Parameter estimation with mixed-state quantum computation

Sergio Boixo*
Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
and University of New Mexico, Albuquerque, New Mexico 87131, USA

Rolando D. Somma†
Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

(Received 14 August 2007; published 19 May 2008)

We present a quantum algorithm to estimate parameters at the quantum metrology limit using deterministic quantum computation with one bit. When the interactions occurring in a quantum system are described by a Hamiltonian $H=\theta H_0$, we estimate θ by zooming in on previous estimations and by implementing an adaptive Bayesian procedure. The final result of the algorithm is an updated estimation of θ whose variance has been decreased in proportion to the time of evolution under H. For the problem of estimating several parameters, we implement dynamical-decoupling techniques and use the results of single parameter estimation. The cases of discrete-time evolution and reference frame alignment are also studied within the adaptive approach.

DOI: 10.1103/PhysRevA.77.052320 PACS number(s): 03.67.Lx, 03.65.Ta, 06.20.Dk

I. INTRODUCTION

Quantum mechanics provides resources that allow us to determine physical properties at the highest possible accuracy established by generalized uncertainty relations [1–3]. Exploiting quantum coherence enables us to estimate parameters [4–6] and expectation values of observables [7] with better resource scaling than classically possible. In this paper, we are interested in the estimation of interaction parameters (e.g., external fields), when the interaction acts independently on n quantum subsystems in a probe [8]. We quantify our resource of interest by $N=nT$, given by the product of the number of subsystems and the interaction time T. The standard quantum limit (SQL) precision in the estimation of an interaction parameter is of order $O(1/\sqrt{N})$, achievable with $O(N)$ independent measurements at a fixed T. The optimal precision for such an estimation, however, is given by the Heisenberg limit and is known to be of order $O(1/N)$. Achieving it requires the preparation of entangled quantum states in the probe [4].

We are interested in estimating parameters at the so-called quantum metrology limit (QML). This can be obtained by a series of estimations performed at different interaction times, while keeping the size of the probe fixed [5,7]. If for a total interaction time T, the precision of the estimation is of order $O(1/T)$, we say that the QML has been achieved. This *sequential* protocol, which is the one exploited in this paper, does not require quantum entanglement in the input state, although the response to uncorrelated decoherence, for an unconstrained interaction time, is the same as that of the entangled protocol [9]. Any method that allows us to achieve the QML clearly provides an improvement over the SQL, since for the same amount of resources (i.e., $N=nT$), the returned precision can be highly enhanced.

Quantum methods (algorithms) designed to beat the SQL could have a wide range of applications, from highly sensitive magnetometry [10] to atomic clock synchronization [11]. In addition, phase estimation, a problem related to parameter estimation, is one of the cornerstones of quantum computation [12,13]. In this paper we show that the QML can be achieved in some cases even if the initial state is the completely mixed state of all except one of the quantum systems, avoiding the complexity associated with initial pure, entangled, state preparation. Although here we consider multiqubit probes, generalization of our algorithms to higher dimensional systems is straightforward.

Specifically, we use deterministic quantum computation with one bit (DQC1), which was initially described in Ref. [14] in the context of high temperature ensemble quantum computation using liquid-state NMR techniques [15]. Although less powerful than the standard model of quantum computation, DQC1 is believed to outperform the classical probabilistic computational model [16]. In DQC1, the initial state ρ_0 of a set of $n+1$ qubits corresponds to having the first (ancilla) qubit a in the pure state $|0\rangle_{\mathsf{a}}$, while the state of the remaining n qubits (probe) is completely mixed. That is,

$$\rho_0 = \frac{1}{2^n}(|0\rangle_{\mathsf{a}}\langle 0| \otimes \mathbb{1}_n). \tag{1}$$

The state ρ_0 is then unitarily evolved and DQC1 returns a noisy expectation value of a Pauli operator on the ancilla qubit. If the evolution is performed by applying a unitary operation controlled by a (i.e., a controlled-U or ${}^cU$ operation), DQC1 allows us to estimate the renormalized trace of U at a certain, fixed precision (Fig. 1).

We assume that a single run of the DQC1 algorithm returns an unbiased renormalized trace estimation with normal distribution $\mathcal{N}(\mathrm{tr}[U]/2^n, \Delta^2)$ of standard deviation Δ [17]. Of course, Δ can be reduced by a factor of $\sqrt{K}$ (i.e., SQL) if the algorithm of Fig. 1 is repeated K times. In fact, this is the situation in NMR where repetition reduces the signal-to-noise ratio (SNR) at the SQL.

Consider now, for example, the typical case where an unknown external magnetic field interacts with the n qubits of

*boixo@unm.edu
†somma@lanl.gov

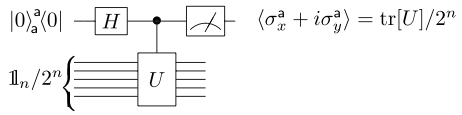

FIG. 1. DQC1 circuit for estimating the trace of a unitary U. The filled circle denotes that U acts on the probe when the state of the ancilla (control qubit) is $|1\rangle_{\mathsf{a}}$, and H is the Hadamard gate. $\langle\sigma_x^{\mathsf{a}}\rangle$ and $\langle\sigma_y^{\mathsf{a}}\rangle$ are the expectation values of the corresponding Pauli operators on the ancilla a.

the probe (Fig. 1), determining an interaction Hamiltonian of the form

$$H' = \theta \sum_{j=1}^{n} \sigma_z^j. \tag{2}$$

We seek to estimate θ. When the probe interacts with the field for time T, the n-qubit state is evolved by applying the corresponding (unitary) evolution operator $W'(T)=e^{-iH'T}$. Replacing ${}^cU$ by ${}^cW'(T)$ in Fig. 1, the final $(n+1)$-qubit state right before the measurement is

$$\rho_f = \frac{1}{2^{n+1}}[|0\rangle_{\mathsf{a}}\langle 0| \otimes \mathbb{1}_n + |1\rangle_{\mathsf{a}}\langle 0| \otimes W'(T) + |0\rangle_{\mathsf{a}}\langle 1| \otimes W'^{\dagger}(T) + |1\rangle_{\mathsf{a}}\langle 1| \otimes \mathbb{1}_n]. \tag{3}$$

Using the trace properties of the Pauli operators, we obtain $\langle\sigma_x^{\mathsf{a}}\rangle = \mathrm{tr}[\rho_f \sigma_x^{\mathsf{a}}] = [\cos(\theta T)]^n$. If $\Delta_x > 0$ denotes the standard deviation in the estimation of $\langle\sigma_x^{\mathsf{a}}\rangle$, a first approximation error formula determines

$$\Delta_\theta \approx \frac{\Delta_x}{|\partial\langle\sigma_x^{\mathsf{a}}\rangle/\partial\theta|}, \tag{4}$$

with Δ_θ the uncertainty in the estimation of θ.

Let $|+\rangle$ denote the single qubit state $(|0\rangle+|1\rangle)/\sqrt{2}$. The output signal of the previous algorithm is $|\langle +_1 \ldots +_{n/2}|W'(T)|+_1 \ldots +_{n/2}\rangle|^2$, which is the probability of measuring $|+_1 \ldots +_{n/2}\rangle$ after its evolution under $W'(T)$. Thus, the output precision of Eq. (4) is upper bounded by $\Delta_\theta = O[1/(\sqrt{n}T)]$, and the Heisenberg limit is not achieved when scaling n [18]. Nevertheless, for fixed (small) n, we obtain $\Delta_\theta = O(1/T)$, yielding the QML.

The previous estimation method has some important disadvantages. The first one concerns the use of the controlled ${}^cW'(T)$ operation which, due to technological difficulties or to the nature of the problem, may be impossible. In fact, this is the case in reference frame alignment, as we discuss in Sec. IV. Second, it is clear that $[\cos(\theta T)]^n$ approaches 0 exponentially with n for $T\theta \neq p\pi$, which is usually the case, as θ is unknown. As a consequence, the SNR of the outcome is weakened, especially for $n \gg 1$.

In this paper we propose a different method to perform multiparameter estimation that achieves the QML scaling $O(1/T)$. Interestingly, our method focuses on the evolution of observables such as tensor products of Pauli operators (i.e., we work in the Heisenberg picture), rather than the evolution of the state of the probe itself. For this reason, in Sec. II A we start by giving a brief description of Hamiltonian evolution in terms of Lie algebras. We then present an adaptive Bayesian estimation method to estimate single parameters at the QML with DQC1 (Sec. II C). Moreover, in Sec. III we show that, by applying dynamical-decoupling techniques and different Suzuki-Trotter approximations, multiparameter estimation can also be performed with DQC1. Here, we deduce that when the amount of short-time evolutions is considered a resource, the QML is asymptotically reached in the order of the approximation. In Sec. IV we discuss the particular example of reference frame alignment and show that to estimate the Euler angles dynamical-decoupling techniques are not required. In Sec. V we discuss the reasons why DQC1 allows us to reach the fundamental quantum limit in some cases, even though this model is less powerful than standard quantum computation. Finally, we present the conclusions in Sec. VI.

In the following, we ignore the effects of decoherence in our quantum algorithms and we assume that all experimental parameters can be controlled with arbitrary precision.

II. SINGLE-PARAMETER ESTIMATION

When the n-qubit probe interaction can be described by a Hamiltonian $H=\theta H_0$, single-parameter estimation aims to return an estimate $\hat{\theta}$ of the unknown θ at the highest precision possible, for some given amount of resources. For fixed n and $\|H_0\|$, our main resource is determined by the total evolution time under H. Let $W(T)=e^{-i\theta H_0 T}$ be the (unitary) evolution operator induced by H, during a time interval T. Clearly, if $W(T)$ acts non-trivially on some operator O, information about θ can be gained by computing $h(\theta T) = \mathrm{tr}[W^{\dagger}(T)OW(T)O]/2^n \in \mathbb{C}$ for different values of T. The form of $h(\theta T)$ can be obtained through the representation theory of Lie algebras (Sec. II A). Contrary to the example given in Sec. I, $h(\theta T)$ can be estimated using DQC1 without controlling the operation $W(T)$ (Sec. II B). Assuming that the accuracy in the estimation of $h(\theta T)$ remains constant regardless of T, an adaptive Bayesian estimator that returns θ at accuracy $O(1/T)$ (i.e., the QML) can be built in some cases of interest. Furthermore, to avoid signal loss due to possible large values of n, we choose the operator O such that $h(\theta T)$ does not depend on n. These points are studied and explained in more detail below.

A. The Heisenberg picture

If $X=X^{\dagger}$ is an observable acting on n qubits, we define (Heisenberg picture)

$$X(T) = W^{\dagger}(T)XW(T), \tag{5}$$

with $W(T)=e^{-i\theta H_0 T}$ as above. Assume now that we are given two observables H_1 and H_2 such that, together with H_0, they generate an su(2) Lie algebra. We obtain (see Appendix A)

$$H_1(T) = W^{\dagger}(T)H_1W(T) = \cos(2\theta T)H_1 - \sin(2\theta T)H_2. \tag{6}$$

If the operators are Schmidt pseudo-orthogonal $(\mathrm{tr}[H_i H_j] = d\delta_{ij})$ we have

$$\cos(2\theta T) = \mathrm{tr}[H_1(T)H_1]/d, \tag{7}$$

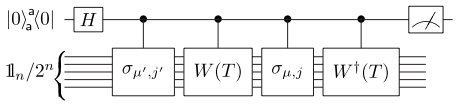

FIG. 2. DQC1 circuit for the estimation of $\mathrm{tr}[W^\dagger(T)\sigma_{\mu,j}W(T)\sigma_{\mu',j'}]/2^n \equiv \langle\sigma_x^a\rangle$. Note that $\langle\sigma_y^a\rangle=0$ in this case. Since every unitary is controlled by a (filled circles), including the evolution operator $W(T)$, the execution of this algorithm may be unfeasible due to the nature of the problem.

$$\sin(2\theta T) = -\mathrm{tr}[H_1(T)H_2]/d. \tag{8}$$

With no loss of generality we expand $H_j=\Sigma_{\mu=1}^{L} e^{\mu,j}\sigma_{\mu,j}$, where $\sigma_{\mu,j}=\sigma_{\mu,j}^1\otimes\cdots\otimes\sigma_{\mu,j}^n$ are tensor products of Pauli operators (henceforth simply called Pauli products), $\sigma_{\mu,j}^j \in \{\mathbb{1},\sigma_x,\sigma_y,\sigma_z\}$, and $e^{\mu,j}\in\mathbb{R}$ are known coefficients [19]. We obtain

$$\cos(2\theta T) = \sum_{\mu,\mu'} e^{\mu,1}e^{\mu',1}\mathrm{tr}[W^\dagger(T)\sigma_{\mu,1}W(T)\sigma_{\mu',1}]/d, \tag{9}$$

$$\sin(2\theta T) = -\sum_{\mu,\mu'} e^{\mu,1}e^{\mu',2}\mathrm{tr}[W^\dagger(T)\sigma_{\mu,1}W(T)\sigma_{\mu',2}]/d. \tag{10}$$

That is, $\cos(2\theta T)$ and $\sin(2\theta T)$ can be estimated at a fixed precision by L^2 runs of the circuit of Fig. 1. Each run returns an estimate of $\mathrm{tr}[U_{\mu,j;\mu',j'}]/2^n$, with unitary $U_{\mu,j;\mu',j'}=W^\dagger(T)\sigma_{\mu,j}W(T)\sigma_{\mu',j'}$. In Sec. II C we show how to estimate θ at the QML from the estimation of $\cos(2\theta T)$ or $\sin(2\theta T)$, for different values of T. Although we are mainly interested in the fundamental scaling achieved by increasing T, it is worth noting that the scaling with L^2 can be largely reduced in some cases of interest (see Appendix A).

B. DQC1 circuits

We now show how to avoid the need of controlled ${}^cW(T)$ operations, when estimating parameters with DQC1. This is of great importance since ${}^cW(T)$ may not be available for our use. To show this, we focus on the estimation of $\mathrm{tr}[W^\dagger(T)\sigma_{\mu,j}W(T)\sigma_{\mu',j'}]$, as required by Eqs. (9) and (10). The circuit that accomplishes this task is shown in Fig. 2. Here, the ${}^cU$ operation of Fig. 1 has been replaced by ${}^cW^\dagger(T){}^c\sigma_{\mu,j}{}^cW(T){}^c\sigma_{\mu',j'}$, where each operation is controlled by the ancilla a. Nevertheless, note that one can accomplish the same task even if the action of the operators $W(T)$ and $W^\dagger(T)$ is not controlled [20]. That is,

$${}^cW^\dagger(T){}^c\sigma_{\mu,j}{}^cW(T){}^c\sigma_{\mu',j'} \equiv W^\dagger(T){}^c\sigma_{\mu,j}W(T){}^c\sigma_{\mu',j'}, \tag{11}$$

which is clearly the identity operator when a is in $|0\rangle_a$. In Fig. 3 we show a simplified circuit that allows us to compute the above trace. The last operation $W^\dagger(T)$ is not included as it does not alter the measurement outcome and it may not be an available resource. Thus, the circuit can be implemented using (known) elementary gates and the available time-evolution operator $W(T)$ only.

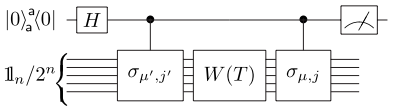

FIG. 3. Simplified version of the DQC1 circuit of Fig. 2. The operators ${}^cW(T)$ and $W^\dagger(T)$ are avoided. The estimation of $\cos(2\theta T)$ or $\sin(2\theta T)$ [Eqs. (9) and (10)] is performed using available resources only.

The next step is to detail a strategy to estimate θ from $\cos(2\theta T)$ or $\sin(2\theta T)$, and to characterize the associated error and resources needed. For this reason, we first make an assumption on the standard deviation Δ of the output returned by the DQC1 circuit of Fig. 3. Obviously, the smaller Δ is, the better the precision of the resulting estimation. We consider $\Delta \ll 1$, which can always be achieved by simple repetition of the computation. In some cases, such as liquid-state NMR quantum computation, where a vast amount of molecules contribute to the output signal, a $\Delta \ll 1$ could be achieved in a single run. Our estimation procedure should take advantage of this property by going beyond just performing a bit by bit estimation of θ, as it is done in several pure-state phase estimation techniques that involve strong (projective) measurements [5,7]. Moreover, we assume that Δ remains approximately constant in a certain region of values of $2\theta T$. This is a consequence of weak measurements [21]. In view of the central limit theorem, we assume that the measurement outcome is distributed according to $\mathcal{N}(\cos(2\theta T),\Delta^2)$ or $\mathcal{N}(\sin(2\theta T),\Delta^2)$(see Sec. I and Ref. [17]).

C. Adaptive Bayesian estimation

In Bayesian estimation, a parameter α to be estimated is considered to be a variable with an associated (known) probability distribution $f(\alpha)$ (i.e., the *prior distribution*). The prior distribution formalizes the experimenter's *state of belief* about α. It is the job of the experimenter to gain access to a sample of data $\{x_1,\ldots,x_K\}$ whose distribution $f(x_K,\ldots,x_1|\alpha)$ depends on α. Thus, the joint distribution of $\{x_1,\ldots,x_K\}$ and α is

$$f(x_K,\ldots,x_1,\alpha) = f(x_K,\ldots,x_1|\alpha)f(\alpha). \tag{12}$$

After observing a set of measurement outcomes $\{x_1,\ldots,x_K\}$, this information is used to obtain a *posterior distribution* $f(\alpha|x_K,\ldots,x_1)$ that corresponds to the experimenter's updated state of belief about the unknown α. This update is done using Bayes' rule

$$f(\alpha|x_K,\ldots,x_1) = \frac{f(x_K,\ldots,x_1,\alpha)}{f(x_K,\ldots,x_1)}, \tag{13}$$

where *the marginal sampling* distribution $f(x_K,\ldots,x_1)$ can be calculated from the joint distribution as

$$f(x_K,\ldots,x_1) = \int f(x_K,\ldots,x_1,\alpha)d\alpha. \tag{14}$$

In standard Bayesian estimation the probability of observing an *i.i.d.* sample $\{x_1,\ldots,x_K\}$ is determined by the total

sampling distribution $f(x_K,\ldots,x_1|\alpha)=f(x_K|\alpha)\cdots f(x_1|\alpha)$. In *adaptive Bayesian estimation* the outcome of the first measurement x_1 can be used to control the sampling distribution of the second measurement, $f(x_2|x_1,\alpha)$ (see below). In general, the sampling distribution of the lth measurement outcome can be conditioned by $\{x_1,\ldots,x_{l-1}\}$. At the end, $f(x_K,\ldots,x_1|\alpha)=f(x_K|x_{K-1},\ldots,x_1,\alpha)\cdots f(x_1|\alpha)$.

The Bayes' risk quantifies the expected penalty to be paid when using a particular estimator $\hat{\alpha}(x_K,\ldots,x_1)$ of α, and a given *cost function*. It is common to search for an $\hat{\alpha}$ that minimizes the Bayes' risk with a quadratic cost function, given by

$$\int (\alpha-\hat{\alpha}(x_K,\ldots,x_1))^2 f(\alpha|x_K,\ldots,x_1)d\alpha. \tag{15}$$

This risk is just the variance of the estimator and is minimized by the expectation value of the posterior distribution, giving the optimal estimator

$$\hat{\alpha}(x_K,\ldots,x_1)=\int \alpha f(\alpha|x_K,\ldots,x_1)d\alpha. \tag{16}$$

While a more detailed explanation of the adaptive Bayesian procedure is given in Appendix B, in the following we present a generic step of the estimation. Denote by $\Delta\ll 1$ the output precision of DQC1 when measuring $\cos(2\theta T)$, which may actually involve many $(L^2>1)$ different runs of the circuit of Fig. 3. Assume that, from l previous estimations, we have obtained an estimator $\hat{\theta}_l$ of θ such that, for known evolution time T_l and integer p_l,

$$2\hat{\theta}_l T_l \approx \pi/2+2p_l\pi. \tag{17}$$

Furthermore, assume that the 95% confidence interval for the estimation is

$$\hat{\theta}_l-1.96\Delta_l/(2T_l)\le\theta\le\hat{\theta}_l+1.96\Delta_l/(2T_l), \tag{18}$$

with $\Delta_l<\Delta$. Next, we show how to *zoom in* on $\hat{\theta}_l$ to obtain an estimator $\hat{\theta}_{l+1}$, so Eqs. (17) and (18) are still satisfied when replacing $l\to l+1$.

To do this, we first find T_{l+1} such that

$$2\hat{\theta}_l T_{l+1}=\pi/2+2p_{l+1}\pi, \tag{19}$$

with $p_{l+1}>p_l$ integer. The $(l+1)$th measurement returns x_{l+1}, an estimate of $\cos(\alpha_{l+1})$, with $\alpha_{l+1}=2\theta T_{l+1}$. This is done by running the algorithm of Fig. 3 with $T=T_{l+1}$. We chose p_{l+1} close enough to p_l such that, for $\alpha'_{l+1}=\alpha_{l+1}-(\pi/2+2p_{l+1}\pi)$, we approximate $\cos\alpha_{l+1}\approx-\alpha'_{l+1}$ (see Appendix B). Using Bayes rule and the joint distribution for the $(l+1)$th measurement, we obtain

$$\begin{aligned} f(x_{l+1},\alpha_{l+1}|x_l,\ldots,x_1) &= f(x_{l+1}|\alpha_{l+1},x_l,\ldots,x_1)f(\alpha_{l+1}|x_l,\ldots,x_1) \\ &\approx \frac{1}{\sqrt{2\pi}\Delta}\exp[-(x_{l+1}+\alpha'_{l+1})^2/2\Delta^2] \\ &\quad\times\frac{1}{\sqrt{2\pi}a_{l+1}\Delta_l}\exp[-(\alpha'_{l+1})^2/2(a_{l+1})^2\Delta_l^2], \end{aligned} \tag{20}$$

with $a_{l+1}=T_{l+1}/T_l$. Equation (20) determines the posterior distribution $f(\alpha_{l+1}|x_{l+1},x_l,\ldots,x_1)=f(x_{l+1},\alpha_{l+1}|x_l,\ldots,x_1)/f(x_{l+1}|x_l,\ldots,x_1)$. This adaptive procedure returns a new estimator $\hat{\theta}_{l+1}$, and standard deviation $\Delta_{l+1}/(2T_{l+1})$, with $\Delta_{l+1}<\Delta$, satisfying

$$\hat{\theta}_{l+1}-1.96\Delta_{l+1}/(2T_{l+1})\le\theta\le\hat{\theta}_{l+1}+1.96\Delta_{l+1}/(2T_{l+1}). \tag{21}$$

The total number K of estimations is chosen such that the final standard deviation is reduced below the desired precision Δ_θ. A sufficient condition is $\Delta/(2T_K)\le\Delta_\theta$. The fact that the standard deviation is reduced by T_l at each step [Eq. (21)] guarantees that the QML is achieved (Appendix B).

Remarkably, the confidence level of estimating the mean with error ϵ in our algorithm increases exponentially as $1-e^{-C(\epsilon/\tau)^2)}$, with $C>0$ and τ the corresponding standard deviation. This is clearly an advantage with respect to the standard pure-state phase estimation algorithm, where the confidence increases as $1-O(1/\epsilon)$ [12].

D. Black-box estimation: discrete time evolution

Imagine now that, instead of being able to evolve under the action of H for any period of time T, we are given a black box whose action is to perform the unitary operation $W_B=e^{-iH}$. As in the previous case, $H=\theta H_0$. That is, we are only allowed to evolve under H for a discrete time by simple concatenation of W_B's. This condition restricts the set of accessible operations to elementary gates and operations of the form

$$\bar{W}(q)=\overbrace{W_B\cdots W_B}^{q\ \text{times}}, \tag{22}$$

only. We seek to estimate θ at the QML using a modification of the previous adaptive Bayesian method.

We now give the generic step for achieving the QML in the discrete time case (see Appendix C for the first step). We assume that $\hat{\theta}_l$ is the mean of the estimator obtained after the lth measurement, performed with $q_l\in\mathbb{N}^*$ uses of W_B. Because we seek to make estimations around $\pi/2$, we take

$$2\hat{\theta}_{l-1}q_l+2\phi_l=\pi/2+2p_l\pi,\quad \forall\, l, \tag{23}$$

with $\hat{\theta}_{l-1}$ the mean of the estimation in the $(l-1)$th measurement, $p_l\in\mathbb{N}^*$, and ϕ_l the phase compensation. In other words, the lth estimation was performed by implementing the algorithm of Fig. 3 with $W(T)\to W_a$ $(q_l,\phi_l)\equiv(W_B)^{q_l}e^{-i\phi_l H_0}$. Because we can always consider $\pi/2\ge\phi_l>-\pi/2$, the phase compensation is made with unit cost. We also assume that $\Sigma_l\le\Delta$ is the standard deviation of the estimator of $2\theta q_l$ [22].

For the $(l+1)$th measurement, we write $q_{l+1}=bq_l$ with small enough b [23]. The $(l+1)$th measurement returns y_{l+1}, an estimate of $\cos(\beta_{l+1})$, with $\beta_{l+1}=2\theta q_{l+1}+2\phi_{l+1}\approx\pi/2+2\pi p_{l+1}$. Thus, to obtain $f(\beta_{l+1}|y_{l+1},\ldots,y_1)$ in the adaptive Bayesian step, we approximate

$$f(y_{l+1}|y_l,\ldots,y_1,\beta_{l+1}) \approx \frac{1}{\sqrt{2\pi}\Delta}\exp[-(y_{l+1}+\beta'_{l+1})^2/2\Delta^2], \tag{24}$$

with

$$\beta'_{l+1} = \beta_{l+1} - (\pi/2 + 2p_{l+1}\pi). \tag{25}$$

Moreover, since

$$f(\beta_{l+1}|y_l,\ldots,y_1) = \frac{1}{\sqrt{2\pi}b\Sigma_l}\exp[-(\beta'_{l+1})^2/2b^2(\Sigma_l)^2], \tag{26}$$

the resulting distribution $f(\beta_{l+1}|y_{l+1},\ldots,y_1)$ is normal. This is a consequence of Bayes' rule. Its mean and standard deviation, obtained by combining the exponents appearing in Eqs. (24) and (26), are

$$\hat{\beta}'_{l+1} = -\frac{(b'_{l+1})^2}{1+(b'_{l+1})^2}y_{l+1}, \tag{27}$$

$$\Sigma_{l+1} = (b'_{l+1}/\sqrt{1+(b'_{l+1})^2})\Delta < \Delta, \tag{28}$$

with $b'_{l+1}=b\Sigma_l/\Delta<b$. Equations (27) and (28) guarantee the success of the induction method. These quantities determine the mean and standard deviation of the new (improved) estimator of θ as

$$\hat{\theta}_{l+1} = \frac{1}{2b^l}\left(\pi/2 + 2p_{l+1}\pi - 2\phi_{l+1} - \frac{(b'_{l+1})^2}{1+(b'_{l+1})^2}y_{l+1}\right),$$

$$\Sigma'_{l+1} = \frac{\Sigma_{l+1}}{2b^l} < \frac{\Delta}{2b^l}. \tag{29}$$

The similarity of these results and those obtained for the continuous time case [Eqs. (B20) and (B21)] is clear.

Summarizing, if Δ_θ denotes the desired precision in the parameter estimation, a sufficient condition is to choose the total number of measurements K such that the final precision satisfies $\Delta/(2b^{K-1}) \le \Delta_\theta$. Since the total number of uses of W_B is given by $1+\cdots+b^{K-1}=(b^K-1)/(b-1)=O[\Delta/(2\Delta_\theta)]$, the QML is also reached in this case.

III. MULTIPARAMETER ESTIMATION

In this section we consider a more general case where the unknown interaction with the n-qubit probe can be described by a Hamiltonian

$$H = \sum_{\nu=1}^{P} \theta^\nu \sigma_\nu. \tag{30}$$

Here, $\theta^\nu \in \mathbb{R}$ and σ_ν are Pauli products. Using the results of Sec. II C, we seek to estimate every (unknown) θ^ν such that the returned precision approaches the QML for a given amount of resources or evolution time. For this reason, we assume that a previous estimate of every parameter, with mean $\pi > \hat{\theta}_0^\nu > 0$, is known.

Using dynamical-decoupling techniques [24], the multiparameter estimation case can be converted into P single-parameter estimations. For simplicity, we consider first the case where there is a Pauli product σ such that (for some ν)

$$[\sigma_\nu, \sigma] = 0, \tag{31}$$

$$\{\sigma_{\nu'}, \sigma\} = 0 \quad \forall \ \nu' \neq \nu. \tag{32}$$

A decoupled interaction is

$$H_\nu = \theta^\nu \sigma_\nu = (H + \sigma H \sigma)/2. \tag{33}$$

(In general, methods such as the one just described can be used to decouple any H_ν from any H.) We define $S_\nu(T) = e^{-iH_\nu T}$ to be the corresponding evolution operator. If such an operator were to be an available resource, θ^ν could be estimated using the scheme of Sec. II C. To do this, we would have to replace $W(T) \to S_\nu(T)$ and $\sigma_{\mu,j} \to \sigma_1$ in the circuit of Fig. 3, with $\{\sigma_\nu, \sigma_1\}$ being Pauli products that generate an su(2) algebra.

We now show how to approximate $S_\nu(T)$ from accessible operations that include $W(T)$ and elementary gates only. To show this, we use a Suzuki-Trotter approximation [25]. Specifically, for $q = 1/\epsilon \in \mathbb{N}^*$, we decompose

$$S_\nu(T) = \overbrace{S_\nu(\epsilon T)\cdots S_\nu(\epsilon T)}^{q \text{ times}}.$$

If $\bar{S}_\nu(\epsilon T)$ denotes a pth order Suzuki-Trotter approximation to $S_\nu(\epsilon T)$, we have

$$\|S_\nu(\epsilon T) - \bar{S}_\nu(\epsilon T)\| = O[\|H\|^p(\epsilon T)^p], \tag{34}$$

with $\|.\|$ some operator norm (e.g., the largest eigenvalue). The error of the approximation is

$$\varepsilon/2 \equiv \|S_\nu(T) - \bar{S}_\nu(T)\| = O[\|H\|^p \epsilon^{p-1} T^p], \tag{35}$$

with $\bar{S}_\nu(T) = [\bar{S}_\nu(\epsilon T)]^q$. Equation (35) was obtained using Eq. (34), together with $\|A^q - B^q\| = \|(A-B)A^{q-1} + B(A-B)A^{q-2} + \cdots + B^{q-1}(A-B)\| \le q\|A-B\|$, for A and B unitaries.

We now show how to build $\bar{S}_\nu(T)$ out of available resources. In the simplest case (i.e., $p=2$) the evolution operator at short times factorizes as

$$\bar{S}_\nu(\epsilon T) = e^{-iH\epsilon T/2}\sigma e^{-iH\epsilon T/2}\sigma. \tag{36}$$

Then, $\bar{S}_\nu(T)$ can be implemented using H evolutions and σ gates only. Similarly, a second-order Suzuki-Trotter approximation (i.e., $p=3$) is given by

$$\bar{S}_\nu(\epsilon T) = e^{-iH\epsilon T/4}\sigma e^{-iH\epsilon T/2}\sigma e^{-iH\epsilon T/4}. \tag{37}$$

Higher order approximations can be constructed in a similar fashion [26], so they can always be implemented with accessible gates. The larger p is, the shorter the actions of H in each step. Thus, we require precise time control in our algorithms (measurements).

Replacing $W(T) \to \bar{S}_\nu(T)$ and $\sigma_{\mu,j} \to \sigma_1$ in the circuit of Fig. 3 allows us to estimate θ^ν. To show this, consider the

measurement output z_l obtained in the lth measurement. z_l will give us an estimator of the parameter θ^ν plus a correction

$$\cos(2\theta^\nu T_l) + \gamma(\epsilon, p, T_l, \vec{\theta}), \tag{38}$$

with $\vec{\theta} = \theta^1, \ldots, \theta^P$. The norm of $\gamma(\epsilon, p, T_l, \vec{\theta}) \in \mathbb{R}$ can be bounded above as

$$|\gamma(\epsilon, p, T_l, \vec{\theta})| = \frac{|\mathrm{tr}[S_\nu^\dagger(T_l)\sigma_1 S_\nu(T_l)\sigma_1 - \bar{S}_\nu^\dagger(T_l)\sigma_1 \bar{S}_\nu(T_l)\sigma_1]|}{2^n}$$

$$\leq \frac{\mathrm{tr}\|S_\nu^\dagger(T_l)\sigma_1 S_\nu(T_l)\sigma_1 - \bar{S}_\nu^\dagger(T_l)\sigma_1 \bar{S}_\nu(T_l)\sigma_1\|}{2^n} \leq \varepsilon. \tag{39}$$

We have used Eq. (35) to obtain Eq. (39). Because $\gamma(\epsilon, p, T_l, \vec{\theta})$ depends on the θ^ν's, we consider it a variable with an associated (worst-case scenario) prior distribution given by $f(\gamma) \equiv \mathcal{N}(0, \Delta_\gamma)$, with $\Delta_\gamma = O(\Delta\varepsilon)$ [27]. The net effect in the adaptive Bayesian procedure is that now, the joint distribution determines a marginal distribution after $\gamma(\epsilon, p, T_l, \vec{\theta})$ is integrated out. More precisely, for the lth estimation, we have

$$f(x_l|x_{l-1}, \ldots, x_1, \alpha_l^\nu) = \int f(x_l|x_{l-1}, \ldots, x_1, \alpha_l^\nu, \gamma) f(\gamma) d\gamma, \tag{40}$$

with $\alpha_l^\nu = 2\theta^\nu T_l$. Making a linear approximation in the cosine function, we obtain

$$f(x_l|x_{l-1}, \ldots, x_1, \alpha_l^\nu) \approx \mathcal{N}(\alpha_1'^\nu, (\Delta')^2), \tag{41}$$

with $\alpha_1'^\nu = \alpha_l^\nu - (\pi/2 + 2p_l\pi)$ (see Sec. II C), and updated variance

$$(\Delta')^2 = \Delta^2 + \Delta_\gamma^2. \tag{42}$$

Thus, we can use the adaptive method of Sec. II C to estimate every parameter θ^ν at the QML, by replacing $\Delta \to \Delta'$.

Notice that one could also implement an adaptive Bayesian approach to learn about $\gamma(\epsilon, p, T_l, \vec{\theta})$. In such a case, the distribution $f(\gamma)$ would need to be updated after each measurement based on the previous measurement outcomes. Nevertheless, we did not consider this approach in the above discussion because we assumed that the Suzuki-Trotter approximation used is good enough for our purposes (i.e., $\Delta_\gamma \ll 1$).

When the resource of interest is the number of calls to $\bar{S}_\nu(\epsilon T)$, the amount of resources to reach a precision Δ_{θ^ν} changes. For this reason, consider the total evolution time $T_t = O[(\Delta + \Delta_\gamma)/\Delta_{\theta^\nu})]$, with $\Delta_\gamma = O(\Delta \|H\|^p \epsilon^{p-1} T_t^p)$. Assume that we want to keep T_t constant, regardless of p. Then, $q = 1/\epsilon = O[(T_t)^{(p/p-1)}]$. That is, $O[1/(\Delta_{\theta^\nu})^{(p/p-1)}]$ actions of $\bar{S}_\nu(\epsilon T)$ are required to attain Δ_{θ^ν}, and the QML is asymptotically reached in p.

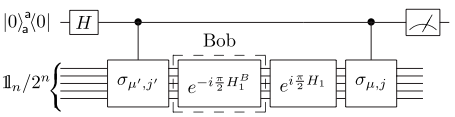

FIG. 4. Elementary step of the frame alignment protocol. The n-qubit probe, whose state remains completely mixed, is exchanged between Alice and Bob. The Pauli products $\sigma_{\mu,j}$ and $\sigma_{\mu',j'}$ appear in the decomposition of H_1 [see Eqs (9) and (10)].

IV. REFERENCE FRAME ALIGNMENT

Imagine that two distant parties, Alice and Bob, suffer some frame misalignment, which is manifest in the way they characterize their operations on equivalent quantum systems. This might be a result of not sharing synchronized clocks or having different spatial reference frames. Aligning both frames requires the exchange of physical systems carrying "unspeakable" information. This information is encoded as frame-dependent parameters that need to be estimated [28]. The resource that limits the quality of this estimation is the number of systems interchanged between Alice and Bob. We will show that Alice and Bob can align their frames within the DQC1 model at the QML. In particular, we propose a modification of a pure-state protocol for frame synchronization [5] based on repeated coherent exchanges of the n-qubit probe only. Remarkably, the state of the probe remains completely mixed and separable from the ancilla at every step [16]. Moreover, in our protocol, Bob never accesses the ancilla that Alice measures.

We consider first the case where the effect of the frame misalignment is uni-parametric. That is, Alice's and Bob's description of operators acting on equivalent Hilbert spaces is known to differ by a unitary transformation $V_\theta = e^{-i\theta H_0}$, with θ unknown. More explicitly, for some operator O we have

$$O^B = V_\theta^\dagger O V_\theta = e^{i\theta H_0} O e^{-i\theta H_0}, \tag{43}$$

where we used the superscript B to denote the action of O in Bob's frame. Since H_0 is known, we assume that we can find pseudo-orthogonal observables H_1 and H_2 such that they form an su(2) algebra [i.e., Eq. (A5) is satisfied in Alice's frame].

An elementary step of the protocol starts with Alice sending the n-qubit probe to Bob. Subsequently, Bob applies the operation $e^{-i\pi H_1^B/2}$ and returns the probe to Alice. Finally, Alice applies the adjoint operation $e^{i\pi H_1/2}$. The resulting operation on the state of the probe in Alice's frame is

$$e^{i\pi H_1/2} e^{-i\pi H_1^B/2} = e^{i\pi H_1/2} e^{i\theta H_0} e^{-i\pi H_1/2} e^{-i\theta H_0} = e^{-2i\theta H_0} \equiv V_{2\theta}. \tag{44}$$

Equation (44) can be obtained by working in any faithful representation of su(2). The global action of each step can be seen as a "black box" implementation of $V_{2\theta}$, whose parameter we want to estimate. Using the results of Sec. II D, the circuit of Fig. 3 can be used to make a first estimate of θ if $W(T)$ is replaced by $V_{2\theta}$ (see Fig. 4). To zoom in on previous estimations requires instead the implementation of the unitary $V_{2m\theta} = (V_{2\theta})^m$, with $m \in \mathbb{N}^*$. This can be done by simple

concatenation of elementary steps, requiring m coherent exchanges of the probe.

Another case of interest is the alignment of spatial reference frames [5]. Let us assume that Bob's operators are related to Alice's through a rotation

$$R = e^{-i\psi H_2} e^{-i\theta H_1} e^{-i\phi H_2}, \tag{45}$$

where $\{\phi, \theta, \psi\}$ are Euler angles and $\{H_1, H_2\}$ generate an su(2) algebra. Consider a synchronization protocol with this elementary step: Alice sends the probe to Bob; Bob applies the operation $e^{-i\pi H_2^B/2}$ and returns the probe to Alice; Alice applies the operation $e^{i\pi H_2/2}$. The effective rotation of this step, in Alice's frame, is given by

$$V' = e^{i\pi H_2/2} e^{-i\pi H_2^B/2} = e^{i\pi H_2/2} R^\dagger e^{-i\pi H_2/2} R. \tag{46}$$

The action of V' on H_2 implies that

$$\mathrm{tr}[V'^\dagger H_2 V' H_2] = d\cos(2\theta), \tag{47}$$

with d some normalization constant [Eq. (9)]. Again, the previous derivation can be carried out in any faithful representation of su(2). Thus, a first estimation of the Euler angle θ can be performed by using the circuit of Fig. 3, replacing $W(T) \to V'$, and with $\sigma_{\mu,j}$ being the Pauli products appearing in the expansion of H_2. Furthermore, since

$$\mathrm{tr}[(V'^\dagger)^m H_2 (V')^m H_2] = d\cos(2m\theta), \tag{48}$$

we can zoom in on the previous estimation by applying V', $m > 1$ times, and using the results of Sec. II D. In this case, m coherent transports of the probe between Alice and Bob are required.

Notice that in this estimation procedure, the action of relevant operations need not be controlled by the ancilla. Finally, the other Euler angles can be estimated in a similar way if Alice and Bob agree to apply other analogous operations.

V. PARAMETER ESTIMATION VS. GROVER'S SEARCH ALGORITHM IN DQC1

So far we have shown that to reach certain precision in the estimation of a parameter, DQC1 requires less resources than other methods. One may wonder if such a quantum speedup can also be attained in problems such as searching for a particular property in a given set (i.e., search problem), which is the case for pure-state quantum algorithms. However, in Ref. [14] the authors proved that DQC1 is strictly less powerful than standard quantum computation in the oracle setting. This implies that DQC1 does not provide a quadratic quantum speedup in the search problem as the one given by Grover's algorithm [29]. To show this, consider the situation in which we are given a black box that implements either the unitary $U_B = e^{i\theta|S\rangle\langle S|}$, with $\theta \neq 0$, or $U_B = \mathbb{1}$, over the state of the n qubits in the probe. Here, $|S\rangle$ encodes the solution to our search problem. We want to determine the existence of a solution; that is, we want to specify if the action of U_B is trivial (i.e., $\mathbb{1}$) or not (i.e., $e^{i\theta|S\rangle\langle S|}$). In fact, this is a phase estimation problem in which we have to decide whether the phase is 0 or θ.

In general, the output of a DQC1 algorithm is given by

$$\langle \sigma_z^{\mathsf{a}} \rangle = \mathrm{tr}[\rho_f \sigma_z^{\mathsf{a}}], \tag{49}$$

with ρ_f the ancilla-probe state right before the measurement. With no loss of generality,

$$\rho_f = W_Q U_B \cdots W_1 U_B W_0 \rho_0 W_0^\dagger U_B^\dagger W_1^\dagger \cdots U_B^\dagger W_Q^\dagger, \tag{50}$$

with $\rho_0 = (|0\rangle_{\mathsf{a}}\langle 0| \otimes \mathbb{1}_n)/2^n$ being the initial state, and Q the number of calls to U_B. Since $|0\rangle_{\mathsf{a}}\langle 0| = (\mathbb{1}_{\mathsf{a}} - \sigma_z^{\mathsf{a}})/2$, we have

$$\langle \sigma_z^{\mathsf{a}} \rangle = \mathrm{tr}[W_Q U_B \cdots U_B W_0 \sigma_z^{\mathsf{a}} W_0^\dagger U_B^\dagger \cdots U_B^\dagger W^\dagger Q \sigma_z^{\mathsf{a}}]/2^{n+1}. \tag{51}$$

Following the proof in Ref. [14], we obtain

$$\left| \langle \sigma_z^{\mathsf{a}} \rangle|_{U_B=\mathbb{1}} - \langle \sigma_z^{\mathsf{a}} \rangle|_{U_B=e^{i\theta|S\rangle\langle S|}} \right| \leq 4Q/2^{n+1}. \tag{52}$$

Since DQC1 returns $\langle \sigma_z^{\mathsf{a}} \rangle$ at accuracy Δ, Q must be exponentially large in n or the algorithm needs to be executed exponentially many times to determine whether there is a solution or not. That is, if J is the amount of times that the algorithm is executed, we would expect that the precision in the estimation scales as $\Delta' = O(\Delta/\sqrt{J})$. To solve the problem, it is necessary (but not sufficient) to choose J and Q such that

$$4Q/2^{n+1} > \Delta', \tag{53}$$

requiring $N \equiv JQ$ uses of the black box. Thus,

$$N > O(2^n). \tag{54}$$

If $Q = O(\sqrt{2^n})$ as in Grover's algorithm, we need $J = O(\sqrt{2^n})$ to satisfy Eq. (54), and no quantum speedup with respect to the classical counterpart is obtained in this case.

The reason for the existence of a quantum-speedup in parameter estimation is that the unitary operators considered act nontrivially in a large-dimensional subspace of the corresponding 2^n-dimensional Hilbert space. In this case, the output signal obtained by executing the DQC1 circuits enables us to distinguish between different unitaries and to estimate the parameter. In the search problem, however, the operator U_B is very close to the identity operator $\mathbb{1}$ in that it only affects the state $|S\rangle$. Therefore, its action over highly mixed states is almost trivial. This is not the case in pure-state algorithms where one usually works in a two-dimensional Hilbert space spanned by the states $|S\rangle$ and $|S^\perp\rangle$, with $\langle S|S^\perp\rangle = 0$ [7,29].

VI. CONCLUSIONS

Parameter estimation at the quantum metrology limit could have a wide range of applications in metrology [10,11]. Further, single-parameter estimation is related to phase estimation, a cornerstone of quantum computation. Mixed-state quantum computation, as formalized in the DQC1 model, is interesting both from a theoretical and from a practical point view [14–16]. We have shown that, under fairly general conditions, it is possible to perform parameter estimation at the quantum metrology limit within the DQC1 model. These conditions are presented using Lie algebraic methods. The algorithm proceeds using adaptive Bayesian

estimation. In each step we zoom in on the previously estimated parameter, while ensuring that the increased variance remains below certain bounds. A measurement reduces such a variance to the previous value and this procedure is repeated until the desired precision in the estimation is reached. In short, the procedure ensures that the phase is kept in the same region, with almost constant variance, but increasing winding number.

The adaptive estimation is clearer when the time of the evolution under the unknown Hamiltonian (parameter) can be controlled at will. When lacking this freedom, as in the case of black box estimation, the algorithm for continuous time can be amended with some straightforward modifications. Moreover, to perform multiparameter estimation, implementation of dynamical-decoupling techniques reduces the problem to several single-parameter estimation procedures. Yet, these techniques are not necessary in some simple cases such as frame alignment between two parties, Alice and Bob. For spatial frame alignment, the Euler angles can be estimated at the quantum metrology limit with Bob having access to the completely mixed, separable, state of the probe only. Surprisingly, although precise estimation is intimately related to the quantum speedup given by Grover's search algorithm, the latter cannot be performed efficiently within the DQC1 model.

ACKNOWLEDGMENTS

We are thankful to E. Bagan, H. Barnum, C.M. Caves, E. Knill, and L. Viola for interesting discussions. This work was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396 and partially supported by ONR Grant No. N00014-07-1-0304.

APPENDIX A: LIE ALGEBRA REPRESENTATIONS

For $X(T)=W^{\dagger}(T)XW(T)$ we can also write $X(T)=X+iT[H,X]-T^2/2[H,[H,X]]+\cdots$, with $[Y,X]=YX-XY$. That is, $X(T)$ is a linear combination of observables belonging to the Lie algebra h generated by H and X [30]. In general, $h\equiv\{O_1,\ldots,O_M\}$ is an M-dimensional (real) semisimple Lie algebra, with $O_j=O_j^{\dagger}$ and $\mathrm{tr}[O_iO_j]=d\delta_{ij}$ ($d\in\mathbb{R}$). A faithful representation of h is the mapping $O_j\rightarrow\bar{O}_j$, with $\bar{O}_j=\bar{O}_j^{\dagger}$ being $(m\times m)$-dimensional matrices that satisfy

$$[O_i,O_j]=\sum_k f_{ij}^k O_k \leftrightarrow [\bar{O}_i,\bar{O}_j]=\sum_k f_{ij}^k \bar{O}_k, \tag{A1}$$

$$\mathrm{tr}[\bar{O}_i\bar{O}_j]=\bar{d}\,\delta_{ij}, \tag{A2}$$

for some $\bar{d}\in\mathbb{R}$. The coefficients f_{ij}^k are the so-called structure constants of h. Then, $X(T)=\Sigma_k c^k O_k$, $c^k\in\mathbb{R}$, and

$$\bar{X}(T)=\bar{W}^{\dagger}(T)\bar{X}\bar{W}(T)=\sum_k c^k\bar{O}_k, \tag{A3}$$

with $\bar{W}(T)=e^{-i\bar{H}T}$. The c^k's depend only on the f_{ij}^k's. Equation (A3) implies that if two different sets of (linearly independent) matrices have the same commutation relations, the coefficients c^k can be determined by working in either matrix representation:

$$c^k=\mathrm{tr}[X(T)O_k]/d=\mathrm{tr}[\bar{X}(T)\bar{O}_k]/\bar{d}. \tag{A4}$$

Because $\{H_0,H_1,H_2\}$ span a su(2) Lie algebra, they satisfy

$$[H_j,H_k]=2i\epsilon_{jkl}H_l, \tag{A5}$$

with $j,k,l\in\{0,1,2\}$ and ϵ_{jkl} the totally antisymmetric symbol. The su(2) Lie algebra can be built upon (2×2)-dimensional Hermitian, traceless, matrices (i.e., Pauli spin-1/2 operators). This allow us to carry out the calculation of Eq. (A3) in a low-dimensional representation.

To provide an example in which the L^2 trace estimation can be reduced [Eqs. (9) and (10)], consider again the situation where $H=\theta\Sigma_{j=1}^n\sigma_z^j$ [Eq. (2)]. Then, for example,

$$\sigma_x^1(T)=e^{iHT}\sigma_x^1e^{-iHT}=\cos(2\theta T)\sigma_x^1-\sin(2\theta T)\sigma_y^1, \tag{A6}$$

as $[\sigma_x^1,\sigma_z^j]=0\,\forall\, j>1$. That is, $\cos(2\theta T)$ can be estimated by computing the renormalized trace of the unitary $U=\sigma_x^1(T)\sigma_x^1$ only. This situation can be generalized to the case when $H_0=\Sigma_{\mu=1}^L e^{\mu,0}\sigma_{\mu,0}$, if there is an $H_1=\sigma_1$, with σ_1 a Pauli product such that (for some μ)

$$\begin{aligned}&[\sigma_{\mu,0},\sigma_{\mu',0}]=0 \quad \forall\ \mu,\mu',\\ &\{\sigma_{\mu,0},\sigma_1\}=0,\\ &[\sigma_{\mu',0},\sigma_1]=0 \quad \forall\ \mu'\neq\mu.\end{aligned} \tag{A7}$$

It follows that

$$\begin{aligned}&[\sigma_{\mu,0},\sigma_1]=2i\sigma_2,\\ &[\sigma_{\mu',0},\sigma_2]=0 \quad \forall\ \mu'\neq\mu,\end{aligned} \tag{A8}$$

with $\sigma_2=-i(\sigma_{\mu,0}\sigma_1)$. Equation (A8) results from the Jacobi identity and the definition of σ_2. Thus, $\{\sigma_{\mu,0},\sigma_1,\sigma_2\}$ satisfies the su(2)-commutation relations, yielding

$$\cos(2\theta e^{\mu,0}T)=\mathrm{tr}[W^{\dagger}(T)\sigma_1W(T)\sigma_1]/2^n, \tag{A9}$$

$$\sin(2\theta e^{\mu,0}T)=\mathrm{tr}[W^{\dagger}(T)\sigma_1W(T)\sigma_2]/2^n. \tag{A10}$$

We estimate Eqs. (A9) and (A10) by executing the circuit of Fig. 1 with $U=W^{\dagger}(T)\sigma_1W(T)\sigma_1$ and $U=W^{\dagger}(T)\sigma_1W(T)\sigma_2$, respectively.

APPENDIX B: THE ADAPTIVE BAYESIAN ESTIMATION PROCEDURE

To reach the QML in our estimation procedure, we first assume an initial estimation of θ given by a prior distribution $\mathcal{N}(\hat{\theta}_0,\Delta')$. For simplicity, in the following we focus on the case where θ is determined from the estimation of $\cos(2\theta T)$ for different values of T, denoted as T_l. This is done using the DQC1 algorithm of Fig. 3, where the corresponding Pauli products $\sigma_{\mu,j}$ are determined by Eq. (9). Thus, x_l denotes an

estimate of $\cos(2\theta T_l)$ and we consider $\pi > \hat{\theta}_0 > 0$. Otherwise, the estimation of $\sin(2\theta T_l)$ is also required to determine, for example, the corresponding quadrant of θ.

To obtain x_1, we choose T_1 such that $2\hat{\theta}_0 T_1 = \pi/2$ [31]. This can be done with an upper bounded initial use of resources when $\hat{\theta}_0 \in [\chi, \pi)$, with $\chi > 0$. That is, $1/4 < T_1 \leq \pi/(4\chi)$. Nevertheless, if $\hat{\theta}_0 \ll 1$ (i.e., $T_1 \gg 1$), a similar analysis as the one carried out below can be performed by measuring $\sin(2\theta T_1')$ instead of $\cos(2\theta T_1)$, with $T_1' = O(1)$ [32]. Therefore, we take $\mathcal{N}(\pi/2, (c\,\Delta)^2)$as the prior distribution of $\alpha_1 = 2\theta T_1$. We define here Δ to be the output precision of DQC1 when measuring $\cos(2\theta T)$, which may actually involve many $(L^2 > 1)$ different runs of the circuit of Fig. 3. We assume $1 \gg c\Delta \gg \Delta$. The measurement outcome x_1 of the first measurement has then a sampling distribution given by $\mathcal{N}(\cos(\alpha_1), \Delta^2)$.

The joint distribution $f(x_1, \alpha_1) = f(x_1 | \alpha_1) f(\alpha_1)$ is

$$\begin{aligned} f(x_1, \alpha_1) &= \frac{1}{\sqrt{2\pi}\Delta} \exp[-(x_1 - \cos(\alpha_1))^2 / 2\Delta^2] \\ &\times \frac{1}{\sqrt{2\pi}c\Delta} \exp[-(\alpha_1 - \pi/2)^2 / 2c^2\Delta^2] \\ &\approx \frac{1}{\sqrt{2\pi}\Delta} \exp[-(x_1 + \alpha_1')^2 / 2\Delta^2] \\ &\times \frac{1}{\sqrt{2\pi}c\Delta} \exp[-(\alpha_1')^2 / 2c^2\Delta^2]. \end{aligned} \tag{B1}$$

Here, $\alpha_1' = \alpha_1 - \pi/2$ and, for simplicity, we approximated $\cos(\alpha_1)$ at first order by $-\alpha_1'$ so that the joint distribution is normal. The error in this approximation can be bounded above, with high confidence, as

$$|\cos(\alpha_1) - (-\alpha_1')| \leq \delta = (c'\Delta)^3/6 \ll 1, \tag{B2}$$

for some $c' \geq c$. For example, if choosing $c' = 1.96c$, Eq. (B2) determines a 95% credible interval for $\cos(\alpha_1)$. Such a confidence can be made exponentially close to 1 as c' increases. Of course this error can be avoided if no approximation is made and other analytical or numerical methods are used. Nevertheless, a linear approximation to the cosine is enough for our purposes, as it will yield the proper results [33]. Moreover, the error of the above approximation will be further corrected by subsequent measurements. This is a consequence of the adaptive method.

The following step is to update the information about α_1 (or θ) based on the measurement outcome x_1. The posterior distribution is $f(\alpha_1 | x_1) = f(x_1, \alpha_1)/f(x_1)$. Using Eq. (B1) this distribution can be shown to be

$$f(\alpha_1 | x_1) \approx \mathcal{N}(\hat{\alpha}_1, \Delta_1^2). \tag{B3}$$

$\hat{\alpha}_1$ and Δ_1 are determined from the exponent $E = -(x_1 + \alpha_1')^2/(2\Delta^2) - \alpha_1'^2/(2c^2\Delta^2)$ in Eq. (B1). We write

$$E = -\frac{\left(\alpha_1' + \frac{c^2}{1+c^2} x_1\right)^2}{2\Delta_1^2} + g(x_1), \tag{B4}$$

implying

$$\Delta_1 = \frac{c}{\sqrt{1+c^2}} \Delta < \Delta, \tag{B5}$$

$$\hat{\alpha}_1 = \pi/2 - \frac{c^2}{1+c^2} x_1. \tag{B6}$$

Summarizing, if $\hat{\theta}_1$ is our estimator of θ after the first measurement, we have

$$\hat{\theta}_1 = \frac{\hat{\alpha}_1}{2T_1} = \frac{1}{2T_1}\left(\pi/2 - \frac{c^2}{1+c^2} x_1\right), \tag{B7}$$

which is only a linear correction in x_1. Moreover, our knowledge about θ has increased such that

$$\hat{\theta}_1 - 1.96\Delta_1/(2T_1) \leq \theta \leq \hat{\theta}_1 + 1.96\Delta_1/(2T_1) \tag{B8}$$

is a 95% credible interval. If the desired output precision Δ_θ in the single-parameter estimation satisfies $\Delta_1/(2T_1) \leq \Delta_\theta$, the estimation procedure stops here. Otherwise, further measurements are required.

In the previous analysis we have neglected other values of α_1 mod 2π that would yield the same measurement outcome. This assumption introduces an extremely small error bounded above by

$$\mathrm{erfc}\left(\frac{2\pi}{c\Delta\sqrt{2}}\right), \tag{B9}$$

where erfc denotes the *complementary error function* of the normal distribution. Since $c\Delta \ll 1$, we have $\mathrm{erfc}(2\pi/(c\Delta\sqrt{2})) \approx 0$. For example, for the unrealistic case of $c\Delta = 1$, we have $\mathrm{erfc}(\sqrt{2}\pi) \approx 3.4\ 10^{-10}$. Thus, our assumption does not affect the final results of the estimation procedure.

To increase the precision we zoom in on the previously estimated θ, so a better estimate is attained. Since $x_1 = O(c\Delta) \ll 1$ [34], we have $|2\hat{\theta}_1 T_1 - \pi/2| \ll 1$. Accordingly, there exists a time period $T_2 > T_1$ such that

$$2\hat{\theta}_1 T_2 = \pi/2 + 2p_2\pi\ , \tag{B10}$$

with $p_2 \in \mathbb{N}^*$. T_2 denotes the evolution time in the second measurement (estimation). The main reason behind Eq. (B10) is that here, as in the first measurement, we seek to make a phase estimation around $\pi/2$ mod2π. In this region, the cosine function is more sensitive to variations in the phase. The second measurement returns x_2, an estimate of $\cos(\alpha_2)$, with $\alpha_2 = 2\theta T_2$.

Since $T_2 = a_2 T_1$, with $a_2 > 1$, we obtain $a_2 \approx (1 + 4p_2)$ [see Eqs. (B7) and (B10)]. The previously estimated θ has a variance of order $\Delta/(2T_1)$ [Eqs. (B5) and (B8)], so the variance of α_2 is of order $a_2\Delta$. To guarantee that this variance is similar to that of α_1 (first measurement), we choose a_2 as large as possible so that [35]

$$0 < c' - 4 < a_2 \leq c'. \tag{B11}$$

This implies that Eq. (B2) is still satisfied when replacing α_1 by α_2, and α_1' by $\alpha_2' = \alpha_2 - (\pi/2 + 2p_2\pi)$. As an example, consider the case when $c' \approx 10$. Therefore, we choose $a_2 \approx 9 \leq c'$, corresponding to $p_2 = 2$ in Eq. (B10).

After the measurement, the outcome x_2 is used to update our information about θ. Since $c'\Delta \ll 1$, $\alpha_2 \in [2p_2\pi, (2p_2+1)\pi]$ with large confidence. That is, we do not consider other values of $\alpha_2 \bmod 2\pi$ and we make the estimation in this region only. The joint distribution is now

$$\begin{aligned} f(x_2, \alpha_2 | x_1) &= f(x_2 | x_1, \alpha_2) f(\alpha_2 | x_1) \\ &\approx \frac{1}{\sqrt{2\pi}\Delta} \exp[-(x_2 + \alpha_2')^2 / 2\Delta^2] \\ &\times \frac{1}{\sqrt{2\pi} a_2 \Delta_1} \exp[-(\alpha_2')^2 / 2(a_2)^2 \Delta_1^2]. \end{aligned} \tag{B12}$$

$f(\alpha_2|x_1)$ has been determined using Eq. (B10). Thus, the posterior density distribution $f(\alpha_2|x_2, x_1) = f(x_2, \alpha_2|x_1)/f(x_2|x_1)$ is a normal $\mathcal{N}(\hat{\alpha}_2, (\Delta_2)^2)$, where $\hat{\alpha}_2$ and Δ_2 are determined by the exponent appearing in Eq. (B12). These are

$$\hat{\alpha}_2 = (\pi/2 + 2p_2\pi) - \frac{(a_2')^2}{1 + (a_2')^2} x_2, \tag{B13}$$

$$\Delta_2 = \frac{a_2'}{\sqrt{1 + (a_2')^2}} \Delta < \Delta, \tag{B14}$$

with

$$a_2' = a_2 \Delta_1 / \Delta < a_2. \tag{B15}$$

Summarizing, the estimator after the second measurement is

$$\hat{\theta}_2 = \frac{1}{2T_2} \left[(\pi/2 + 2p_2\pi) - \frac{(a_2')^2}{1 + (a_2')^2} x_2 \right], \tag{B16}$$

with a 95% credible interval

$$\hat{\theta}_2 - 1.96\Delta_2/(2T_2) \leq \theta \leq \hat{\theta}_2 + 1.96\Delta_2/(2T_2). \tag{B17}$$

The standard deviation in the estimation of θ has been reduced by a factor of order a_2 with respect to the one returned in the first measurement [Eq. (B8)].

If the desired output precision satisfies $\Delta_2/(2T_2) \leq \Delta_\theta$, the estimation stops here. Otherwise, we continue with the adaptive procedure. At each step l we find T_l such that

$$2\hat{\theta}_{l-1} T_l = \pi/2 + 2p_l\pi, \tag{B18}$$

where $\hat{\theta}_{l-1}$ is the Bayes' estimator determined by the previous measurement outcomes. Since $p_l > p_{l-1}$ are positive integers, we write $T_l = a_l T_{l-1}$, with $a_l \approx (1+4p_l)/(1+4p_{l-1})$. The lth measurement returns x_l, an estimate of $\cos(\alpha_l)$, with $\alpha_l = 2\theta T_l$. This is done by running the algorithm of Fig. 3 with $T = T_l$. To keep the variance of α_l at order $c'\Delta$ we choose a_l such that $c' - 4 \leq a_l \leq c'$ [35]. With this choice, Eq. (B2) is still satisfied when replacing α_1 by α_l, and α_1' by $\alpha_l' = \alpha_l - (\pi/2 + 2p_l\pi)$. Using Bayes' rule, and considering

$$\begin{aligned} f(x_l, \alpha_l | x_{l-1}, \ldots, x_1) &\approx \frac{1}{\sqrt{2\pi}\Delta} \exp[-(x_l + \alpha_l')^2 / 2\Delta^2] \\ &\times \frac{1}{\sqrt{2\pi} a_l \Delta_{l-1}} \exp[-(\alpha_l')^2 / 2(a_l)^2 \Delta_{l-1}^2], \end{aligned} \tag{B19}$$

we obtain for the lth estimation

$$2\hat{\theta}_l T_l = \pi/2 + 2p_l\pi - \frac{(a_l')^2}{1 + (a_l')^2} x_l, \tag{B20}$$

$$\frac{\Delta_l}{2T_l} = \frac{a_l'}{\sqrt{1 + (a_l')^2}} \frac{\Delta}{2T_l} < \frac{\Delta}{2T_l}, \tag{B21}$$

with

$$a_l' = a_l \Delta_l / \Delta < a_l. \tag{B22}$$

That is, the variance of the lth estimator has been reduced by a factor $1/T_l$.

We now show that the QML has been achieved. Consider the total evolution time $T_t = T_1 + \cdots + T_K$, for K estimations, with $T_l = (\Pi_{l'=2}^{l} a_{l'}) T_1$. Then,

$$T_t = \left[\frac{1}{a_2 \ldots a_K} + \cdots + 1 \right] T_K. \tag{B23}$$

Moreover, since $(1/a_l) \leq 1/(c'-4) < 1$, we have

$$T_t \leq [(c'-4)^{-(K-1)} + \cdots + 1] T_K < \frac{c'-4}{c'-5} T_K. \tag{B24}$$

Equation (B24) gives the total resource scaling $T_t = O[\Delta/(2\Delta_\theta)]$ (i.e., the total evolution time under the action of H), implying that the QML is attained.

APPENDIX C: FIRST STEP IN BLACK-BOX ESTIMATION

For simplicity, assume $\pi/4 > \hat{\theta}_0 > 0$, with $\hat{\theta}_0$ the mean of the prior (normal) distribution of θ. Then, there exists a ϕ_1 such that

$$2\hat{\theta}_0 + 2\phi_1 = \pi/2. \tag{C1}$$

The first estimation of θ can be done by running the algorithm of Fig. 3 replacing $W(T) \to W_a(1, \phi_1) = W_B e^{-i\phi_1 H_0}$. Since H_0 is known, $W_a(1, \phi_1)$ can be implemented with available gates. The output of the first measurement, denoted by y_1, is a measurement of $\cos(\beta_1)$, with $\beta_1 = 2(\theta + \phi_1)$. The *a priori* distribution of β_1 is then given by

$f(\beta_1)=\mathcal{N}(\pi/2,(c\Delta)^2)$ [36]. Similar to the continuous time evolution case, this distribution yields the joint distribution

$$f(y_1,\beta_1)=\frac{1}{\sqrt{2\pi}\Delta}\exp[-(y_1-\cos(\beta_1))^2/2\Delta^2]\times\frac{1}{\sqrt{2\pi}c\Delta}\exp[-(\beta_1-\pi/2)^2/2c^2\Delta^2]\approx\frac{1}{\sqrt{2\pi}\Delta}\exp[-(y_1+\beta_1')^2/2\Delta^2]\times\frac{1}{\sqrt{2\pi}c\Delta}\exp[-(\beta_1')^2/2c^2\Delta^2], \quad \text{(C2)}$$

with $\beta_1'=\beta_1-\pi/2$. The first measurement returns an estimator of β_1, with mean $\hat{\beta}_1=\pi/2-[c^2/(1+c^2)]y_1$, and standard deviation $\Sigma_1=(c/\sqrt{1+c^2})\Delta<\Delta$. Therefore, the first estimation of θ has mean and variance determined by

$$\hat{\theta}_1=\frac{1}{2}\left(\pi/2-\frac{c^2}{1+c^2}y_1\right)-\phi_1, \quad \text{(C3)}$$

$$\Sigma_1'=\Sigma_1/2=\frac{c}{\sqrt{1+c^2}}\frac{\Delta}{2}<\Delta/2\ . \quad \text{(C4)}$$

Clearly, the accuracy of the estimation has increased after the first measurement. Thus, we continue with the adaptive Bayesian method by zooming in on the previously estimated parameters.

[1] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam 1982).

[2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[3] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Ann. Phys. **247**, 135 (1996).

[4] See, for example, V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004), and F. Toscano, D. A. R. Dalvit, L. Davidovich, and W. H. Zurek, Phys. Rev. A **73**, 023803 (2006).

[5] T. Rudolph and L. Grover, Phys. Rev. Lett. **91**, 217905 (2003); M. de Burgh and S. D. Bartlett, Phys. Rev. A **72**, 042301 (2005).

[6] A. Luis, Phys. Lett. A **329**, 8 (2004); S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Phys. Rev. Lett. **98**, 090401 (2007); A. Luis, Phys. Rev. A **76**, 035801 (2007); A. M. Rey, L. Jiang, and M. D. Lukin, *ibid.* **76**, 053617 (2007); S. Boixo, A. Datta, S. T. Flammia, A. Shaji, E. Bagan, and C. M. Caves, *ibid.* **77**, 012317 (2008); S. Choi and B. Sundaram, e-print arXiv:cond-mat/0709.3842.

[7] E. Knill, G. Ortiz, and R. D. Somma, Phys. Rev. A **75**, 012328 (2007).

[8] If, on the contrary, the interaction determines an energy range that scales like n^k, the QML precision achieved by these protocols is $O(1/n^kT)$ [6].

[9] A. Shaji and C. M. Caves Phys. Rev. A **76**, 032111 (2007); S. Boixo, A. Datta, A. Shaji, and C. M. Caves, Laser Phys. **16**, 1525 (2006).

[10] JM Geremia, J. K. Stockton, and H. Mabuchi, Science **304**, 270 (2004).

[11] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys. Rev. A **54**, R4649 (1996).

[12] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. London, Ser. A **454**, 339 (1998).

[13] P. Wocjan and S. Zhang, e-print arXiv:quant-ph/0606179.

[14] E. Knill and R. Laflamme, Phys. Rev. Lett. **81**, 5672 (1998).

[15] R. Laflamme *et al.*, e-print arXiv:quant-ph/0207172.

[16] See, for example, D. Poulin, R. Blume-Kohout, R. Laflamme, and H. Ollivier, Phys. Rev. Lett. **92**, 177906 (2004); A. Datta, S. T. Flammia, and C. M. Caves, Phys. Rev. A **72**, 042316 (2005); D. Poulin, R. Laflamme, G. J. Milburn, and J. P. Paz, *ibid.* **68**, 022302 (2003); P. W. Shor and S. P Jordan, e-print arXiv:0707.2831.

[17] Since $|\mathrm{tr}[U]/2^n|\leq 1$, the normal distribution is an approximation to the actual distribution. The error of such an approximation is determined by the corresponding *error function* and is exponentially small in Δ^{-2}.

[18] On the contrary, for fixed $T=1$, $\Delta_\theta=O(1/\sqrt{n})$ so nothing is gained with respect to the SQL in this case.

[19] Such a decomposition is always possible because Pauli products constitute a basis of su(2^n).

[20] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **65**, 042323 (2002).

[21] S. Lloyd and Jean-Jacques E. Slotine, Phys. Rev. A **62**, 012307 (2000).

[22] All of the above assumptions are satisfied in the first measurement, as seen in Appendix C.

[23] We could choose $q_{l+1}=b_{l+1}q_l$. To guarantee that the next estimation will be made under similar conditions as the previous one, it is sufficient to choose $b_{l+1}=b\equiv\lfloor c'\rfloor$ for all l, with $\lfloor . \rfloor$ the *floor function*. With this choice, $b\Sigma_l<c\Delta$ and $q_{l+1}=b^l$, determining p_{l+1} as

$$p_{l+1}=[2\hat{\theta}_l b^l+2\phi_{l+1}-\pi/2]/(2\pi). \quad \text{(C5)}$$

[24] See, for example, L. Viola and E. Knill, Phys. Rev. Lett. **94**, 060502 (2005), and references therein.

[25] M. Suzuki, Phys. Rev. B **31**, 2957 (1985).

[26] M. Suzuki, Phys. Lett. A **165**, 387 (1992).

[27] To obtain Δ_γ we have assumed first that $c\Delta$ is the standard deviation associated with the prior distribution of every θ^ν. Since $|\gamma(\epsilon,p,T_l,\vec{\theta})|\leq\varepsilon$, it is expected that $\Delta_\gamma<c\Delta\epsilon$. Moreover, since $\gamma(\epsilon,p,T_l,\vec{\theta})$ depends on many stochastic variables, we also assumed that $f(\gamma)$ is unbiased. This assumption may certainly not be right in a general scenario and we would expect that, for some simple cases, such a bias could be calculated and corrected. However, the error introduced by approxi-

mating the mean of $\gamma(\epsilon,p,T_l,\vec{\theta})$ by 0 is also of order ε, with negligible effects in our calculation.

[28] E. Bagan, M. Baig, and R. Munoz-Tapia, Phys. Rev. A **70**, 030301(R) (2004); G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. Lett. **93**, 180503 (2004); G. Chiribella, G. M. D'Ariano, and M. F. Sacchi, Phys. Rev. A **72**, 042338 (2005); S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Rev. Mod. Phys. **79**, 555 (2007); J. Kahn, Phys. Rev. A **75**, 022326 (2007).

[29] L. K. Grover, *Proceedings of the 28th Annual ACM Symposium on the T. of Comp. (STOC)*, 212 (1996).

[30] J. F. Cornwell, *Group Theory in Physics* (Academic, London, 1997).

[31] We have chosen $\pi/2$ because the cosine function in this region is very sensitive to variations in the phase. Nevertheless, our analysis can still be performed choosing another T_1, as long as $\sin(2\hat{\theta}_0 T_1)$ is not very close to 0.

[32] This can be done by choosing the Pauli products of Eq. (10) in the algorithm of Fig. 3, or by compensating the phase by $(\pi/2)-\hat{\theta}_0$ in the first measurement. Such a phase compensation will be modified as we learn about θ (see Sec. II D).

[33] Numerical and asymptotic analyses show that the linear approximation results in an error $O[(c\Delta/T_1)^2] \ll \Delta/T_1$ in the estimator of Eq. (B7) and an error of $O[(c\Delta)^4]$ in the computed standard deviation.

[34] It is expected that the first measurement outcome lies, in the worst case scenario, within a few standard deviations $c\Delta$ from 0.

[35] The main necessary condition for choosing a_2 is that the different phases mod2π do not get aliased when measuring $\cos(2\theta T_2)$. Nevertheless, we require $a_2\Delta \ll 1$, so a linear approximation to the cosine function is still possible. We have chosen, for simplicity, $a_2=O(c')$. The offset of 4 in Eq. (B11) is due to $a_2 \approx (1+4p_2)$ and $p_2 \in \mathbb{N}^*$.

[36] The constant c in this case may be even smaller than the one introduced in Sec. B. The normal distribution is due to β_1 being a linear transformation of θ.