (QFI), defined in Eq. (1) of the main text.

Therefore, by definition, the classical Fisher information (CFI) is upper-bounded by the QFI. However, it has been readily employed in quantum experiments to prove the presence of multipartite entanglement or enhanced metrological sensitivity with respect to the classical cases [37–39]. The latter is due to the fact that it is easier to measure in general, since it does not need the full spectral resolution of the density matrix, even though one is faced with the challenge of finding good measurement observables and measurement basis. In this work we proposed a method to estimate

directly the QFI, that is intrinsically not affected by the choice of the measurement basis, employing a converging series of lower bounds to it. In this section, for the sake of completeness we draw a comparison between the CFI, for some fixed choice of measurement, and the QFI in noisy GHZ states.

The CFI, in this instance, can be calculated analytically. We write the GHZ state of a N -qubits system as $|\psi\rangle = (|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$. We perform the evolution under the operator $U = e^{-i\theta A}$ with $A = \frac{1}{2} \sum_i \sigma_j^z$ and we measure the qubits along the y -axis. We remark that the latter means applying a phase gate $S^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}^\dagger$ and a Hadamard gate $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ to the state, before measuring along the z -axis. The probability of obtaining as an outcome an N -bit string \mathbf{s} is given by

$$P(\mathbf{s}|\theta) = |\langle \mathbf{s} | (HS^\dagger)^{\otimes N} e^{-i\theta A} |\psi\rangle|^2. \quad (\text{G2})$$

Performing the calculation one obtains

$$(HS^\dagger)^{\otimes N} e^{-i\theta A} |\psi\rangle = \frac{1}{\sqrt{2}} (HS^\dagger)^{\otimes N} \left(e^{-i\theta N/2} |0\rangle^{\otimes N} + e^{i\theta N/2} |1\rangle^{\otimes N} \right) = \frac{1}{\sqrt{2^{N+1}}} \sum_{\mathbf{s}} \left(e^{-i\theta N/2} + e^{i\theta N/2} e^{i\pi N/2} e^{i\pi |\mathbf{s}|} \right) |\mathbf{s}\rangle \quad (\text{G3})$$

leading to

$$P(\mathbf{s}|\theta) = \frac{1}{2^N} \left[1 + (-1)^{|\mathbf{s}|} \left(\cos(\theta N) \cos\left(\frac{\pi}{2}N\right) + \sin(\theta N) \sin\left(\frac{\pi}{2}N\right) \right) \right]. \quad (\text{G4})$$

We observe that in $\theta = \frac{\pi}{2N} \equiv \theta_0$ and assuming that N is even one gets

$$P(\mathbf{s}|\theta_0) = \frac{1}{2^N}, \quad \left. \frac{\partial P(\mathbf{s}|\theta)}{\partial \theta} \right|_{\theta=\theta_0} = -\frac{N}{2^N} (-1)^{|\mathbf{s}|+N/2}. \quad (\text{G5})$$

Eventually, the CFI reads

$$F(\theta_0) = \sum_{\mathbf{s}} \frac{1}{P(\mathbf{s}|\theta_0)} \left(\left. \frac{\partial P(\mathbf{s}|\theta)}{\partial \theta} \right|_{\theta=\theta_0} \right)^2 = N^2. \quad (\text{G6})$$

Therefore, employing this measurement scheme, the CFI reaches the maximum possible value for a system comprised of N qubits and coincides with the QFI. We will show here that in the presence of global depolarization, the CFI (obtained with the same, fixed, measurement as considered above) decreases faster than the QFI, whose explicit functional dependence with respect to the strength of the noise has been given in Eq. (20) and discussed in Sec. B. In the presence of global depolarization of strength p_D , then $P(\mathbf{s}|\theta)$ calculated above is changed to $(1-p_D)P(\mathbf{s}|\theta) + p_D \frac{1}{2^N}$, so that now one has

$$P(\mathbf{s}|\theta_0) = \frac{1}{2^N}, \quad \left. \frac{\partial P(\mathbf{s}|\theta)}{\partial \theta} \right|_{\theta=\theta_0} = -(-1)^{|\mathbf{s}|+N/2} (1-p_D) \frac{N}{2^N} \quad (\text{G7})$$

which leads to the following value for the CFI

$$F(\theta_0) = \frac{1}{2^N} \sum_{\mathbf{s}} (1-p_D)^2 N^2 = (1-p_D)^2 N^2 \quad (\text{G8})$$

With respect to Eq. (20) we observe that the CFI decays faster (as $\sim (1-p_D)^2$ instead of $\sim (1-p_D)$) in the limit of large system sizes N . Then, in the presence of global depolarizing noise in the system, we can argue that the CFI is always strictly a lower bound to the true value of the quantum Fisher information with the fixed measurement setting considered here.

In Fig. 11 we plot numerical results for comparing the QFI and the CFI in the presence of depolarization errors. We consider two cases, in Fig. 11(a) we calculate those quantities on the state defined as

$$\rho = (1-p_D) |\psi\rangle \langle \psi| + p_D \frac{\mathbb{1}}{d}, \quad (\text{G9})$$

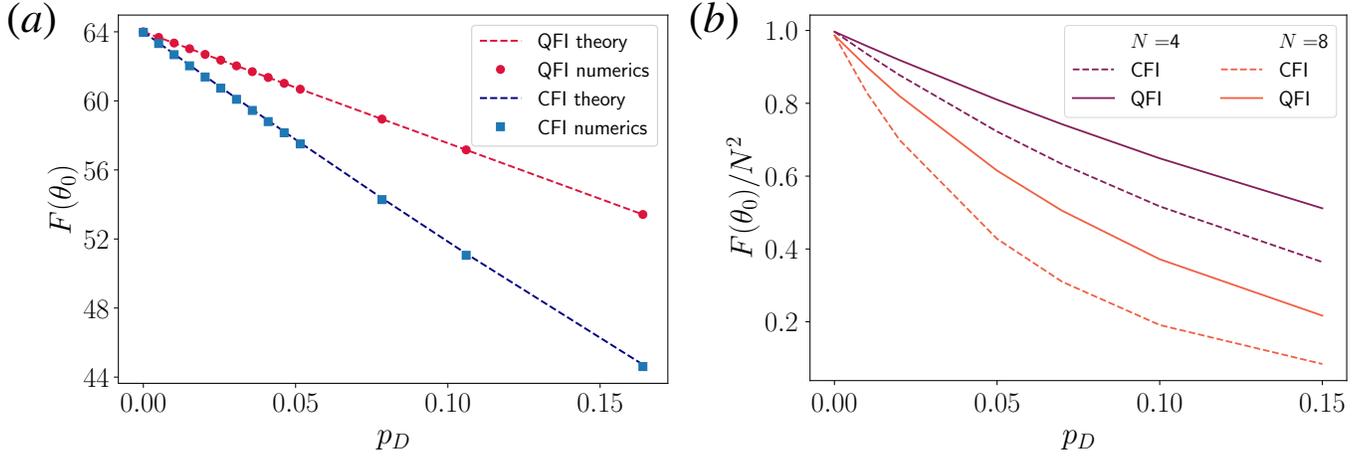


Figure 11: (a) QFI and CFI of the GHZ state for $N = 8$ qubits, in the presence of global depolarization of strength p_D . The points represent the numerical results while the dashed lines correspond to Eq. (B1) and Eq. (G8). (b) QFI and CFI of the GHZ state simulated using 'qasm_simulator'. Noise, in the form of a depolarizing channel, is added to each CNOT gate necessary to prepare the GHZ state. Its strength is labeled by p_D .

where d is the Hilbert space dimension ($d = 2^N$). For the state ρ we have analytical predictions (dashed lines according to Eq. (20) and Eq. (G8)) that can be compared with numerical results according to the Hellinger method for estimating the CFI [39]. We observe that the CFI decreases faster than QFI, enforcing our statement that it is important to find a direct and reliable estimator of the QFI. In Fig. 11(b) we prepare the GHZ state on the simulator 'qasm_simulator' [33], by means of a Hadamard gate and $N-1$ CNOT gates. On each CNOT we add depolarizing noise whose strength p_D we can tune. We observe that the CFI and QFI both decrease, as expected when increasing p_D and the number of qubits in the system. Also, as in the previous case, the CFI is always strictly smaller than the QFI for any value of $p_D > 0$.

2. \mathcal{G}_j as a function of the number of unitaries and readout error

In this section, we study numerically the estimator $\hat{\mathcal{G}}_j$ in Eq. (11). We employ the IBM quantum simulator for providing an estimate of the scaling of $\hat{\mathcal{G}}_j$ as a function of the number of unitaries N_U in the randomized measurement protocol in the calibration step. We induce noise in the circuit as a readout error p_{meas} according to the noise model employed in Appendix D 2. In Fig. 12 we plot $\hat{\mathcal{G}}_j$ for different values of p_{meas} , as a function of N_U . The estimation is compatible with the theoretical values (dashed lines) within error bars, for any value of p_{meas} . We observe the error bars on the estimation decreases with N_U , for fixed $N_M = 1000$. For the value of N_U used in our experimental protocol ($N_U \sim 200$) we observe an uncertainty of $\sim 1\%$ on the estimation of $\hat{\mathcal{G}}_j$. Increasing the number of unitaries used does not improve the estimation significantly. Hence, we choose $N_U = 200$.

3. Scaling of the measurement budget for the lower bound F_2

In Fig. 13, we provide numerical simulations to extract the scalings of the statistical errors on our highest measured lower bounds F_2 . We consider an N -qubit pure GHZ state and consider once again the Hermitian operator $A = \frac{1}{2} \sum_{j=1}^N \sigma_j^z$. We simulate the protocol by applying N_U local random unitaries $U^{(r)}$ with $r = 1, \dots, N_U$ with $N_M = 1000$ projective computational basis measurements per unitary to obtain batch estimates \hat{F}_2 using $N_B = 10$ batches. The estimation is realized using Eq. (15) of the main text, however we do not consider here common randomized measurement, i.e. we take $\sigma = 0$ in Eq. (6). The average statistical error \mathcal{E} is calculated by averaging the relative error $\mathcal{E} = |\hat{F}_2 - F_2|/F_2$ over 100 numerically simulated experimental runs for different values of N_U . We find the maximum value of N_U for which we obtain $\mathcal{E} \leq 0.1$ for different system sizes N by using a linear interpolation function.

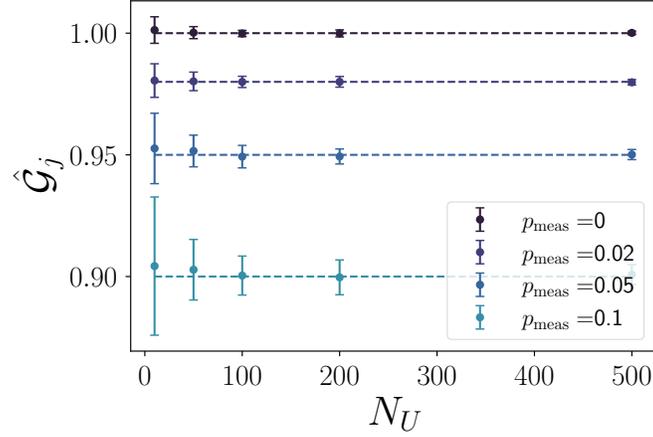


Figure 12: Numerical simulation of the calibration protocol on the IBMQ quantum simulator for a two qubit system. We plot \hat{G}_j as a function of N_U , for $N_M = 1000$ and varying the readout error p_{meas} .

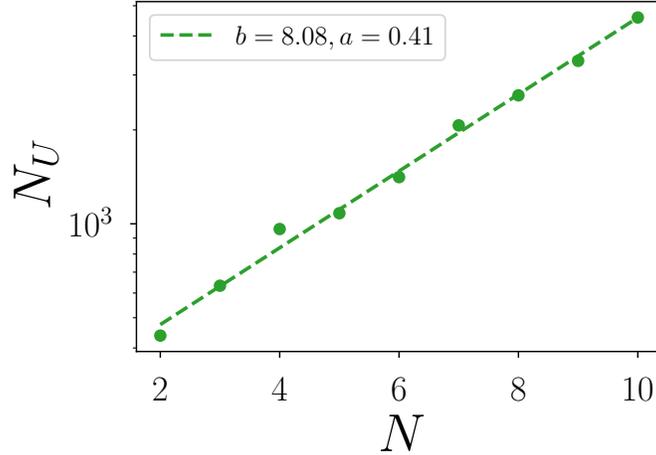


Figure 13: Numerical simulations to provide the number unitaries N_U required to estimate F_2 below an error of $\mathcal{E} = 10\%$ for a GHZ state with respect to $A = \frac{1}{2} \sum_{j=1}^N \sigma_j^z$. We simulate $N_M = 1000$ computational basis measurements per unitary. The dashed line is an exponential fit of the type 2^{b+aN} highlighting the scaling as a function of the system size N .

4. Numerical simulation of the experiment for perfect GHZ states and readout errors

In this last section, we provide the measurement of the lower bounds via a classical numerical experiment for GHZ states prepared without any state preparation errors (perfect GHZ states). We take the same measurement budget as applied in the experimental procedure (c.f Sec. III A of the main text). Here again, we consider $\sigma = 0$ in Eq. (6). Additionally, we consider that the single qubit random unitary operations are done perfectly and take into account only readout errors with a probability of $p_{\text{meas}} = 1.4\%$ as recorded for the IBM superconducting qubit device ‘ibm_prague’ [33]. The results are shown in Fig. 14.

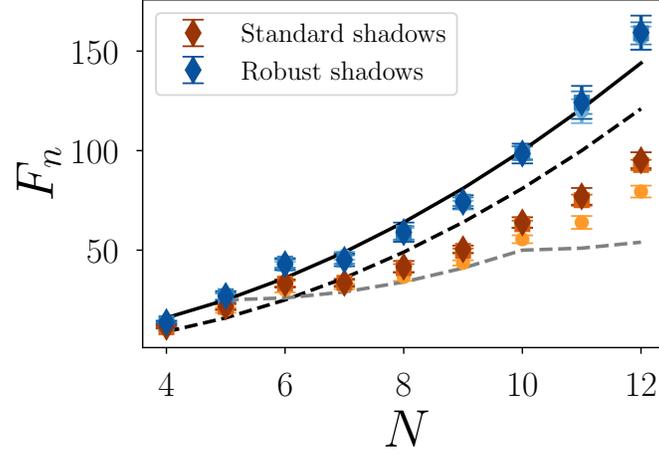


Figure 14: Numerical simulation of the experimental procedure, for perfect GHZ states but including readout errors. As also mentioned in the main text, this figure shows F_0, F_1, F_2 (light to dark with circle, square and diamond respectively) as a function of the number of qubits N , where we fix as always the operator $A = \frac{1}{2} \sum_{j=1}^N \sigma_j^z$. The solid line is the exact value of the QFI $F_Q = N^2$ for pure GHZ states. The dashed black line corresponds to the entanglement witness $\Gamma(N, k = N - 1) = (N - 1)^2$. The dashed grey line corresponds to the entanglement witness $\Gamma(N, k = 5)$.