Enhanced measurements on quantum computers via the simultaneous probing of non-commuting Pauli operators

Rick P. A. Simon[1], Zheng Shi[2], Charlie Nation[1], Andrew Jena[2,3], and Luca Dellantonio[1]

[1]Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, UK
[2]Institute for Quantum Computing, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[3]Department of Combinatorics & Optimization, University of Waterloo, Waterloo, ON N2L 3G1, Canada

arXiv:2509.01482v3 [quant-ph] 22 Sep 2026

Measuring the state of quantum computers is a highly non-trivial task, with implications for virtually all quantum algorithms. A promising avenue is multi-copy schemes, where identical copies of a quantum state are measured jointly so that all Pauli operators within the considered observable can be simultaneously assessed. Here, we present a first implementation of such a two-copy scheme in a measurement protocol. Based on Bayesian statistics, it accurately estimates not only the average of the desired observable but also the error en route. This enables an adaptive shot-allocation algorithm that preferentially samples the most uncertain Pauli terms. In regimes with many non-commuting Pauli operators, this "double" scheme can outperform the state-of-the-art measurement protocol in minimizing total shots for a given precision. We also numerically confirm the finding in previous theoretical works that the two-copy scheme incurs an overhead due to the square-root relationship between the variance of measured quantities and the number of measurement shots.

1 Introduction

An essential task for all quantum algorithms is to accurately estimate the expectation value of an observable [1–3], expressed in terms of Pauli strings (PS), within a given measurement budget M. Groups of commuting PS can be measured simultaneously to lower the required number of measurements. Nevertheless, to attain the desired estimation accuracy, in practice one often needs to repeatedly probe many such groups of commuting PS. This requires an efficient strategy to allocate measurement budgets to different groups [4–30].

Various such schemes have been developed. One class of methods groups PS into commuting subsets to measure them simultaneously [4–14]. Another approach, broadly called classical shadows [15–25], uses randomized measurements to build a classical representation of the state from which observables can be estimated. These techniques have greatly enhanced the estimation accuracies of observables measured on quantum computers. However, many successful approaches such as the variational ones [7, 31, 32] need better measurement protocols. These variational algorithms prepare a parameterized quantum state and measure the constituent Pauli strings of the observable to learn a cost function, which is minimised by classical optimizers that iteratively update the state parameters. Tackling larger and larger systems [33–37] restricts the number of available cost function evaluations, and thus jeopardizes the future of these algorithms unless efficient measurement protocols are employed [38].

Of the strategies proposed so far, multi-copy schemes [22–25] take advantage of the following observation. Two PS $\hat{P}_i$ and $\hat{P}_j$ composed of qubit operators (Pauli matrices, see Sec. 2) either commute or anti-commute with each other. Therefore, their "double" versions – the tensor product of each PS with itself, $\hat{P}_i \otimes \hat{P}_i$ and $\hat{P}_j \otimes \hat{P}_j$ – always commute and can therefore be measured simultaneously, the advantage of which is illustrated in Fig. 1(a). In this work, we implement the theo-

Rick P. A. Simon: r.p.a.simon@exeter.ac.uk
Zheng Shi: zheng.shi@uwaterloo.ca
Charlie Nation: c.nation2@exeter.ac.uk
Andrew Jena: ajjena@uwaterloo.ca
Luca Dellantonio: l.dellantonio@exeter.ac.uk

retical idea in Refs. [22–25] in a "double" scheme, which combines the multi-copy principle with a Bayesian approach to accurately estimate both the mean *and* the statistical error of an arbitrary multi-qubit observable.

A major step forward compared to Refs. [22–25] is the ability to monitor the estimated error, which allows our scheme to predict how to allocate each measurement shot (either a group of commuting PS $\hat{P}_i$ or all of their double $\hat{P}_i \otimes \hat{P}_i$) so that the resulting error is minimized. We show numerically that when many PS do not commute with each other (so grouping is hard), our double scheme outperforms state-of-the-art measurement protocols, particularly for a limited measurement budget M. A paradigmatic class of observables that are characterized by long-range interactions between qubits and therefore includes several non-commuting PS is molecular Hamiltonians. These are a prime target for both near- and long-term quantum simulations [35, 39–42].

2 Theory

The aim of this work is to build an accurate estimator $\widetilde{O} \approx \langle \hat{O} \rangle$, where $\hat{O} = \sum_{i=1}^{p} c_i \hat{P}_i$ is an observable defined on q qubits and expressed as a weighted sum of p PS $\hat{P}_i = \bigotimes_{j=1}^{q} \hat{\sigma}_j^{(i)}$. Here, c_i are real constants and the Pauli operator $\hat{\sigma}_j^{(i)} \in \{\hat{I}_j, \hat{X}_j, \hat{Y}_j, \hat{Z}_j\}$ acts on qubit j. Calling θ_i the probability of obtaining $+1$ when measuring $\hat{P}_i$ we have $\widetilde{O} = \sum_{i=1}^{p} c_i(2\widetilde{\theta}_i - 1)$. As explained in Sec. 3, we achieve high accuracy in $\widetilde{O}$ by continuously monitoring its estimated error $(\Delta\widetilde{O})^2$. Differently from elsewhere [4, 5, 43], where $(\Delta\widetilde{O})^2$ is calculated from the variances and covariances of the PS, in this work $(\Delta\widetilde{O})^2$ is the variance of the estimate of $\langle \hat{O} \rangle$. The quantities $\widetilde{O}$ and $(\Delta\widetilde{O})^2$, central in this work, are thus determined via

$$\widetilde{O} = \sum_{i=1}^{p} c_i \widetilde{P}_i = \sum_{i=1}^{p} c_i(2\widetilde{\theta}_i - 1) \tag{1a}$$

$$\begin{aligned}(\Delta\widetilde{O})^2 &= \widetilde{\langle \hat{O} \rangle^2} - \left(\widetilde{O}\right)^2 \\ &= 4\sum_i c_i^2(\widetilde{\theta_i^2} - \widetilde{\theta}_i^2) + 4\sum_{i \neq j} c_i c_j(\widetilde{\theta_i\theta_j} - \widetilde{\theta}_i\widetilde{\theta}_j),\end{aligned} \tag{1b}$$

where we highlight that $(\Delta\widetilde{O})^2$ in Eq. (1) approximates the "statistical" variance of the estimator $\widetilde{O}$, and *not* the "quantum mechanical" variance of the observable $\langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2$ [indeed Eq. (1b) calculates the difference between the estimation of the square of the average (first term in the right hand side) and the square of the estimated average (second term)]. A few key observations must be made here. Firstly, in M repeated measurements of $\hat{P}_i$ alone, in agreement with the central limit theorem [4, 44] we find in Sec. 2.1 that $(\Delta\widetilde{O})^2 \overset{M\mapsto\infty}{=} (\langle \hat{P}_i^2 \rangle - \langle \hat{P}_i \rangle^2)/M$. Furthermore, when two PS $\hat{P}_i$, $\hat{P}_j$ with $i \neq j$, have never been measured together $\widetilde{\theta_i\theta_j} - \widetilde{\theta}_i\widetilde{\theta}_j$ in Eq. (1) is zero, and only the scaled variances contribute to the error, as should be the case. In this scenario, the posterior that we use in Sec. 2.2 to calculate $\widetilde{\theta_i\theta_j}$ is the product of the posteriors used for $\widetilde{\theta}_i$ and $\widetilde{\theta}_j$ individually, implying $\widetilde{\theta_i\theta_j} = \widetilde{\theta}_i\widetilde{\theta}_j$.

In the following, we explain how $\widetilde{O}$ and $(\Delta\widetilde{O})^2$ are practically determined from the experimental measurement outcomes of both the PS $\hat{P}_i$ and their double $\hat{P}_i \otimes \hat{P}_i$. To do so, we build Bayesian probability distributions [4, 5, 45, 46] to estimate $\widetilde{\theta}_i$, $\widetilde{\theta_i^2}$ and $\widetilde{\theta_i\theta_j}$ and therefore the quantities in Eq. (1). As can be seen in Eqs. (3) and (5), our approach requires distinguishing the cases in which $\widetilde{P}_i$ and $\widetilde{P}_j$ are measured together, alone, or in their double version. The procedure for calculating $\widetilde{\theta}_i$ and $\widetilde{\theta_i^2}$, conceptually simpler, is given in Sec. 2.1. Sec. 2.2 explains how to find $\widetilde{\theta_i\theta_j}$ when $i \neq j$ and $\hat{P}_i$ commutes with $\hat{P}_j$.

2.1 Settings and estimation of $\widetilde{\theta}_i$ and $\widetilde{\theta_i^2}$

As depicted in Fig. 1(b), for an input state $|\varphi\rangle$ and an observable $\hat{O}$ to be measured, the double scheme requires preparing $|\psi\rangle = |\varphi\rangle \otimes |\varphi\rangle$ and simultaneously probing all $\hat{P}_i \otimes \hat{P}_i$ (where $\hat{P}_i$ is in $\hat{O}$). Calling θ_i and ϕ_i the probabilities of obtaining $+1$ when measuring $\hat{P}_i$ and $\hat{P}_i \otimes \hat{P}_i$, respectively, their relationship is

$$\left.\begin{aligned}\langle\varphi|\,\hat{P}_i\,|\varphi\rangle &= 2\theta_i - 1 \\ \underbrace{\langle\psi|\,\hat{P}_i \otimes \hat{P}_i\,|\psi\rangle}_{\langle\varphi|\hat{P}_i|\varphi\rangle^2} &= 2\phi_i - 1\end{aligned}\right\} \Rightarrow \phi_i = \theta_i^2 + (1-\theta_i)^2. \tag{2}$$

This equation confirms the intuitive fact that obtaining $+1$ $[-1]$ in the measurement of $\hat{P}_i \otimes \hat{P}_i$ is equivalent to obtaining $(+1,+1)$ or $(-1,-1)$ $[(+1,-1)$ or $(-1,+1)]$ in two independent measurements of $\hat{P}_i$.

As can be understood from Eq. (2) and Fig. 1(a), measuring the double PS $\hat{P}_i \otimes \hat{P}_i$ enhances the estimation's accuracy of the absolute values of $\widetilde{P}_i$ in Eq. (1), yet provides no

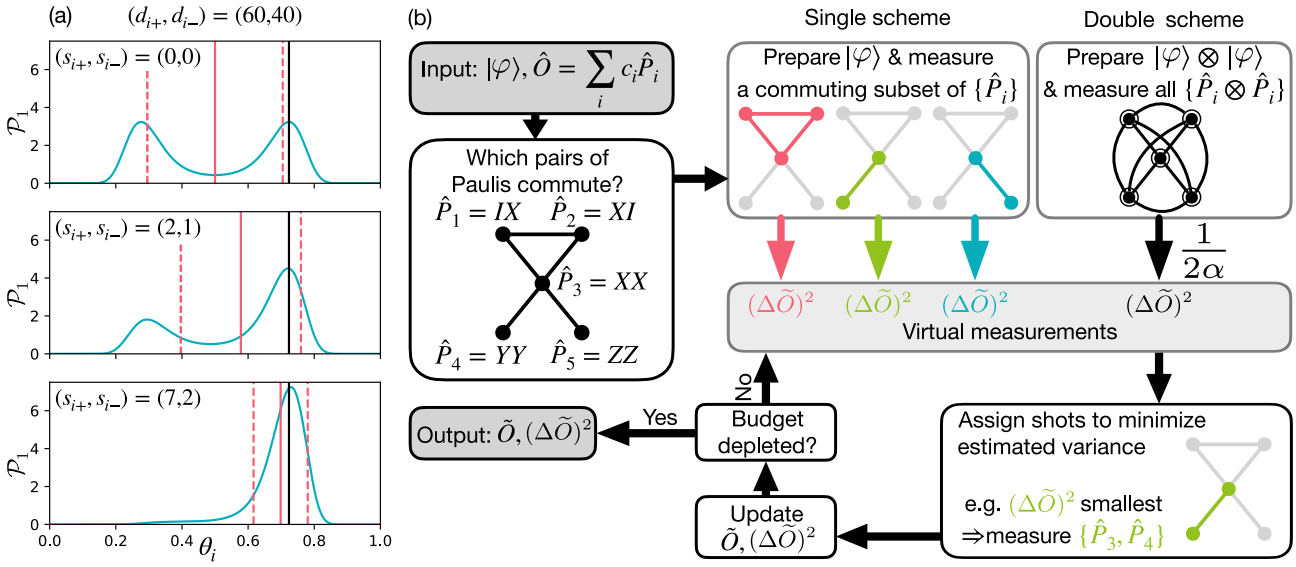


Figure 1: (a) The posterior $\mathcal{P}_1$ from Eq. (3b) (blue line). The sign uncertainty is evident from the two peaks, which rapidly become one when shots are allocated without the double scheme (increase of $s_{i\pm}$, top to bottom). The vertical black line is the value $\theta_i = 0.725$ [$\implies \phi_i = 0.60125$, see Eq. (2)] used to randomly generate the measurement outcomes $s_{i\pm}$ and $d_{i\pm}$. The pink vertical lines are the averages $\widetilde{\theta}_i$ (full) and the confidence interval $\widetilde{\theta}_i \pm \sqrt{(\Delta\widetilde{O})^2}$ (dashed), calculated from Eq. (3a). (b) Scheme of our algorithm in Sec. 3, for the toy example $\hat{O} = I\hat{X} + \hat{X}I + \hat{X}\hat{X} + Y\hat{Y} + \hat{Z}\hat{Z}$ ($c_i = 1$ for all $i = 1, \ldots, 5$). These PS are depicted as vertices of a graph, connected when they commute [4, 5]. All-to-all connected groups (pink, green, and blue) can be simultaneously measured. Alternatively, one can employ the more expensive (factor $1/(2\alpha)$, see main text) double scheme to assess the magnitude of all $\hat{P}_i$. To choose which group to probe, we assign a virtual measurement [Sec. 3 and Eqs. (8)] that predicts the group that minimizes $(\Delta\widetilde{O})^2$ (based on previous measurements, green in the figure as an example). After the real measurement, $\widetilde{O}$ and $(\Delta\widetilde{O})^2$ are updated following the methods in Sec. 2, and are either outputted by the algorithm (if the budget M is depleted) or employed to assign the next measurement.

information about their signs – see top row in Fig. 1(a): the distribution is symmetric around 1/2. Therefore, our efficient measurement protocol must distribute the shots between the double scheme $\hat{P}_i \otimes \hat{P}_i$ (that targets *all* $|\widetilde{P}_i|$) and the single scheme $\hat{P}_i$ (where every shot helps determine the signs and magnitudes of *some* $\widetilde{P}_i$). Next, we derive the probability distribution function $\mathcal{P}_1$ to estimate $\widetilde{\theta}_i$ and $\widetilde{\theta_i^2}$ from the measurement outcomes. Alongside $\mathcal{P}_2$ (see Sec. 2.2), $\mathcal{P}_1$ will be used by our algorithm [see Sec. 3 and Fig. 1(b)] for the error $(\Delta\widetilde{O})^2$ in Eq. (1), and hence the measurement allocation.

Let us denote $s_{i\pm}$ and $d_{i\pm}$ the number of ± 1 outcomes obtained when probing $\hat{P}_i$ and $\hat{P}_i \otimes \hat{P}_i$, respectively. Following the standard Bayesian approach [45, 46], the expected value $\widetilde{f}$ of any function f of probabilities θ_i can be determined from the posterior probability $\mathcal{P}_1(\theta_i|s_{i\pm}, d_{i\pm})$ by

$$\widetilde{f}(\theta_i) = \int_0^1 f(\theta_i)\mathcal{P}_1(\theta_i|s_{i\pm}, d_{i\pm})d\theta_i, \tag{3a}$$

$$\mathcal{P}_1(\theta_i|s_{i\pm}, d_{i\pm}) \propto \theta_i^{s_{i+}}(1-\theta_i)^{s_{i-}}\phi_i^{d_{i+}}(1-\phi_i)^{d_{i-}}. \tag{3b}$$

Here, we assumed a constant prior for $\mathcal{P}_1$, and that the posterior is normalized: $\int_0^1 \mathcal{P}_1(\theta_i|s_{i\pm}, d_{i\pm})d\theta_i = 1$. Furthermore, ϕ_i is related to θ_i through Eq. (2).

Using Eqs. (3), one can substitute $f \mapsto \theta_i$ [$f \mapsto \theta_i^2$] to calculate $(2\widetilde{\theta}_i - 1)$ [$(\widetilde{\theta_i^2} - \widetilde{\theta}_i^2)$], as required by Eq. (1) to determine $\widetilde{O}$ [$(\Delta\widetilde{O})^2$].

Looking at the features of $\mathcal{P}_1$ in Eq. (3b) helps to understand the advantages of our double scheme and the logic behind the allocation algorithm in Sec. 3 and Fig. 1(b). As can be seen in Fig. 1(a), when $\hat{P}_i$ has never been measured alone ($s_{i+} = s_{i-} = 0$), its associated $\mathcal{P}_1$ often[1] exhibits two symmetric peaks around $\theta_i = 1/2$. This reflects the sign uncertainty that characterizes the double scheme [see Eq. (2)]. As long as $d_{i+}+d_{i-}$ is large enough, very few direct measurements of $\hat{P}_i$ suffice to lift this sign uncertainty, collapse one of the peaks, and obtain an accurate

[1]The distance $|2\theta_i - 1|$ between the two peaks must be larger than their widths $\propto (d_{i+} + d_{i-})^{-x}$, $x \in \left[\frac{1}{4}, 1\right]$ (depending on $\langle \hat{P}_i \rangle$).

estimate for $\widetilde{P}_i$ [top to bottom, Fig. 1(a)]. As discussed in Sec. 4, our measurement algorithm employs the double scheme to gain information about the magnitudes $|\widetilde{P}_i|$ of all PS, and allocates few extra measurements to discern the signs of $\widetilde{P}_i$ and further enhance their estimation accuracies.

2.2 Estimation of $\widetilde{\theta_i\theta_j}$

The Bayesian estimation of $\widetilde{\theta_i\theta_j}$ is conceptually similar to the process outlined in the previous Sec. 2.1. Here, we focus on the main differences, arising from the fact that we have data from multiple random variables that must be accounted for simultaneously. Namely, we need to determine a posterior $\mathcal{P}_2$ that describes the cases of two commuting PS $\hat{P}_i$ and $\hat{P}_j$ that can be measured (a) simultaneously in the double scheme, (b) simultaneously in the single scheme, or (c) individually[2].

Since the outcomes of two PS can be correlated only if they are measured together, whenever $[\hat{P}_i, \hat{P}_j] \neq 0$ we have $\widetilde{\theta_i\theta_j} = \widetilde{\theta}_i\widetilde{\theta}_j$ and therefore $(\widetilde{\theta_i\theta_j} - \widetilde{\theta}_i\widetilde{\theta}_j) = 0$ in Eq. (1)[3]. As such, we can compute $\widetilde{\theta_i\theta_j}$ working in the simultaneous eigenbasis of $\hat{P}_i$ and $\hat{P}_j$. We denote the probability of the outcome $(\pm1, \pm1)$ when probing $\hat{P}_i$ and $\hat{P}_j$ $[\hat{P}_i \otimes \hat{P}_i$ and $\hat{P}_j \otimes \hat{P}_j]$ with $\theta_{i\pm j\pm}$ $[\phi_{i\pm j\pm}]$. By enumerating all possible measurement outcomes, we find relations between $\theta_{i\pm j\pm}$ and $\phi_{i\pm j\pm}$ similar to Eq. (2):

$$\phi_{i+j+} = \theta^2_{i+j+} + \theta^2_{i+j-} + \theta^2_{i-j+} + \theta^2_{i-j-}, \tag{4a}$$
$$\phi_{i+j-} = 2\theta_{i+j+}\theta_{i+j-} + 2\theta_{i-j+}\theta_{i-j-}, \tag{4b}$$
$$\phi_{i-j+} = 2\theta_{i+j+}\theta_{i-j+} + 2\theta_{i+j-}\theta_{i-j-}, \tag{4c}$$
$$\phi_{i-j-} = 2\theta_{i+j+}\theta_{i-j-} + 2\theta_{i+j-}\theta_{i-j+}. \tag{4d}$$

From Eqs. (4), it follows that normalization of $\theta_{i\pm j\pm}$ leads to normalization of $\phi_{i\pm j\pm}$.

[2] Indeed, in the single scheme we probe simultaneously groups of commuting PS. Say that $\hat{P}_i$ and $\hat{P}_j$ commute. Then, in some groups, they belong together and when measured their outcomes are correlated. However, it is possible (and likely) that there are groups in which only one of them belongs. Suppose $\hat{P}_k$ commutes with $\hat{P}_i$ but not $\hat{P}_j$. Then, there is a group which includes $\hat{P}_k$ and $\hat{P}_i$, but not $\hat{P}_j$. When that group is probed, we obtain uncorrelated (from $\hat{P}_j$) outcomes for $\hat{P}_i$, that must be distinguished from the previous, correlated ones. Hence, the existence of scenarios (b) and (c).

[3] This follows from the fact that when $s_{i\pm j\pm} = 0$ for all combinations $\pm\pm$, $\mathcal{P}_2(\vec{\theta}|s_{i\pm}, s_{j\pm}, s_{i\pm j\pm}, d_{i\pm j\pm}) = \mathcal{P}_1(\theta_i|s_{i\pm}, d_{i\pm})\mathcal{P}_1(\theta_j|s_{j\pm}, d_{j\pm})$, see Eqs. (3b) and (5b).

As mentioned in the beginning of this section and in Fig. 1(b), we distinguish the cases in which $\hat{P}_i$ and $\hat{P}_j$ are measured with the double scheme (a) and without, and for the latter scenario when they are simultaneously (b) or independently (c) probed (see footnote 2). We denote $d_{i\pm,j\pm}$ $[s_{i\pm,j\pm}]$ $\{s_{i\pm}$ and $s_{j\pm}\}$ the number of $(\pm1, \pm1)$ $[(\pm1, \pm1)]$ $\{\pm1$ for $\hat{P}_i$ and ±1 for $\hat{P}_j\}$ outcomes in these three cases, (a), [(b)], and {(c)}, respectively[4]. Using $\vec{\theta} = \{\theta_{i+j+}, \theta_{i-j+}, \theta_{i+j-}\}$, Eqs. (3) can be generalized to yield the estimate $\widetilde{g}$ of an arbitrary function g of $\vec{\theta}$ by

$$\widetilde{g}(\vec{\theta}) = \int g(\vec{\theta})\mathcal{P}_2(\vec{\theta}|s_{i\pm}, s_{j\pm}, s_{i\pm j\pm}, d_{i\pm j\pm})d\vec{\theta}, \tag{5a}$$

$$\mathcal{P}_2(\vec{\theta}|s_{i\pm}, s_{j\pm}, s_{i\pm j\pm}, d_{i\pm j\pm}) \propto \begin{cases} \phi^{d_{i+j+}}_{i+j+}\phi^{d_{i+j-}}_{i+j-}\phi^{d_{i-j+}}_{i-j+}\phi^{d_{i-j-}}_{i-j-} & \leftarrow \text{(a)} \\ \theta^{s_{i+j+}}_{i+j+}\theta^{s_{i+j-}}_{i+j-}\theta^{s_{i-j+}}_{i-j+}\theta^{s_{i-j-}}_{i-j-} & \leftarrow \text{(b)} \\ \theta^{s_{i+}}_i(1-\theta_i)^{s_{i-}}\theta^{s_{j+}}_j(1-\theta_j)^{s_{j-}} & \leftarrow \text{(c)} \end{cases} \tag{5b}$$

where $\phi_{i\pm j\pm}$ is given by Eq. (4), $\theta_i = \theta_{i+j+} + \theta_{i+j-}$ and $\theta_j = \theta_{i+j+} + \theta_{i-j+}$.

Similarly to Eq. (3b), Eq. (5b) must be normalized: $\int \mathcal{P}_2(\vec{\theta}|s_{i\pm}, s_{i\pm j\pm}, d_{i\pm j\pm})d\vec{\theta} = 1$. Here and in Eq. (5a), the integral is to be taken over three independent probabilities within $\vec{\theta}$ such that their sum never exceeds unity. We note that the posterior $\mathcal{P}_2$ in Eq. (5a) is the product of the individual posteriors (a), (b) and (c), as these experiments are carried out independently.

By plugging

$$g \mapsto \theta_i\theta_j = (\theta_{i+j+} + \theta_{i+j-})(\theta_{i+j+} + \theta_{i-j+}) \tag{6}$$

into Eqs. (5), together with the results from Sec. 2.1, we find $\widetilde{\theta_i\theta_j}$ and can determine $(\Delta\widetilde{O})^2$ as described in Eq. (1b). This quantity will be central to our algorithm in Sec. 3.

3 Algorithm

In this section, we explain the algorithm that allocates the M shots available to the measurement. Each time, it will choose between simultaneously probing all double PS $\hat{P}_i \otimes \hat{P}_i$ or a group of commuting PS $\hat{P}_i$, see Fig. 1(b). Since the double scheme requires twice the number of qubits to

[4] Notice that $s_{i\pm}$ ($s_{j\pm}$) in Eqs. (5), in agreement with case (c), only includes outcomes where $\hat{P}_i$ ($\hat{P}_j$) has been measured independently from $\hat{P}_j$ ($\hat{P}_i$). This is in contrast to Eqs. (3), where $s_{i\pm}$ includes all outcomes obtained for $\hat{P}_i$.

measure $|\psi\rangle = |\varphi\rangle \otimes |\varphi\rangle$ compared to $|\varphi\rangle$ ($|\varphi\rangle$ being the state to be measured), it costs twice as many measurement shots, and at least twice the amount of resources. To assess the advantage yielded by the double scheme, we define the effective measurement shots

$$M_{\text{eff}} = M_{\text{single}} + 2\alpha M_{\text{double}}, \tag{7}$$

where M_{double} and M_{single} are shots assigned and not assigned to the double scheme, respectively. This definition of M_{eff} is equivalent to saying that measuring all $\hat{P}_i \otimes \hat{P}_i$ requires 2α-times the physical resources. Thus, in order to be advantageous, probing all $\hat{P}_i \otimes \hat{P}_i$ *once* must lower the estimation error $(\Delta\widetilde{O})^2$ in Eq. (1) more than probing 2α groups of commuting PS $\hat{P}_i$ (or the same group 2α times). The factor $\alpha \geq 1$ takes into account the additional resources required to simultaneously diagonalize and measure all $\hat{P}_i \otimes \hat{P}_i$ [4, 9]. As we will see in Sec. 4, despite this extra weight, the double scheme is often advantageous provided the observable $\hat{O}$ is not trivial (in the sense that the p PS within do not commute with each other).

The pipeline of the algorithm is shown in Fig. 1(b). Before each measurement, we determine the *expected* improvement to the estimated error $(\Delta\widetilde{O})^2$ depending on whether the shot is assigned to the double scheme or one of the groups containing commuting PS $\hat{P}_i$. This is done by allocating *virtual* measurements that yield outcomes in agreement with the knowledge already acquired. In the toy example presented in Fig. 1(b), our algorithm considers the blue, green, pink, and black (for the double) groups of PS and assesses how their contribution to the error $(\Delta\widetilde{O})^2$ changes when this virtual measurement is assigned to them.

In practice, the virtual measurement is achieved by adding to $s_{i\pm}$ or $d_{i\pm}$ within $\mathcal{P}_1$ in Eq. (3b) [$s_{i\pm}$, $s_{j\pm}$, $s_{i\pm j\pm}$ or $d_{i\pm j\pm}$ within $\mathcal{P}_2$ in Eq. (5b)] the expected outcome (based on the current knowledge). This is equivalent to a change in the prior, effectively positing that, through a virtual measurement, the outcomes are expected to change according to our prior observations. An heuristic that does this prescribes the following substitutions:

$$\begin{cases} d_{i+} \mapsto d_{i+} + \frac{\widetilde{\phi}_i}{2\alpha} \\ d_{i-} \mapsto d_{i-} + \frac{1-\widetilde{\phi}_i}{2\alpha} \end{cases} \quad \begin{array}{l} \text{with } \widetilde{\phi}_i \\ \text{from Eq. (3)}, \end{array} \tag{8a}$$

$$d_{i\pm j\pm} \mapsto d_{i\pm j\pm} + \frac{\widetilde{\phi}_{i\pm j\pm}}{2\alpha} \quad \begin{array}{l} \text{with } \widetilde{\phi}_{i\pm j\pm} \\ \text{from Eq. (5)}, \end{array} \tag{8b}$$

$$\begin{cases} s_{i+} \mapsto s_{i+} + \widetilde{\theta}_i \\ s_{i-} \mapsto s_{i-} + (1 - \widetilde{\theta}_i) \end{cases} \quad \begin{array}{l} \text{with } \widetilde{\theta}_i \\ \text{from Eq. (3)}, \end{array} \tag{8c}$$

$$s_{i\pm j\pm} \mapsto s_{i\pm j\pm} + \widetilde{\theta}_{i\pm j\pm} \quad \begin{array}{l} \text{with } \widetilde{\theta}_{i\pm j\pm} \\ \text{from Eq. (5)}, \end{array} \tag{8d}$$

where $\widetilde{\theta}_i$ and $\widetilde{\phi}_i$ [$\widetilde{\theta}_{i\pm j\pm}$ and $\widetilde{\phi}_{i\pm j\pm}$] are obtained from $f(\theta_i) \mapsto \theta_i$ and $f(\theta_i) \mapsto \phi_i$ in Eq. (3a) [$g(\vec{\theta}) \mapsto \theta_{i\pm j\pm}$ and $g(\vec{\theta}) \mapsto \phi_{i\pm j\pm}$ in Eq. (5a)], respectively, and making use of Eq. (2) [Eqs. (4)]. Furthermore, the factors 2α dividing $\widetilde{\phi}_i$, $(1-\widetilde{\phi}_i)$, and $\widetilde{\phi}_{i\pm j\pm}$ in Eqs. (8a) and (8b) take into account the extra resources required by the double scheme (see Eq. (7)).

Once the *expected* improvements for all possible choices (all $\hat{P}_i \otimes \hat{P}_i$ for the double scheme or any group of commuting PS $\hat{P}_i$) available to the algorithm are determined, we assign the measurement shot. The choice that leads to the greatest improvement, that is, the smallest resulting error $(\Delta\widetilde{O})^2$ in Eq. (1), will be the one taken by our algorithm. For the toy example in Fig. 1(b), that is the green group. Ideally, this should be done at each of the M steps of the process; in practice, however, it is possible to improve the resource scaling by updating the error prediction only intermittently once we accept minor compromises in terms of obtained errors (see next paragraph). Taking into account the Markov Chain Monte Carlo process (see Refs. [5, 47, 48] for details) employed to estimate $\widetilde{\theta}_i$ and $\widetilde{\theta_i\theta_j}$ in Eq. (1b) ($p^2/2$ terms that need to be calculated) in Sec. 2, the scaling of the entire measurement algorithm is at most $\mathcal{O}(p^2M^3)$ (one factor of M for updating the errors at every shot and M^2 for the Monte Carlo process to reach the required precision). We also point out that assigning virtual measurements does not present a bottleneck of the algorithm. From Eq. (8), it follows that the terms $\widetilde{\theta}_i$, $\widetilde{\theta_i^2}$, and $\widetilde{\theta_i\theta_j}$ in Eq. (1b) need to be updated four times at most, reflecting the possible cases (a), (b), and (c) (twice for the latter).

The time scaling of the algorithm can be drastically improved since the independent contributions $\widetilde{\theta_i\theta_j} - \widetilde{\theta}_i\widetilde{\theta}_j$ in Eq. (1) can be calculated in

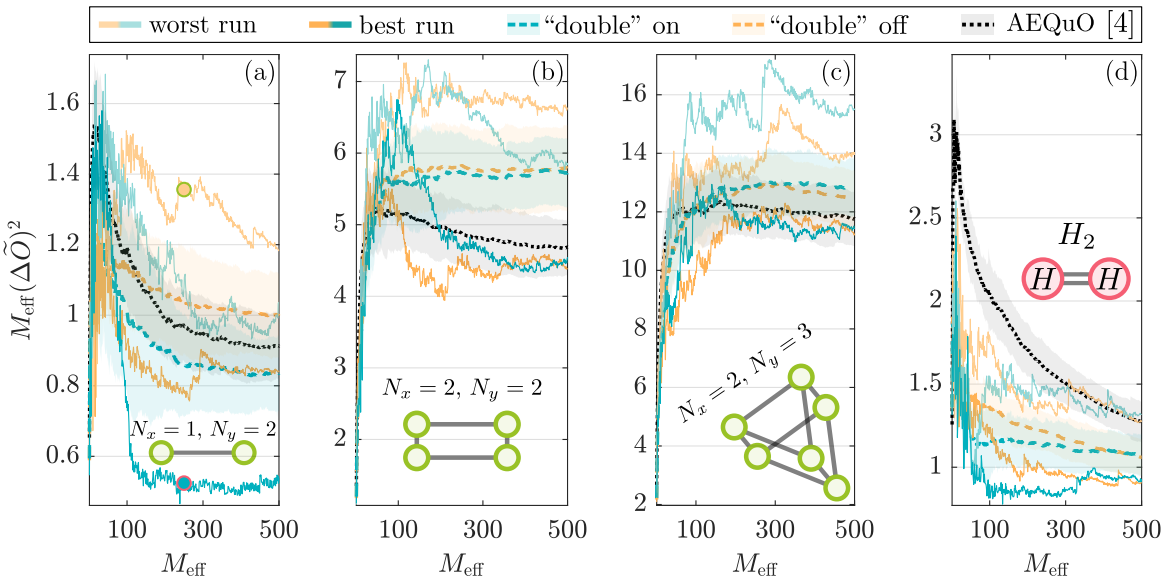


Figure 2: Numerical results with Ising (a-c) and chemistry (d) Hamiltonians – see main text – and their ground states as input $|\varphi\rangle$. Dashed (our method) and dotted (AEQuO [4]) lines are the average from 25 repeated measurement processes, and their shadows correspond to their root mean squares. Blue (Yellow) is used when the double scheme is (not) employed. Dark and light lines are the best and worst runs at large M_{eff}, respectively, within the 25 repetitions. The histograms in Figs. 3(c) refer to the two dots at $M_{\text{eff}} = 250$.

parallel and one can use multiple Monte Carlo chains for each estimation [49, 50]. Furthermore, instead of assigning shots one-by-one, another viable option is to allocate multiple shots at each step of the algorithm, such that the error contributions in Eq. (1b) do not have to be computed at each shot. This can be done based on the expected error improvements either deterministically (allocating shots as to mimic the relative size of the errors for each group) or probabilistically (normalizing the calculated errors and using them as a probability distribution). Hence, a realistic time scaling sits in between $\mathcal{O}(M)$ and $\mathcal{O}(p^2M^3)$, depending on the available computational resources and the chosen allocation strategy.

4 Numerical results

In this section, we benchmark our scheme by turning on and off the possibility of allocating measurement to the double scheme in Fig. 1(b). We therefore unambiguously determine whether it is advantageous to sacrifice some measurement shots [see Sec. 3 and the definition of M_{eff}, Eq. (7)] to probe all PS $\hat{P}_i \otimes \hat{P}_i$ during the measurement process. Furthermore, we compare the estimation errors $(\Delta\widetilde{O})^2$ obtained through these two methods with the errors yielded by AEQuO [4] and by a naive strategy. AEQuO is, to our knowledge, the state of the art in measurement protocols and until this work the only approach that provides single run estimates for both the average $\widetilde{O}$ and the error $(\Delta\widetilde{O})^2$ of the input observable $\hat{O}$, see Eq. (1). With "single run" we mean that the assessment of the error that affects $\widetilde{O}$ does not require repeating the entire measurement process multiple times. Instead, it is produced together with $\widetilde{O}$ in a single run. On the other hand, the naive strategy consists in allocating a fixed percentage of the shots to the double scheme, and randomly assigning the others to groups of commuting PS, weighting each choice with the sum of their coefficients squared – similar to Refs. [14, 20].

For our numerical results, we consider two physical systems. The first is the Hydrogen molecule H_2 in its 6-31G, JW encoding [51, 52], with $q = 8$ qubits and $p = 185$ PS. The second is a quantum Ising model of nuclear spins on rectangular lattices with periodic boundary conditions, N_x and N_y vertices along the x and y directions, with higher-order spin-spin tensor interactions [53] included perturbatively. The three instances considered here are shown in the insets in Figs. 2(a-c) and have (N_x, N_y) equal to $(1, 2)$, $(2, 2)$, and $(2, 3)$, resulting in Hamiltonians $\hat{\mathcal{H}}_{\text{I}}$ with $q = 2, 4, 6$ qubits and $p = 15, 48, 99$ PS, re-

spectively. The $\hat{\mathcal{H}}_{\mathrm{I}}$ used in this work is

$$\begin{aligned}\hat{\mathcal{H}}_{\mathrm{I}} = &\sum_i B_i^z \hat{Z}_i + \sum_{\langle i,k\rangle} J_{ik}\hat{Z}_i\hat{Z}_k \qquad \leftarrow \text{Ising} \\ &+ \sum_i \boldsymbol{b}_i \cdot \hat{\boldsymbol{\sigma}}_i + \sum_{\langle i,k\rangle} \underline{j_{ik}} \cdot \hat{\boldsymbol{\sigma}}_i \otimes \hat{\boldsymbol{\sigma}}_k, \quad \leftarrow \text{tensor}\end{aligned} \tag{9}$$

where $\langle i,k\rangle$ indicates connected vertices i,k in the lattice, B_i^z, J_{ik}, $\boldsymbol{b}_i$ and $\underline{j_{ik}}$ are coefficients characterizing the system, and we explicitly divided single- and two-body terms. Bold indicates vectors (e.g., $\hat{\boldsymbol{\sigma}}_i = \{\hat{X}_i, \hat{Y}_i, \hat{Z}_i\}$) and $\underline{j_{ik}}$ is a symmetric matrix for all i,k. Numerical values used for the results reported in Fig. 2 are listed in App. A.

To assess the advantage yielded by the double scheme, we use the effective measurement shots M_{eff} in Eq. (7). As discussed above, to utilize the double scheme one requires at least twice the resources compared to standard methods. This is also taken into account when we estimate the expected improvement to $(\Delta\widetilde{O})^2$ – see factors $1/(2\alpha)$ in Eqs. (8) and the related discussion. In the following, we will assume the ideal case $\alpha = 1$. We remark, however, that our scheme is implementation ready: it also provides the measurement Clifford circuits for both the single and the double measurements. From these, it is possible to precisely assess the extra resources required to implement the double scheme, and therefore the factor α. In practical scenarios, we expect $\alpha \approx 1$, due to the fact that measurement circuits are often shallower or comparable in depth to the algorithmic circuits preceding them.

The numerical results in Fig. 2 suggest that the advantage of the double scheme is highly dependent on the properties of the observable $\hat{O}$ and input state $|\varphi\rangle$. Looking at panels (a-c) – associated to the Ising Hamiltonians in Eq. (9) – it is possible to notice that the (rescaled) error $M_{\mathrm{eff}}(\Delta\widetilde{O})^2$ obtained with the double scheme (light blue dashed line) is significantly better than the one without (yellow dashed line) only for the case $(N_x = 1, N_y = 2)$. This results from the PS $\hat{P}_i$ in the Hamiltonian being less likely to commute between each other compared to the instances $(N_x = 2, N_y = 2)$ and $(N_x = 2, N_y = 3)$[5].

This conclusion is reinforced by the results in Fig. 2(d), where the observable is the Hydrogen molecule H_2. Molecular Hamiltonians are characterized by many-body interactions such that several of their PS $\hat{P}_i$ have few identities. This means that the double scheme is more advantageous and, for small M_{eff} (see below why), yields better errors $(\Delta\widetilde{O})^2$. Generally, the more likely it is for PS $\hat{P}_i$ to be non-commuting, and the more PS p there are in the decomposition of $\hat{O}$ in Eq. (1), the more advantageous the double scheme is. This is further supported by the number of times M_{double} our algorithm chooses to employ the double scheme, see Fig. 3(d) and discussion below.

An insightful comparison to be made is between AEQuO [4] (black dotted line) and our schemes, with and without the double feature activated. AEQuO is, to our knowledge, the state of the art when it comes to estimating observables on quantum computers while providing statistical errors $(\Delta\widetilde{O})^2$ within a single run. It also allocates measurements based on the instantaneous knowledge it has on the estimated variance $(\Delta\widetilde{O})^2$, although that knowledge is less precise than ours [5] and it cannot access the double scheme. As can be seen in Fig. 2(a), sometimes AEQuO yields better errors $(\Delta\widetilde{O})^2$, despite a lower estimation accuracy of the covariances and equivalent terms $\widetilde{\theta_i\theta_j} - \widetilde{\theta}_i\widetilde{\theta}_j$ in Eq. (1), see Ref. [5]. We believe that this results from different strategies in the shot allocation due to the instantaneous knowledge available at each step. A similar behaviour was observed in Ref. [4] in the comparison of machine-learning and bucket filling approaches. We remark, however, that in the scenarios where the double scheme is advantageous (many non-commuting PS $\hat{P}_i$, Figs. 2(a,d), see above), the double scheme performs considerably better than AEQuO.

Beyond the comparison with AEQuO, it is interesting to understand how well our adaptive allocation scheme works. In Fig. 3(a) we benchmark our double scheme, which allocates shots one by one to the most promising group of commuting PS or all double PS (see Fig. 1(b) and Sec. 3), against a naive strategy that assigns fixed percentages p_{double} and $1 - p_{\mathrm{double}}$ of the measurement budget to the double scheme and to groups of commuting PS, respectively. The three grey lines in the figure are obtained for p_{double} equal to 0.25 (darkest), 0.5 and 0.75 (lightest). This means that, taking $M_{\mathrm{eff}} = 100$ and

[5] Since the Ising Hamiltonian in Eq. (9) is characterized by single- and two-body interactions, the larger the lattice (N_x, N_y) the more identities within each PS $\hat{P}_i$.

$p_{\text{double}} = 0.5$ as an example, the naive strategy assigns 25 shots to the double scheme and 50 to groups of commuting observables (see Eq. (7) with $\alpha = 1$). For the latter, it chooses randomly and non-adaptively (meaning that it does not refine the selection based on previously gathered knowledge, see Sec. 3), preferring groups whose sums of squares of the PS coefficients $\sum_i c_i^2$ are larger [14, 20][6].

As can be seen in the figure, our adaptive approach yields significantly better results. Compared to the naive strategies with $p_{\text{double}} = 0.25$, 0.50 and 0.75, at $M_{\text{eff}} = 500$ our approach yields variances $(\Delta\widetilde{O})^2$ that are on average 1.3, 1.6, and 2.1 times smaller, respectively. Unsurprisingly, the naive strategy that yields better results is the one with $p_{\text{double}} = 0.25$, which approximates best the number of shots assigned to the double scheme by our adaptive approach, see Fig. 3(d). As discussed below, the optimal number of times to employ the double scheme is highly problem dependent. This works in favour of our adaptive approach, where instead of using fixed percentages, it chooses what to measure at each step depending on the available knowledge at that time. An interesting observation is that the naive strategy with $p_{\text{double}} = 0.25$ is more stable than the others with higher $p_{\text{double}} = 0.5$ and 0.75. This can be understood from the shaded area (representative of the variations between runs) and the smaller fluctuations of the average. We believe that this is a consequence of the fact that, with fewer shots assigned to groups of PS, the signs of individual PS are less certain for larger p_{double}.

An important question is whether the error $(\Delta\widetilde{O})^2$ in Eq. (1) is representative of the statistical uncertainty of the estimate $\widetilde{O}$. This is investigated in the scatter plot and the histograms in Figs. 3(b,c), respectively. The first collects single run estimates $\widetilde{O}$ from our approach with (blue) and without (yellow) the double scheme. Error bars are $\sqrt{(\Delta\widetilde{O})^2}$ and, except for the first point at $M_{\text{eff}} = 25$, all intersect the black full line that represents $\langle\hat{O}\rangle$. This suggests that our estimates for both the average and its error are statistically meaningful. The latter, in particular, is not over- or underestimated.

To further confirm this, the histograms in Figs. 3(c) collect 5000 values of

$$\frac{\widetilde{O} - \langle\hat{O}\rangle}{\sqrt{(\Delta\widetilde{O})^2}} \tag{10}$$

obtained in repeated measurement processes with a budget of $M_{\text{eff}} = 250$. For each of those, $\widetilde{O}$ and $(\Delta\widetilde{O})^2$ are calculated using Eq. (1) and the methodology described in Sec. 2. As can be seen in the figure, both histograms with (pink) and without (green) the double scheme are centred nearby zero (black dash-dotted vertical line) and have a root mean square (numbers reported in the plots) that is just below unity. This follows from the Bayesian procedure in Sec. 2. For small measurement budgets M_{eff}, the chosen priors ensure that $(\Delta\widetilde{O})^2$ is overestimated to compensate for biases [5, 45, 46] (due to the non-commutativity of the PS $\hat{P}_i$, for $M_{\text{eff}} = 250$ some PS are measured approximately ten times). For reference, the vertical, full lines in Figs. 3(c) are the averages from the corresponding data points on the single runs highlighted in Fig. 2(a), and their shadows cover $\pm\sqrt{(\Delta\widetilde{O})^2}$. This illustrates that with or without the double scheme, our Bayesian framework correctly reproduces the statistical error. In other words, employing $(\Delta\widetilde{O})^2$ to assess the quality of our estimator is statistically equivalent to $\sum_i(\widetilde{O}_i - \langle\hat{O}\rangle)^2/N_{\text{runs}}$, i.e. the variance obtained from N_{runs} repeated evaluations $\widetilde{O}_i$ of $\langle\hat{O}\rangle$, with the differences that (a) we do not need to perform repeated runs and (b) we do not rely on the exact expectation value $\langle\hat{O}\rangle$, which is generally unknown.

In Fig. 3(d) we show M_{double} as a function of M for all instances considered in Fig. 2. A few characteristics are common to all lines. First, our algorithm [Sec. 3 and Fig. 1(b)] starts with an exploratory phase, where the signs of the PS $\hat{P}_i$ are assessed. This corresponds to the flat part of the lines and avoids situations in which the posterior $\mathcal{P}_1$ presents two peaks [see Fig. 1(a)], resulting in large error contributions. Second, the assignment of shots to the double scheme scales approximately linearly with M. This is a consequence of

[6] After allocating all shots and performing the measurements, we check whether any group of commuting PS has a negative error contribution, similarly to [4], and in that case discard the result. These negative errors, clearly unphysical, arise from insufficient statistics for the sign estimation, and occur for large values of p_{double} (our adaptive allocation ensures the signs are estimated correctly before the double scheme is employed). We are currently working to further enhancing the Bayesian estimation framework [5, 54] to avoid these cases and ensure better estimation overall.

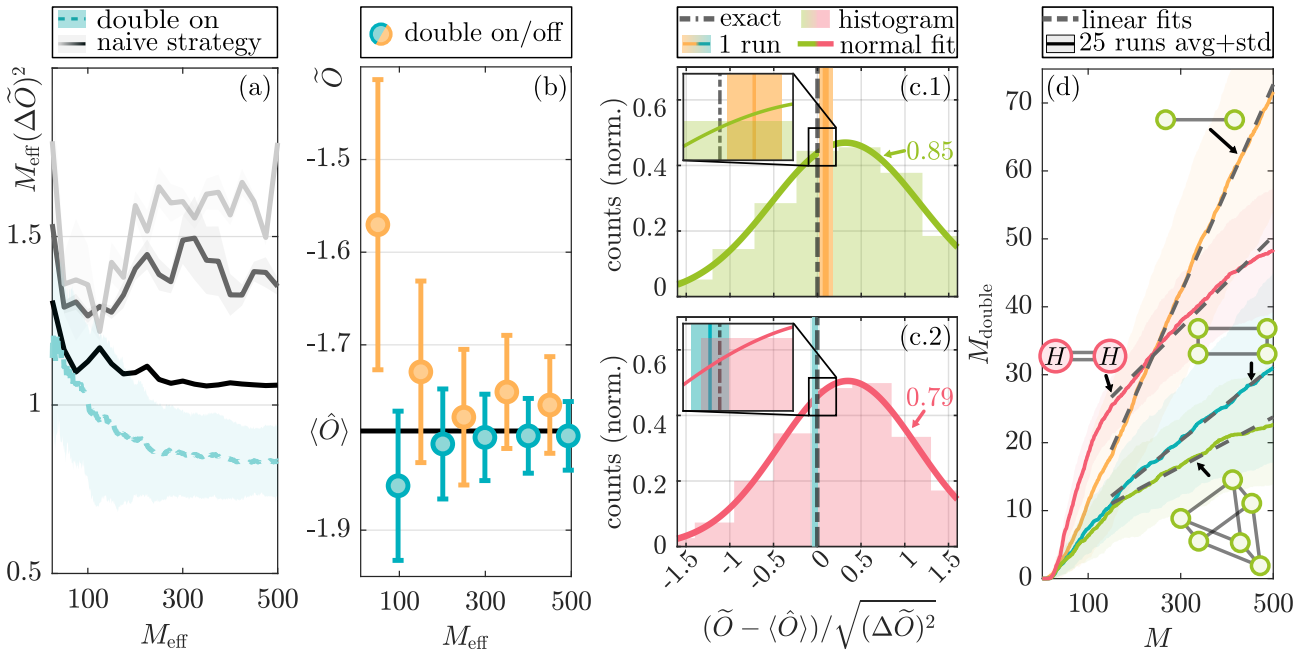


Figure 3: (a) Comparison between our protocol with the double (blue) and the naive strategy (grey, see main text); same settings as in Fig. 2(a). The percentages of shots allocated to measurements without the double scheme is 75% (light grey), 50% (grey), and 25% (dark grey). Panels (b) and (c) demonstrate the statistical validity of the estimators $\widetilde{O}$ and $(\Delta\widetilde{O})^2$ in Sec. 2. (b) reports intermediate values of $\widetilde{O}$ within a single run, with error bars $\sqrt{(\Delta\widetilde{O})^2}$. Blue (Yellow) does (not) employ the double scheme; exact value $\langle\hat{O}\rangle$ is the horizontal black line. In (c), we simulate the measurement process 5000 times with the allocations from the points at $M_{\text{eff}} = 250$ highlighted in Fig. 2(a), and collect the values $(\widetilde{O} - \langle\hat{O}\rangle)/\sqrt{(\Delta\widetilde{O})^2}$ in the histograms. Full lines are Gaussian fits; their root mean squares reported. Vertical full lines and their shadows (visible in the insets) are the averages $\widetilde{O} - \langle\hat{O}\rangle$ and errors $\sqrt{(\Delta\widetilde{O})^2}$ (respectively) extracted from the points highlighted in Fig. 2(a). For (a-c), we used the Ising Hamiltonian with $N_x = 2$ and $N_y = 1$ and the ground state as input $|\varphi\rangle$. In (d), we report M_{double} [see discussion below Eq. (7)] as a function of M for all instances investigated in Fig. 2. Dashed lines are linear fits, with slopes 0.154, 0.055, 0.036 [Ising Hamiltonians in Figs. 2(a-c)] and 0.068 [chemistry molecule in Fig. 2(d)].

the sign uncertainty that limits the accuracy after repeatedly measuring with the double scheme. The slopes of the grey dashed lines in Fig. 3(d) mainly depend on the non-commuting terms in the observable. With more, only a small fraction of the PS $\hat{P}_i$ is probed when the double scheme is not used, and many shots must be employed to collect signs' information. This makes the double scheme more convenient, as indicated by the steeper linear fit of the $(N_x = 1, N_y = 2)$ Ising (yellow line) Hamiltonian. To support this observation, note that the yellow, blue, and green lines (referring to the Ising Hamiltonians) are all flat until around $M \approx 15$. This suggests that (despite larger lattices (N_x, N_y) have more PS p) measuring all PS $\hat{P}_i$ can be done with a similar number of measurements in all Ising instances.

Finally, we draw attention to the results obtained for the chemistry molecules in Figs. 2(d) and 3(d). As can be seen in the latter, the double scheme is chosen more often when the number of shots M is between 20 and 150. Indeed, the slope of the pink like is suddenly reduced for larger $M \gtrsim 150$. Coincidentally, the errors $(\Delta\widetilde{O})^2$ obtained with the double scheme are considerably better within this regime, see Fig. 2(d). This is a consequence of the fact that PS whose averages are approximately zero are less convenient to measure with the double scheme [25], as their error contribution $(\Delta\widetilde{O})^2$ scales $\propto 1/\sqrt{M}$ rather than $\propto 1/M$. In fact, due to the square in Eqs. (2), the accuracy enhancement from the double scheme depends on the expectation values of the PS. The larger $\langle\hat{P}_i\rangle$ is in absolute value, the better. This can be understood from the probability distribution in the top row of Fig. 1(a), where the number of single measurements $s_{i+} + s_{i-}$ is 0. When the peaks are dis-

tinct, which happens for sufficiently many measurements and an average of $\langle \hat{P}_i \rangle$ that is large enough, we get two individual peaks, each with variance $\propto 1/M$ as if we measured only one Pauli with the single measurement scheme. However, with the peaks approaching each other, the variance increases until it becomes $\propto 1/\sqrt{M}$ when $\langle \hat{P}_i \rangle = 0$. In the chemistry example, at around $M \approx 150$ the error contributions from PS with low expectation values start to be important, and therefore the double scheme becomes less effective.

5 Conclusions

We presented an application-ready, multi-copy measurement protocol [22–25] that we call the double measurement scheme, based on a Bayesian enhanced framework that exploits joint PS measurements to improve the estimation of observables on quantum computers. Our algorithm adaptively allocates future measurements based on instantaneous knowledge, yields estimates for both the average and the error of the desired quantity, and exploits general commutation relations between PS. Our results show that the double scheme is often advantageous, in the sense that it requires fewer shots to reach the desired precision. Notably, when the observable is characterized by many non-commuting Pauli terms, the double scheme performs particularly well and greatly outperforms AEQuO [4].

As a first outlook, it would be highly desirable to adapt the code to run on GPU architectures. The resulting acceleration would greatly improve the scaling of the algorithm [55], thus allowing us to tackle larger observables for higher values of shots M. These regimes, with more non-commuting PS, are expected to be more advantageous for the double scheme than the small-scale instances considered in this work. Secondly, we want to extend the double scheme to incorporate sign information more directly. This could be achieved by ancillary systems [29], with the caveat of more complex single shot measurements. It would also be valuable to analyse the performance of this approach in the presence of hardware noise [5], and to integrate it with error-mitigation techniques. Finally, the concept of double measurement could be generalized beyond Pauli observables to other operator families or even continuous variable systems, as well as other measurement schemes that use different posteriors.

Author Contributions

All authors contributed to the manuscript. This work was developed by R.S based on previous work done by A.J., with analytical support from Z.S. and numerics carried out by C.N. under the guidance and direction of L.D. Generative AI was not used in the construction of this manuscript and was at most used sparingly as a resource to look up information otherwise hard to access.

Acknowledgments

We thank Margaritis Voliotis for fruitful probabilistic discussions. LD, CN, and RS acknowledge the EPSRC quantum career development grant EP/W028301/1 and the EPSRC Standard Research grant EP/Z534250/1. We would also like to thank the reviewers for their insightful comments and questions that elevated the manuscript considerably.

References

[1] Seth Lloyd. Universal Quantum Simulators. *Science*, 273(5278):1073–1078, August 1996. ISSN 1095-9203. DOI: 10.1126/science.273.5278.1073.

[2] I. M. Georgescu, S. Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153–185, March 2014. ISSN 1539-0756. DOI: 10.1103/revmodphys.86.153.

[3] Andrew J. Daley, Immanuel Bloch, Christian Kokail, Stuart Flannigan, Natalie Pearson, Matthias Troyer, and Peter Zoller. Practical quantum advantage in quantum simulation. *Nature*, 607(7920):667–676, July 2022. ISSN 1476-4687. DOI: 10.1038/s41586-022-04940-6.

[4] Ariel Shlosberg, Andrew J. Jena, Priyanka Mukhopadhyay, Jan F. Haase, Felix Leditzky, and Luca Dellantonio. Adaptive estimation of quantum observables. *Quantum*, 7:906, January 2023. ISSN 2521-327X. DOI: 10.22331/q-2023-01-26-906.

[5] Rick P. A. Simon, Michael Meth, Francesco Martini, Peter Tirler, Andrew Jena, Martin Ringbauer, and Luca Dellantonio. An error-aware and adaptive method for the estimation of quantum observables on qudit-based quantum computers, 2026. URL https://doi.org/10.48550/arXiv.2605.00682.

[6] Artur F. Izmaylov, Tzu-Ching Yen, Robert A. Lang, and Vladyslav Verteletskyi. Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method. *Journal of Chemical Theory and Computation*, 16(1): 190–195, November 2019. ISSN 1549-9626. DOI: 10.1021/acs.jctc.9b00791.

[7] Jarrod R. McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2): 023023, February 2016. ISSN 1367-2630. DOI: 10.1088/1367-2630/18/2/023023.

[8] Bujiao Wu, Jinzhao Sun, Qi Huang, and Xiao Yuan. Overlapped grouping measurement: A unified framework for measuring quantum states. *Quantum*, 7:896, January 2023. ISSN 2521-327X. DOI: 10.22331/q-2023-01-13-896.

[9] Andrew Jena, Scott Genin, and Michele Mosca. Pauli partitioning with respect to gate sets, 2019. URL https://doi.org/10.48550/arXiv.1907.07859.

[10] Pranav Gokhale, Olivia Angiuli, Yongshan Ding, Kaiwen Gui, Teague Tomesh, Martin Suchara, Margaret Martonosi, and Frederic T. Chong. Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families, 2019. URL https://doi.org/10.48550/arXiv.1907.13623.

[11] Vladyslav Verteletskyi, Tzu-Ching Yen, and Artur F. Izmaylov. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *The Journal of Chemical Physics*, 152(12), March 2020. ISSN 1089-7690. DOI: 10.1063/1.5141458.

[12] Tzu-Ching Yen, Vladyslav Verteletskyi, and Artur F. Izmaylov. Measuring all compatible operators in one series of single-qubit measurements using unitary transformations. *Journal of Chemical Theory and Computation*, 16(4):2400–2409, March 2020. ISSN 1549-9626. DOI: 10.1021/acs.jctc.0c00008.

[13] Ikko Hamamura and Takashi Imamichi. Efficient evaluation of quantum observables using entangled measurements. *npj Quantum Information*, 6(1), June 2020. ISSN 2056-6387. DOI: 10.1038/s41534-020-0284-2.

[14] Ophelia Crawford, Barnaby van Straaten, Daochen Wang, Thomas Parks, Earl Campbell, and Stephen Brierley. Efficient quantum measurement of pauli operators in the presence of finite sampling error. *Quantum*, 5:385, January 2021. ISSN 2521-327X. DOI: 10.22331/q-2021-01-20-385.

[15] Charles Hadfield. Adaptive pauli shadows for energy estimation, 2021. URL https://doi.org/10.48550/arXiv.2105.12207.

[16] Charles Hadfield, Sergey Bravyi, Rudy Raymond, and Antonio Mezzacapo. Measurements of quantum hamiltonians with locally-biased classical shadows. *Communications in Mathematical Physics*, 391(3): 951–967, March 2022. ISSN 1432-0916. DOI: 10.1007/s00220-022-04343-8.

[17] Zhenyu Cai, Adrian Chapman, Hamza Jnane, and Bálint Koczor. Biased estimator channels for classical shadows. *Physical Review A*, 111(3):l030402, March 2025. ISSN 2469-9934. DOI: 10.1103/physreva.111.l030402.

[18] Hong-Ye Hu, Andi Gu, Swarnadeep Majumder, Hang Ren, Yipei Zhang, Derek S. Wang, Yi-Zhuang You, Zlatko Minev, Susanne F. Yelin, and Alireza Seif. Demonstration of robust and efficient quantum property learning with shallow shadows. *Nature Communications*, 16(1), March 2025. ISSN 2041-1723. DOI: 10.1038/s41467-025-57349-w.

[19] Scott Aaronson. Shadow tomography of quantum states, 2017. URL https://arxiv.org/abs/1711.01053.

[20] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Efficient estimation of pauli observables by derandomization. *Physical Review Letters*, 127(3):030503, July 2021. ISSN 1079-7114. DOI: 10.1103/physrevlett.127.030503.

[21] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few mea-

surements. *Nature Physics*, 16(10):1050–1057, June 2020. ISSN 1745-2481. DOI: 10.1038/s41567-020-0932-7.

[22] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li. Exponential separations between learning with and without quantum memory, 2021. URL https://arxiv.org/abs/2111.05881.

[23] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, and Jarrod R. McClean. Quantum advantage in learning from experiments. *Science*, 376(6598):1182–1186, 2022. DOI: 10.1126/science.abn7293.

[24] Sitan Chen, Weiyuan Gong, and Qi Ye. Optimal tradeoffs for estimating pauli observables, 2024. URL https://arxiv.org/abs/2404.19105.

[25] Robbie King, David Gosset, Robin Kothari, and Ryan Babbush. *Triply efficient shadow tomography*, pages 914–946. 2024. DOI: 10.1137/1.9781611978322.27.

[26] Masaya Kohda, Ryosuke Imai, Keita Kanno, Kosuke Mitarai, Wataru Mizukami, and Yuya O. Nakagawa. Quantum expectation-value estimation by computational basis sampling. *Physical Review Research*, 4(3):033173, September 2022. ISSN 2643-1564. DOI: 10.1103/physrevresearch.4.033173.

[27] Andrew Arrasmith, Lukasz Cincio, Rolando D. Somma, and Patrick J. Coles. Operator sampling for shot-frugal optimization in variational algorithms, 2020. URL https://doi.org/10.48550/arXiv.2004.06252.

[28] Stefan Hillmich, Charles Hadfield, Rudy Raymond, Antonio Mezzacapo, and Robert Wille. Decision diagrams for quantum measurements with shallow circuits. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 24–34. IEEE, October 2021. DOI: 10.1109/qce52317.2021.00018.

[29] Liubov A. Markovich, Attaallah Almasi, Sina Zeytinoğlu, and Johannes Borregaard. Quantum memory assisted observable estimation. *Quantum*, 9:1655, March 2025. ISSN 2521-327X. DOI: 10.22331/q-2025-03-06-1655.

[30] Giacomo Torlai, Guglielmo Mazzola, Giuseppe Carleo, and Antonio Mezzacapo. Precise measurement of quantum observables with neural-network estimators. *Physical Review Research*, 2(2):022060, June 2020. ISSN 2643-1564. DOI: 10.1103/physrevresearch.2.022060.

[31] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, August 2021. ISSN 2522-5820. DOI: 10.1038/s42254-021-00348-9.

[32] R. R. Ferguson, L. Dellantonio, A. Al Balushi, K. Jansen, W. Dür, and C. A. Muschik. Measurement-Based Variational Quantum Eigensolver. *Physical Review Letters*, 126(22):220501, June 2021. ISSN 1079-7114. DOI: 10.1103/physrevlett.126.220501.

[33] Shouvanik Chakrabarti, Rajiv Krishnakumar, Guglielmo Mazzola, Nikitas Stamatopoulos, Stefan Woerner, and William J. Zeng. A Threshold for Quantum Advantage in Derivative Pricing. *Quantum*, 5:463, June 2021. ISSN 2521-327X. DOI: 10.22331/q-2021-06-01-463.

[34] Koki Kitai, Jiang Guo, Shenghong Ju, Shu Tanaka, Koji Tsuda, Junichiro Shiomi, and Ryo Tamura. Designing metamaterials with quantum annealing and factorization machines. *Physical Review Research*, 2(1):013319, March 2020. ISSN 2643-1564. DOI: 10.1103/physrevresearch.2.013319.

[35] Y. Cao, J. Romero, and A. Aspuru-Guzik. Potential of quantum computing for drug discovery. *IBM Journal of Research and Development*, 62(6):6:1–6:20, November 2018. ISSN 0018-8646. DOI: 10.1147/jrd.2018.2888987.

[36] Danny Paulson, Luca Dellantonio, Jan F. Haase, Alessio Celi, Angus Kan, Andrew Jena, Christian Kokail, Rick van Bijnen, Karl Jansen, Peter Zoller, and Christine A. Muschik. Simulating 2D Effects in Lattice Gauge Theories on a Quantum Computer. *PRX Quantum*, 2(3):030334, August 2021. ISSN 2691-3399. DOI: 10.1103/prxquantum.2.030334.

[37] Aurora Maurizio and Guglielmo Mazzola. Quantum computing for genomics: Con-

ceptual challenges and practical perspectives. *PRX Life*, 3:047001, Oct 2025. DOI: 10.1103/h49j-bsc6.

[38] Adelina Bärligea, Benedikt Poggel, and Jeanette Miriam Lorenz. Scalability challenges in variational quantum optimization under stochastic noise. *Phys. Rev. A*, 112: 032407, Sep 2025. DOI: 10.1103/rgyh-8xw8.

[39] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum Chemistry in the Age of Quantum Computing. *Chemical Reviews*, 119(19):10856–10915, August 2019. ISSN 1520-6890. DOI: 10.1021/acs.chemrev.8b00803.

[40] Mario Motta and Julia E. Rice. Emerging quantum computing algorithms for quantum chemistry. *WIREs Computational Molecular Science*, 12(3), December 2021. ISSN 1759-0884. DOI: 10.1002/wcms.1580.

[41] Ryan V. Mishmash, Tanvi P. Gujarati, Mario Motta, Huanchen Zhai, Garnet Kin-Lic Chan, and Antonio Mezzacapo. Hierarchical Clifford Transformations to Reduce Entanglement in Quantum Chemistry Wave Functions. *Journal of Chemical Theory and Computation*, 19(11):3194–3208, May 2023. ISSN 1549-9626. DOI: 10.1021/acs.jctc.3c00228.

[42] Philip Richerme, Melissa C. Revelle, Christopher G. Yale, Daniel Lobser, Ashlyn D. Burch, Susan M. Clark, Debadrita Saha, Miguel Angel Lopez-Ruiz, Anurag Dwivedi, Jeremy M. Smith, Sam A. Norrell, Amr Sabry, and Srinivasan S. Iyengar. Quantum Computation of Hydrogen Bond Dynamics and Vibrational Spectra. *The Journal of Physical Chemistry Letters*, 14 (32):7256–7263, August 2023. ISSN 1948-7185. DOI: 10.1021/acs.jpclett.3c01601.

[43] R. M. Dudley. Central Limit Theorems for Empirical Measures. *The Annals of Probability*, 6(6):899 – 929, 1978. DOI: 10.1214/aop/1176995384.

[44] Hans Fischer. *A history of the central limit theorem: from classical to modern probability theory*, volume 4. Springer, 2011. DOI: 10.1007/978-0-387-87857-7.

[45] Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 1995. DOI: 10.1201/9780429258411.

[46] John Kruschke. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press, 2014. DOI: 10.1016/B978-0-12-405888-0.09999-2.

[47] M. E. J. Newman and G. T. Barkema. *Monte Carlo Methods in Statistical Physics*. Clarendon Press, 1999. DOI: 10.1093/oso/9780198517962.001.0001.

[48] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of markov chain monte carlo*. CRC press, 2011. DOI: 10.1201/b10905.

[49] Andrew Gelman and Donald B. Rubin. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457 – 472, 1992. DOI: 10.1214/ss/1177011136. URL `https://doi.org/10.1214/ss/1177011136`.

[50] Stephen P. Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455, 1998. DOI: 10.1080/10618600.1998.10474787.

[51] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, 2019. DOI: 10.1021/acs.chemrev.8b00803.

[52] Jarrod R McClean, Nicholas C Rubin, Kevin J Sung, Ian D Kivlichan, Xavier Bonet-Monroig, Yudong Cao, Chengyu Dai, E Schuyler Fried, Craig Gidney, Brendan Gimby, Pranav Gokhale, Thomas Häner, Tarini Hardikar, Vojtěch Havlíček, Oscar Higgott, Cupjin Huang, Josh Izaac, Zhang Jiang, Xinle Liu, Sam McArdle, Matthew Neeley, Thomas O'Brien, Bryan O'Gorman, Isil Ozfidan, Maxwell D Radin, Jhonathan Romero, Nicolas P D Sawaya, Bruno Senjean, Kanav Setia, Sukin Sim, Damian S Steiger, Mark Steudtner, Qiming Sun, Wei Sun, Daochen Wang, Fang Zhang, and Ryan

Babbush. Openfermion: the electronic structure package for quantum computers. *Quantum Science and Technology*, 5(3):034014, jun 2020. DOI: 10.1088/2058-9565/ab8ebc.

[53] J. W. Blanchard, T. F. Sjolander, J. P. King, M. P. Ledbetter, E. H. Levine, V. S. Bajaj, D. Budker, and A. Pines. Measurement of untruncated nuclear spin interactions via zero- to ultralow-field nuclear magnetic resonance. *Phys. Rev. B*, 92:220202, Dec 2015. DOI: 10.1103/PhysRevB.92.220202.

[54] Ariel Shlosberg et al. In preparation.

[55] Alexander Terenin, Shawfeng Dong, and David Draper. GPU-accelerated Gibbs sampling: A case study of the Horseshoe Probit model. *Statistics and Computing*, 29(2): 301–310, March 2019. ISSN 1573-1375. DOI: 10.1007/s11222-018-9809-3.

A Numerical simulations

In this section, we include the randomly generated coefficients c_i and PS $\hat{P}_i$ for the Ising Hamiltonians, Eq. (9) and (N_x, N_y) equal to $(1,2)$ [table 1], $(2,2)$ [table 2], and $(2,3)$ [table 3].

Coefficient c_i	**PS** $\hat{P}_i$
-0.005	$\hat{X}\hat{I}$
-0.044	$\hat{Y}\hat{I}$
1.028	$\hat{Z}\hat{I}$
0.198	$\hat{I}\hat{X}$
-0.053	$\hat{I}\hat{Y}$
1.100	$\hat{I}\hat{Z}$
0.121	$\hat{X}\hat{X}$
-0.022	$\hat{X}\hat{Y}$
-0.104	$\hat{X}\hat{Z}$
-0.022	$\hat{Y}\hat{X}$
0.034	$\hat{Y}\hat{Y}$
-0.143	$\hat{Y}\hat{Z}$
-0.104	$\hat{Z}\hat{X}$
-0.143	$\hat{Z}\hat{Y}$
0.416	$\hat{Z}\hat{Z}$

Table 1: Coefficients and PS for the $(N_x, N_y) = (1,2)$ Ising Hamiltonian.

Coefficient c_i	**PS** $\hat{P}_i$	**Coefficient** c_i	**PS** $\hat{P}_i$
-0.005	$\hat{X}\hat{I}\hat{I}\hat{I}$	-0.071	$\hat{I}\hat{X}\hat{I}\hat{Y}$
-0.044	$\hat{Y}\hat{I}\hat{I}\hat{I}$	-0.096	$\hat{I}\hat{X}\hat{I}\hat{Z}$
1.028	$\hat{Z}\hat{I}\hat{I}\hat{I}$	-0.071	$\hat{I}\hat{Y}\hat{I}\hat{X}$
0.198	$\hat{I}\hat{X}\hat{I}\hat{I}$	-0.055	$\hat{I}\hat{Y}\hat{I}\hat{Y}$
-0.053	$\hat{I}\hat{Y}\hat{I}\hat{I}$	-0.039	$\hat{I}\hat{Y}\hat{I}\hat{Z}$
1.100	$\hat{I}\hat{Z}\hat{I}\hat{I}$	-0.096	$\hat{I}\hat{Z}\hat{I}\hat{X}$
-0.144	$\hat{I}\hat{I}\hat{X}\hat{I}$	-0.039	$\hat{I}\hat{Z}\hat{I}\hat{Y}$
0.161	$\hat{I}\hat{I}\hat{Y}\hat{I}$	0.515	$\hat{I}\hat{Z}\hat{I}\hat{Z}$
0.893	$\hat{I}\hat{I}\hat{Z}\hat{I}$	-0.059	$\hat{I}\hat{I}\hat{X}\hat{X}$
-0.090	$\hat{I}\hat{I}\hat{I}\hat{X}$	-0.063	$\hat{I}\hat{I}\hat{X}\hat{Y}$
0.110	$\hat{I}\hat{I}\hat{I}\hat{Y}$	0.101	$\hat{I}\hat{I}\hat{X}\hat{Z}$
0.805	$\hat{I}\hat{I}\hat{I}\hat{Z}$	-0.063	$\hat{I}\hat{I}\hat{Y}\hat{X}$
0.033	$\hat{X}\hat{I}\hat{X}\hat{I}$	-0.126	$\hat{I}\hat{I}\hat{Y}\hat{Y}$
0.014	$\hat{X}\hat{I}\hat{Y}\hat{I}$	0.054	$\hat{I}\hat{I}\hat{Y}\hat{Z}$
0.067	$\hat{X}\hat{I}\hat{Z}\hat{I}$	0.101	$\hat{I}\hat{I}\hat{Z}\hat{X}$
0.014	$\hat{Y}\hat{I}\hat{X}\hat{I}$	0.054	$\hat{I}\hat{I}\hat{Z}\hat{Y}$
-0.028	$\hat{Y}\hat{I}\hat{Y}\hat{I}$	0.645	$\hat{I}\hat{I}\hat{Z}\hat{Z}$
-0.162	$\hat{Y}\hat{I}\hat{Z}\hat{I}$	0.067	$\hat{Z}\hat{I}\hat{X}\hat{I}$
0.413	$\hat{Z}\hat{I}\hat{Z}\hat{I}$	-0.149	$\hat{X}\hat{X}\hat{I}\hat{I}$
-0.154	$\hat{X}\hat{Y}\hat{I}\hat{I}$	-0.060	$\hat{X}\hat{Z}\hat{I}\hat{I}$
-0.154	$\hat{Y}\hat{X}\hat{I}\hat{I}$	-0.116	$\hat{Y}\hat{Y}\hat{I}\hat{I}$
0.027	$\hat{Y}\hat{Z}\hat{I}\hat{I}$	-0.060	$\hat{Z}\hat{X}\hat{I}\hat{I}$
0.027	$\hat{Z}\hat{Y}\hat{I}\hat{I}$	0.429	$\hat{Z}\hat{Z}\hat{I}\hat{I}$
-0.076	$\hat{I}\hat{X}\hat{I}\hat{X}$		

Table 2: Coefficients and PS for the $(N_x, N_y) = (2,2)$ Ising Hamiltonian.

Coefficient c_i	**PS** $\hat{P}_i$	**Coefficient** c_i	**PS** $\hat{P}_i$
-0.005	$\hat{X}\hat{I}\hat{I}\hat{I}\hat{I}\hat{I}$	0.136	$\hat{I}\hat{Y}\hat{I}\hat{I}\hat{Z}\hat{I}$
-0.044	$\hat{Y}\hat{I}\hat{I}\hat{I}\hat{I}\hat{I}$	0.004	$\hat{I}\hat{Z}\hat{I}\hat{I}\hat{X}\hat{I}$
1.028	$\hat{Z}\hat{I}\hat{I}\hat{I}\hat{I}\hat{I}$	0.136	$\hat{I}\hat{Z}\hat{I}\hat{I}\hat{Y}\hat{I}$
0.198	$\hat{I}\hat{X}\hat{I}\hat{I}\hat{I}\hat{I}$	0.559	$\hat{I}\hat{Z}\hat{I}\hat{I}\hat{Z}\hat{I}$
-0.053	$\hat{I}\hat{Y}\hat{I}\hat{I}\hat{I}\hat{I}$	0.122	$\hat{I}\hat{X}\hat{X}\hat{I}\hat{I}\hat{I}$
1.100	$\hat{I}\hat{Z}\hat{I}\hat{I}\hat{I}\hat{I}$	0.055	$\hat{I}\hat{X}\hat{Y}\hat{I}\hat{I}\hat{I}$
-0.144	$\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}\hat{I}$	0.156	$\hat{I}\hat{X}\hat{Z}\hat{I}\hat{I}\hat{I}$
0.161	$\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}\hat{I}$	0.055	$\hat{I}\hat{Y}\hat{X}\hat{I}\hat{I}\hat{I}$
0.893	$\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}\hat{I}$	0.127	$\hat{I}\hat{Y}\hat{Y}\hat{I}\hat{I}\hat{I}$
-0.090	$\hat{I}\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}$	0.087	$\hat{I}\hat{Y}\hat{Z}\hat{I}\hat{I}\hat{I}$
0.110	$\hat{I}\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}$	0.156	$\hat{I}\hat{Z}\hat{X}\hat{I}\hat{I}\hat{I}$
0.805	$\hat{I}\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}$	0.087	$\hat{I}\hat{Z}\hat{Y}\hat{I}\hat{I}\hat{I}$
0.061	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{X}\hat{I}$	0.584	$\hat{I}\hat{Z}\hat{Z}\hat{I}\hat{I}\hat{I}$
-0.007	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Y}\hat{I}$	0.034	$\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}\hat{X}$
0.950	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Z}\hat{I}$	0.118	$\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}\hat{Y}$
0.013	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{I}\hat{X}$	0.071	$\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}\hat{Z}$
-0.264	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{I}\hat{Y}$	0.118	$\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}\hat{X}$
0.876	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{I}\hat{Z}$	0.001	$\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}\hat{Y}$
-0.076	$\hat{X}\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}$	0.12	$\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}\hat{Z}$
-0.071	$\hat{X}\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}$	0.071	$\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}\hat{X}$
-0.096	$\hat{X}\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}$	0.12	$\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}\hat{Y}$
-0.071	$\hat{Y}\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}$	0.539	$\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}\hat{Z}$
-0.055	$\hat{Y}\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}$	0.078	$\hat{I}\hat{I}\hat{I}\hat{X}\hat{X}\hat{I}$
-0.039	$\hat{Y}\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}$	0.06	$\hat{I}\hat{I}\hat{I}\hat{X}\hat{Y}\hat{I}$
-0.096	$\hat{Z}\hat{I}\hat{I}\hat{X}\hat{I}\hat{I}$	0.068	$\hat{I}\hat{I}\hat{I}\hat{X}\hat{Z}\hat{I}$
-0.039	$\hat{Z}\hat{I}\hat{I}\hat{Y}\hat{I}\hat{I}$	0.06	$\hat{I}\hat{I}\hat{I}\hat{Y}\hat{X}\hat{I}$
0.515	$\hat{Z}\hat{I}\hat{I}\hat{Z}\hat{I}\hat{I}$	0.064	$\hat{I}\hat{I}\hat{I}\hat{Y}\hat{Y}\hat{I}$
-0.059	$\hat{X}\hat{X}\hat{I}\hat{I}\hat{I}\hat{I}$	0.067	$\hat{I}\hat{I}\hat{I}\hat{Y}\hat{Z}\hat{I}$
-0.063	$\hat{X}\hat{Y}\hat{I}\hat{I}\hat{I}\hat{I}$	0.068	$\hat{I}\hat{I}\hat{I}\hat{Z}\hat{X}\hat{I}$
0.101	$\hat{X}\hat{Z}\hat{I}\hat{I}\hat{I}\hat{I}$	0.067	$\hat{I}\hat{I}\hat{I}\hat{Z}\hat{Y}\hat{I}$
-0.063	$\hat{Y}\hat{X}\hat{I}\hat{I}\hat{I}\hat{I}$	0.437	$\hat{I}\hat{I}\hat{I}\hat{Z}\hat{Z}\hat{I}$
-0.126	$\hat{Y}\hat{Y}\hat{I}\hat{I}\hat{I}\hat{I}$	0.061	$\hat{I}\hat{I}\hat{I}\hat{X}\hat{I}\hat{X}$
0.054	$\hat{Y}\hat{Z}\hat{I}\hat{I}\hat{I}\hat{I}$	0.111	$\hat{I}\hat{I}\hat{I}\hat{X}\hat{I}\hat{Y}$
0.101	$\hat{Z}\hat{X}\hat{I}\hat{I}\hat{I}\hat{I}$	0.076	$\hat{I}\hat{I}\hat{I}\hat{X}\hat{I}\hat{Z}$
0.054	$\hat{Z}\hat{Y}\hat{I}\hat{I}\hat{I}\hat{I}$	0.111	$\hat{I}\hat{I}\hat{I}\hat{Y}\hat{I}\hat{X}$
0.645	$\hat{Z}\hat{Z}\hat{I}\hat{I}\hat{I}\hat{I}$	0.169	$\hat{I}\hat{I}\hat{I}\hat{Y}\hat{I}\hat{Y}$
-0.003	$\hat{X}\hat{I}\hat{X}\hat{I}\hat{I}\hat{I}$	0.03	$\hat{I}\hat{I}\hat{I}\hat{Y}\hat{I}\hat{Z}$
-0.001	$\hat{X}\hat{I}\hat{Y}\hat{I}\hat{I}\hat{I}$	0.076	$\hat{I}\hat{I}\hat{I}\hat{Z}\hat{I}\hat{X}$
-0.121	$\hat{X}\hat{I}\hat{Z}\hat{I}\hat{I}\hat{I}$	0.03	$\hat{I}\hat{I}\hat{I}\hat{Z}\hat{I}\hat{Y}$
-0.001	$\hat{Y}\hat{I}\hat{X}\hat{I}\hat{I}\hat{I}$	0.484	$\hat{I}\hat{I}\hat{I}\hat{Z}\hat{I}\hat{Z}$
-0.059	$\hat{Y}\hat{I}\hat{Y}\hat{I}\hat{I}\hat{I}$	0.152	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{X}\hat{X}$
-0.042	$\hat{Y}\hat{I}\hat{Z}\hat{I}\hat{I}\hat{I}$	0.038	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{X}\hat{Y}$
-0.121	$\hat{Z}\hat{I}\hat{X}\hat{I}\hat{I}\hat{I}$	0.011	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{X}\hat{Z}$
-0.042	$\hat{Z}\hat{I}\hat{Y}\hat{I}\hat{I}\hat{I}$	0.038	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Y}\hat{X}$
0.569	$\hat{Z}\hat{I}\hat{Z}\hat{I}\hat{I}\hat{I}$	0.146	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Y}\hat{Y}$
0.094	$\hat{I}\hat{X}\hat{I}\hat{I}\hat{X}\hat{I}$	0.012	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Y}\hat{Z}$
-0.144	$\hat{I}\hat{X}\hat{I}\hat{I}\hat{Y}\hat{I}$	0.011	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Z}\hat{X}$
0.004	$\hat{I}\hat{X}\hat{I}\hat{I}\hat{Z}\hat{I}$	0.012	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Z}\hat{Y}$
-0.144	$\hat{I}\hat{Y}\hat{I}\hat{I}\hat{X}\hat{I}$	0.601	$\hat{I}\hat{I}\hat{I}\hat{I}\hat{Z}\hat{Z}$
-0.138	$\hat{I}\hat{Y}\hat{I}\hat{I}\hat{Y}\hat{I}$		

Table 3: Coefficients and PS for the $(N_x, N_y) = (2,3)$ Ising Hamiltonian.