Universal implementation of projective measurement of energy

Shojun Nakayama,1 Akihito Soeda,1, 2 and Mio Murao1, 3

1Department of Physics, Graduate School of Science, the University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan
2Centre for Quantum Technologies, National University of Singapore, Singapore
3Institute for Nano Quantum Information Electronics, the University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo, Japan
(Dated: October 14, 2013)

We present a scheme to asymptotically implement a projective measurement in the energy eigenbasis on a finite dimensional system driven by an unknown Hamiltonian $H$ based on the quantum phase estimation algorithm. Our scheme also provides an outcome associated with an energy eigenvalue of $H$. Two new algorithms are introduced to apply the quantum phase estimation algorithms for unknown Hamiltonian systems. One is for asymptotically but universally implementing a controlled unitary operation $C_{U(t)}$ of a unitary operation $U(t) = e^{-iHt}$ up to the global phase of $U(t)$ for an unknown Hamiltonian $H$. This algorithm utilizes random unitary operations achieve a decoupling effect. The other is an algorithm for evaluating the absolute value of the trace of $U(t)$ without using $C_{U(t)}$ required to run the first algorithm. We analyze the accuracy and the running time of our scheme and show that the running time of our scheme is independent of the dimension of the system for a given accuracy.

PACS numbers: 03.67.Ac, 03.67.Hk, 03.67.-a

Introduction.— A unique feature of quantum mechanics is that a measurement not only returns an outcome, but also changes the state of the measured system according to the outcome. Suppose a system evolves by a certain Hamiltonian. Using an appropriate measurement operation, it projects the system to an energy eigenstate and returns the associated energy eigenvalue. We refer to this measurement as the projective measurement of energy (PME). This operation plays a crucial role in the recent theories on the statistical foundations of thermodynamical relations [1, 3]. The implementation of certain quantum thermalization algorithms (e.g., [4, 5]) assumes that PME is implementable. In addition, PME freezes the evolution of the system as eigenstates are also stationary states—a useful feature when the system needs to be stabilized.

Generally, the dynamics of a physical system is non-trivial to control. For instance, the evolution of an isolated system is completely determined by its Hamiltonian. On the other hand, the controllability of some systems has improved significantly over the past few years. Let us imagine a perfectly controllable system, and exploit its controllability to implement PME on the Hamiltonian system. The time resource of the implementation is lower bounded by the time needed to run the Hamiltonian system.

We may regard the controllable system as a quantum computer. If eigenvalues and eigenstates of the Hamiltonian are known, then implementing PME is simple: We transfer the state of the system into the quantum computer, perform PME within the device (which is possible because it implements any quantum operation), and transfer back the resulting state. PME on a quantum computer is realized in several ways. The most straightforward method, but extremely inefficient, involves a diagonalization of the Hamiltonian and transformation from the computational basis to the energy eigenbasis. If a local decomposition of Hamiltonian is known, there is a polynomial implementation [3, 5] using Kitaev’s phase estimation algorithm [6].

On the other hand, if the Hamiltonian is completely unknown, the only known solution uses an identification of the Hamiltonian as discussed in Ref. [2]. A perfect Hamiltonian identification is possible, e.g., via process tomography [8], but requires an infinite time resource. For process tomography, the necessary resource to achieve a desired accuracy for a system of $d$ dimensions increases on the order of $O(d^4)$. The scheme discussed in Ref. [8] implements PME by using phase estimation and without a Hamiltonian identification, but it requires special assumptions only satisfied in particular setups such as in linear optical quantum computation using photon qubits.

In this paper, we approximately implement PME on any finite dimensional system of an unknown Hamiltonian, achieving the ideal PME in the limit of infinite time resource, but its running time for a given accuracy is independent of the system dimension.

Projective measurement of energy.— Given a Hamiltonian $H$ of a $d$-dimensional system $\mathcal{H}$, we denote its spectral decomposition by $H = \sum_{k=1}^{L} E_k P_k$, where $L$ is a natural number of $L \leq d$, $E_k$ are the eigenvalues in the increasing order, i.e., $E_1 < E_2 < \cdots < E_L$, $P_k$ are projectors on the eigenspace of $E_k$. A mathematical description of its projective measurement of energy is given by a set of pairs $\{(E_i, P_i)\}$ of outcomes $\{E_i\}$ and a set of projective completely positive maps called instruments $\{P_i\}$, where $P_i$ is a completely positive (CP) map that projects its input state to the eigenspace associated with $E_i$, i.e., $P_i \rho = P_i \rho P_i$. For a density matrix $\rho \in \mathcal{S}(\mathcal{H})$, the probability to obtain the outcome $E_i$ for a projective
measurement of energy \((\{E_i, P_i\})\) is performed is given by \(\text{Tr}[\rho]\).

**Phase estimation.**— The phase estimation algorithm (PEA) runs on a system consisting of \(N\) control qubits and a \(d\)-dimensional target system. Let \(U\) be a \(d \times d\) unitary matrix with \(U\) with a spectral decomposition \(\sum_{k=1}^{d} \exp(i\theta_k)|\theta_k\rangle\langle\theta_k|\), where \(0 \leq \theta_k < 2\pi\). A controlled-unitary operation \(C_U\) of \(U\) is defined by \(C_U := |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U\). PEA uses \(N\) controlled unitary operations of \(C_U^\otimes_k\) for \(k = 0, 1, \ldots, N - 1\). When the target system is initially in an eigenstate \(|\theta_k\rangle\), PEA returns a \(N\)-bit string \(\{n_1, \ldots, n_N\}\), where \(n_k \in \{0, 1\}\), according to the probability distribution

\[
P_N(2\pi f_N|\theta_k) = \left(\frac{\sin[2N(\theta_k - 2\pi f_N)/2]}{2N\sin[(\theta_k - 2\pi f_N)/2]}\right)^2,
\]

where \(f_N := 0.n_1 \cdots n_N\) is a binary decimal number [11]. For \(N \to \infty\), \(f_N\) can be regarded as a continuous variable \(f\) and \(P_N(2\pi f_N|\theta_k)\) approaches the delta function \(\delta(\theta_k - f)\) [12], hence \(2\pi f_N\) provides a good estimate of \(\theta_k\) when \(N\) is sufficiently large. In the same limit, PEA becomes a projective measurement on the target system in the basis \(|\theta_k\rangle\) [11]. In case of degenerate eigenvalues, the input state is projected to its respective eigensubspace, e.g., \(|\varphi\rangle \to P_k|\varphi\rangle\). For each \(f_N\) is associated a CP map \(\mathcal{I}_{f_N}\).

**Universal controllization.**— With finite time resource, \(C_{U(t)}\) for \(U(t) = \exp(-iHt)\) cannot be perfectly implemented when \(H\) is unknown. We address this issue by implementing an approximated \(C_{U(t)}\).

Let us denote the Hilbert spaces of the control and the target system by \(\mathcal{H}_c\) and \(\mathcal{H}_t\), respectively. First, we add an ancilla system, with Hilbert space \(\mathcal{H}_a = \mathbb{C}^d\), and initially prepare it in the completely mixed state \(I/d\). We introduce the pseudo controlled unitary operation \(W_{U(t/m)}\) of \(C_{U(t/m)}\), which is a unitary operation on \(\mathcal{H}_c \otimes \mathcal{H}_a \otimes \mathcal{H}_t\), defined by

\[
W_{U(t/m)} = \tilde{F}(I \otimes I \otimes U(t/m))\tilde{F},
\]

where \(\tilde{F} = |0\rangle\langle 0| \otimes I \otimes I + |1\rangle\langle 1| \otimes SWAP\). For an input density operator \(\rho := \sum_{j,k} |j\rangle\langle k| \otimes I/d \otimes \rho_{j,k}\) on \(\mathcal{H}_c \otimes \mathcal{H}_a \otimes \mathcal{H}_t\), we see that

\[
W_{U(t/m)}\rho W_{U(t/m)}^\dagger = \sum_{j,k} |j\rangle\langle k| \otimes \frac{1}{d} U((j-k)t/m) \otimes U(jt/m) \rho_{j,k} U^\dagger((kt/m),
\]

where \(j, k = 0, 1\). The operation \(W_{U(t/m)}\) generally entangles the ancilla and the rest of the system.

The entanglement is broken by the use of the following randomization. We perform \(m\) iterations of unitary operation \(V^{(r)}_{U(t/m)} = (I \otimes \sigma_r \otimes I)W_{U(t/m)}(I \otimes \sigma_r \otimes I)\) on \(\mathcal{H}_c \otimes \mathcal{H}_a \otimes \mathcal{H}_t\), where \(\sigma_r\) is uniformly and randomly chosen for each iteration from a set of \(D\) operations \(\{\sigma_r\}\) such that \(\frac{1}{D} \sum_{r} \sigma_r = \text{Tr} U - I\), for all unitary operators \(U\) on \(\mathcal{H}_t\). We denote the CPTP map of the randomized \(V^{(r)}_{U(t/m)}\) by

\[
\mathcal{V}_{U(t/m)}(\rho) := \frac{1}{D} \sum_{r} V^{(r)}_{U(t/m)}\rho V^{(r)*}_{U(t/m)}
\]

and define the reduced CPTP map \(\Gamma_{U(t/m)}\) of \(\mathcal{V}_{U(t/m)}\) on \(\mathcal{H}_c \otimes \mathcal{H}_t\) for the reduced state \(\rho_{ct} := \text{Tr}_a\rho\) by

\[
\Gamma_{U(t/m)}(\rho_{ct}) := \text{Tr}\left[\mathcal{V}_{U(t/m)}(\rho_{ct} \otimes I/d)\right].
\]

The randomization process transforms all components on \(\mathcal{H}_a\) to states proportional to the completely mixed state, hence \(\mathcal{V}_{U(t/m)}(\rho_{ct} \otimes I/d) = \Gamma_{U(t/m)}(\rho_{ct}) \otimes I/d\). Random operations are used in a similar spirit when decoupling a system from an interacting environment [13], but in this case the random operations are applied on the system. See Fig. 1(b) for a quantum circuit representation of the algorithm.

**Accuracy of controllization.**— Let us analyze the accuracy of the approximation. For \(m\) iterations of the maps \(\Gamma_{U(t/m)}\) and \(\mathcal{V}_{U(t/m)}\) (denoted by \(\Gamma^m_{U(t/m)}\) and \(\mathcal{V}^m_{U(t/m)}\) respectively), we have

\[
\Gamma^m_{U(t/m)}(\rho_{ct}) = \text{Tr}\left[\mathcal{V}^m_{U(t/m)}(\rho_{ct} \otimes I/d)\right].
\]

We define a controlled unitary “up to phases” by

\[
C^{(g_U)}_{U(t)} := |0\rangle\langle 0| \otimes I + e^{i g_U} |1\rangle\langle 1| \otimes U,
\]

where \(g_U\) is a real function of \(U\). Let \(C^{(g_U)}_{U(t)}\) be the CPTP map representing unitary operation \(C_U^{(g_U)}\). The difference between the map \(C^{(m g_U)}_{U(t)}\) and \(\Gamma^m_{U(t/m)}\) evaluated by the diamond norm [14] is

\[
\left\|C^{(m g_U)}_{U(t)} - \Gamma^m_{U(t)}\right\| = 1 - (a_{U(t/m)})^m,
\]

where \(a_{U(t/m)} := |\text{Tr}[U(t/m)]/d|\). We refer to \((a_{U(t/m)})^m\) as the coherence factor. We also define a phase factor \(\varphi_{U(t/m)} \in [E_t/t_m, E_t/t_m]\) so that

\[
e^{-i \varphi_{U(t/m)}} = \text{Tr}[U(t/m)]/|\text{Tr}[U(t/m)]|.
\]

Let \(\Delta_{\text{max}}\) be the maximum difference in the energy eigenvalues, i.e., \(\Delta_{\text{max}} = E_2 - E_1\), then \(\varphi_{U(t/m)}\) is determined uniquely, whenever \(\Delta_{\text{max}}t \leq \pi/2\).
For $\Delta_{\text{max}} t < \pi/2$, the quantity $a_{U(t/m)}$ satisfies
\[ a_{U(t/m)} \geq \sqrt{\cos(\Delta_{\text{max}} t/m)} \geq 1 - (\Delta_{\text{max}})^2 t^2 / 4m^2. \tag{9} \]

By Ref. [15], $m\varphi_{U(t/m)}$ converges to $\text{Tr}[H] t / d$, thus $\Gamma_{U(t/m)}$ converges to $C_{U(t)}$ for $m \to \infty$.

**PEA with approximated controllization.**— A spectral decomposition of $U(t) = \exp(-iHt)$ is given by $U(t) = \sum_{k=1}^{L} e^{-i\epsilon_k t} P_k$. Define $\theta_k \in [0, 2\pi)$ so that $\theta_k = -\epsilon_k t + 2\pi \nu$ for $\nu \in \mathbb{Z}$. Let us replace each $C_{U(t/m)}$ with the approximated map $\Gamma_{U(t/m)}$. The probability distribution of $f_N$ given an eigenvector $|\theta_k\rangle$ of $\theta_k$ is
\[ Q_N(2\pi f_N | \theta_k, U(t/m)) = \prod_{k=1}^{N} \frac{1}{2} [1 + (a_{U(t/m)})^{2m^2} \cos^{2k-1} \{ \theta_k + m\varphi_{U(t/m)} - 2\pi f_N \}], \tag{10} \]
according to Ref. [16]. We observe that if the coherence factor satisfies
\[ 1 - (a_{U(t/m)})^{2m^2} \leq \delta / N \tag{11} \]
for a fixed $\delta \in [0, 1/2]$ and $\forall N \in \mathbb{N}$, the probability distribution $Q_N$ satisfies
\[ Q_N(2\pi f_N | \theta_k, U(t/m)) - P_N(2\pi f_N | \theta_k + m\varphi_{U(t/m)}) \leq e^\delta - 1 \leq 2\delta. \tag{12} \]
To satisfy Eq. (11), it suffices to have
\[ m \geq (\Delta_{\text{max}})^2 t^2 N^2 / 4\delta. \tag{13} \]

**Accuracy of approximated PME.**— The exact PME of unknown $H$ can only be approximated under finite time resource. An approximated PME will show deviations from the ideal behaviors. The accuracy of implementation can be measured by quantifying these deviations.

We define a measurement by $\mathcal{M} = \{(x_i, \mathcal{I}_i) | i \in X\}$ where $\{x_i\}$ is a set of real valued outcomes, $X$ is a countable set of indices, and $\{\mathcal{I}_i\}$ denotes an instrument. For example, given $H$ with spectral decomposition $H = \sum_k E_k P_k$, its PME is described by $\mathcal{M} = \{(E_i, \mathcal{I}_i)\}$ where a set of real valued outcomes $\{E_i\}$ representing energy eigenvalues and $\{\mathcal{I}_i\}$ is an instrument that maps an in put state $\rho$ to an energy eigenstate $|E_i\rangle$ (if the Hamiltonian is not degenerate) with probability $\text{Tr}[\mathcal{I}_i(\rho)]$.

We require that a good approximation of PME should give a good estimate of the expectation value of the energy for any given input state, and that, for any input state, the resulting state after the process is close enough to the one we would have obtained via an ideal PME. As a necessary condition, we may require these properties for the case when the input is an eigenstate. Furthermore, the reference point of energy is arbitrary and that

a physical system evolving by $H$ cannot be distinguished from the one by $H + \lambda I$. PME for $H$ and $H + \lambda I$ should be considered equivalent.

Let us introduce the following functions $R_1$ and $R_2$. We define $R_1$ that evaluates fluctuation of the outcome by
\[ R_1(\mathcal{M} | H) = \max_{\lambda, i, k} \sum_{j \in X} \text{Tr} \left[ \mathcal{I}_j (|E_i^\lambda\rangle \langle E_i^\lambda|) (E_i - x_j)^2 \right], \tag{14} \]
where $\lambda$ is a continuous variable which identifies a state from the eigenspace and $|E_i^\lambda\rangle$ is the corresponding state. We define $R_2$ that evaluates the irreproducibility of energy eigenstate of the measurement
\[ R_2(\mathcal{M} | H) = \max_{\lambda, i, j} \sum_{k \in K} \left| \sum_{\nu \in X} \mathcal{I}_j (|E_i^\lambda\rangle \langle E_i^\lambda|) - |E_i^\lambda\rangle \langle E_i^\lambda| \right|. \tag{15} \]
The larger values indicate poorer performance. These functions apply even for degenerate Hamiltonians.

We estimate the energy eigenvalues for a given $f_N$ according to
\[ E(f_N) = \begin{cases} -2\pi f_N / t & \text{for } f_N \in [0, 1/2) \\ -(2\pi f_N - 2\pi) / t & \text{for } f_N \in [1/2, 1) \end{cases}. \tag{16} \]
Let $\mathcal{M}_{\text{PEA}}$ denote the measurement implemented by our scheme and define $\tilde{H} = H - (m\varphi_{U(t/m)} / t) I$. A spectral decomposition of $\tilde{H}$ is given by $\tilde{H} = \sum_{k=1}^{L} E_k \tilde{P}_k$. Notice that $\tilde{E}_k = E_k - m\varphi_{U(t/m)} / t$ and $P_k \equiv \tilde{P}_k$. The distance between $\mathcal{M}_{\text{PEA}}$ and PME of $\tilde{H}$ in terms of $R_1$ is
\[ R_1(\mathcal{M}_{\text{PEA}} | \tilde{H}) = \max_{\lambda, k} \sum_{f_N} \text{Tr} \left[ \mathcal{I}_f (|E_k^\lambda\rangle \langle E_k^\lambda|) (\tilde{E}_k - E(f_N))^2 \right]. \tag{17} \]
Note that $P_k \equiv \tilde{P}_k$ implies $|E_k^\lambda\rangle = |E_i^\lambda\rangle$, thus
\[ \text{Tr} \left[ \mathcal{I}_f (|E_k^\lambda\rangle \langle E_k^\lambda|) \right] = Q_N(2\pi f_N | \theta_k, U(t/m)). \]
In addition, $\max_x f(x) \leq \max_x g(x) = \max_y f(y) - g(y)$. If $\Delta_{\text{max}} t < \pi / 2$, we have $\tilde{E}_k \in (-\pi / t, \pi / t)$, thus $|\tilde{E}_k - E(f_N)| \leq 2\pi / t$. Therefore, when Eq. (12) holds,
\[ R_1(\mathcal{M}_{\text{PEA}} | \tilde{H}) \leq R_1(\mathcal{M}_{\text{CU(t)}} | \tilde{H}) + 2^N (2\delta) 4\pi^2 / t^2, \tag{18} \]
where $\mathcal{M}_{\text{CU(t)}}$ denotes the measurement implemented using the ideal controlled unitary operation $C_{U(t)}$. By Ref. [17], the first term is bounded as
\[ R_1(\mathcal{M}_{\text{CU(t)}} | \tilde{H}) \leq \varepsilon + 16\pi^2 / t^4 2^N \varepsilon \tag{19} \]
for $\varepsilon \in (0, 1]$.

**Evaluation of the maximum energy difference.**— Even when the upper bound of $\Delta_{\text{max}}$ is not given, we can estimate $\Delta_{\text{max}}$ for large dimensional systems if the following two conditions can be assumed. The one is that $U(t) = e^{-iHt}$ is sampled according to the Haar measure
The probability that \( \theta_i + 2n_i \pi \) is taken uniformly randomly from a set \( n_i \in \{0, 1, \cdots, \lfloor \Delta_{\text{max}} \pi / (2\pi) \rfloor \} \).

Our method of obtaining the bound of \( \Delta_{\text{max}} \) consists of iterating of evaluation of the value of \( \langle a_{U(t/2)} \rangle^2 \) by using the algorithm evaluating the coherence factor for \( k = 2, 3, \cdots \) starting from arbitrarily chosen \( t \). The task of this algorithm is similar to the deterministic quantum computation with one pure qubit (DQC1) algorithm calculating \( \text{Tr}U \), but our algorithm employs pseudo controlled unitary operations in stead of the ideal ones. A circuit representation of the algorithm is given in Fig. 2.

A simple calculation reveals that the reduced state of the control qubit after the last Hadamard gate is given by \( \rho_c = \frac{1}{d} + \frac{a_{U(t)} \sigma_z}{2} \). The coherence factor can be calculated from the average value of \( \sigma_z \) via the measurement on the control system, i.e., \( \text{Tr}[\rho_c \sigma_z] = a_{U(t)} \). Therefore, \( M \) iterations of this algorithm estimate the coherence factor with error of \( O(1/\sqrt{M}) \).

As we see below, for all \( k \geq 2 \) satisfying \( \Delta_{\text{max}} / 2^k \geq 4\pi \), the probability that \( \langle a_{U(t)} \rangle^2 \geq 1/2 \) decreases exponentially with \( d \); on the other hand, if \( \Delta_{\text{max}} < \pi / 2 \), \( \langle a_{U(t)} \rangle^2 \) is always guaranteed to be larger than 1/2. Hence if we take the smallest \( k' \) such that \( \langle a_{U(t/2^{k'})} \rangle^2 \) is larger than 1/2, we can bound \( \Delta_{\text{max}} \) as \( \Delta_{\text{max}} < 2^{k' + 2}\pi / t \) with success probability more than 1 \(- O(\exp(-d)) \). See Ref. 19 for the proof.

The sketch of the proof is as follows. For the CUE, the expectation value of any power \( r \) of \( \langle a_{U(t)} \rangle^2 \) is given by using the correspondence between CUE and random permutations 19, 20, i.e.,

\[
\langle (a_{U(t)})^{2r} \rangle_{\text{CUE}} \leq \frac{r!}{d^{2r}}. \tag{20}
\]

The probability that \( \langle a_{U(t)} \rangle^2 \geq 1/2 \) is bounded as

\[
P \{ \langle a_{U(t)} \rangle^2 \geq 1/2 \} \leq \exp \left( -\frac{3(d - 2)^2}{4d + 4} \right) + \left( \frac{\log d/2}{d} \right)^d \frac{\log d/2}{d - \log d/2} \left( \frac{2}{d} \right)^{d - 1} \tag{21}
\]

by using Bernstein’s inequality 19. Next we consider the expectation value of \( \langle a_{U(t/2)} \rangle^2 \). If \( n_i \) is odd, the \( i \)-th eigenvalue of \( U(t/2) \) is \( e^{i\theta_i} / 2 \). Although the probability distribution of \( n_i \) is uniformly random, when \( \Delta_{\text{max}} / \pi > 4 \), the probability depends on the value of \( \lfloor \Delta_{\text{max}} / 2\pi \rfloor \). In the most biased case, they are 2/3 and 1/3 for even and odd cases, respectively. Then the average value of \( \langle a_{U(t/2)} \rangle^2 \) is less than or equal to 1/3. Substituting this average property to Bernstein’s inequality 19, the probability can be bounded as

\[
P \{ \langle a_{U(t/2)} \rangle^2 \geq 1/2 \} \leq \exp \left( -\frac{3(24 \pi - 23\sqrt{2})d}{24} \right) + \exp \left( -\frac{74 + 27\sqrt{2}}{328}d \right). \tag{22}
\]

Thus we can estimate \( \Delta_{\text{max}} \) by searching \( k' \) using the algorithm evaluating the coherence factor.

Analysis of the running time.— The running time of our algorithm is given by \( O((\Delta_{\text{max}})^3 / \varepsilon^2) \), which is independent of the dimension of the system \( d \). See 17 for the details of estimation of the running time. On the other hand, as shown in Ref 11, the runtime increases twice for adding a digit to the measurement outcome. Thus the running time increases exponentially in the number of the decimal places of the outcome.

Conclusion.— In this letter, we presented a new scheme to universally and asymptotically implement the projective measurement of energy on a finite dimensional system driven by an unknown Hamiltonian by coupling the system to a quantum computer of which time scale is much faster than that of the Hamiltonian dynamics. Our algorithm is based on two new subroutines to apply the phase estimation algorithm for such a system. One is a universal controllization algorithm and the other is an algorithm evaluating the absolute value of the trace of a unitary operator representing Hamiltonian dynamics. The time cost of our algorithm to achieve a given accuracy does not depend on the dimension of the system, whereas the brute-force measurement scheme based on process tomography and diagonalization of the Hamiltonian cannot avoid such cost.

Acknowledgments: The authors thanks T. Sugiyama for suggesting the use of Bernstein’s inequality and H. Nishiura for pointing out the Robinson-Schensted correspondence required to proof that the algorithm evaluating the coherence factor can also used for estimate \( \Delta_{\text{max}} \) for CUE. This work is supported by the Project for Developing Innovation Systems of MEXT, Japan, the Global COE Program of MEXT Japan, and JSPS KAKENHI (Grant No. 23540463, and No. 23240001). The authors also gratefully acknowledge to the ELC project (Grant-in-Aid for Scientific Research on Innovative Areas MEXT KAKENHI (Grant No. 24106009)) for encouraging the research presented in this paper. After the completion of our work, we were notified that our algorithm for calculating \( |\text{Tr}[U(t)]| \) based on DQC1 has been independently discovered by J. Thompson, M. Gu, K. Modi, and V. Vedral in “Quantum Computing with Black-box Subroutines” 21. We thank these authors for drawing their work to our attention.
Appendix A: Summary of phase estimation algorithm

The phase estimation algorithm is a quantum algorithm to estimate the phase factor $0 \leq \theta_i < 2\pi$ of the eigenvalue $e^{i\theta_i}$ of a finite dimensional unitary operation $U$ when an eigenstate $|\theta_i\rangle$ is given. In the followings, we assume that there is no degeneracy for $U$. The algorithm is presented in Fig. 3. As shown in the figure, this algorithm uses controlled-unitary operations of $U, U^2, U^2, \ldots, U^{2^k}$ where $N$ denotes the number of control qubits. A controlled-unitary operation $C_U$ of an unitary operation $U$ is defined by

$$C_U := |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes U$$  \hspace{1cm} (A1)

on $\mathcal{H}_c \otimes \mathcal{H}_t$ where the Hilbert spaces of the control system and the target system are represented by $\mathcal{H}_c = \mathbb{C}^2$ and $\mathcal{H}_t = \mathbb{C}^d$ ($d < \infty$), respectively.

The final measurement of $N$ control qubits in the computational base provides a sequence of outcomes $\{n_1, \ldots, n_N\}$ where $n_k \in \{0, 1\}$ corresponds to the measurement outcome of the $k$-th control qubit. Then the estimated phase is given by $2\pi f_N$ where $f_N := 0.n_1 \cdots n_N$ is the binary representation of a decimal number defined by $0.n_1 \cdots n_N := \sum_{k=1}^{N} 2^{-k} n_k$. When an input state of the target system is $|\theta_i\rangle$, the probability distribution of obtaining $2\pi f_N$ is

$$P_N(2\pi f_N|\theta_i) = \left(\frac{\sin[2^N(\theta_i - 2\pi f_N)/2]}{2^N \sin[\theta_i - 2\pi f_N/2]}\right)^2.$$  \hspace{1cm} (A2)

The output state of the target system remains in $|\theta_i\rangle$. The probability distribution of the estimation error $y := \theta_i - 2\pi f_N$ is given by

$$p_N(y) := P_N(\theta_i - y|\theta_i) = \left(\frac{\sin[2^N y/2]}{2^N \sin[y/2]}\right)^2.$$  \hspace{1cm} (A3)

This means that $p_N(y)$ for $y \neq 0$ decreases exponentially in $N$. Since $f_N$ can be regarded as a continuous variable for $N \to \infty$, $p_N(y)$ is transformed to a delta function (the proof is given in the next section of this supplemental material) and $2\pi f_N$ converges to $\theta_i$ as $N \to \infty$.

If we apply the phase estimation algorithm to an arbitrarily superposed input state $|\phi\rangle = \sum_i \alpha_i |\theta_i\rangle$ where $\sum_i |\alpha_i|^2 = 1$, the probability distribution of obtaining the outcomes $\{n_1, \ldots, n_N\}$ represented in terms of $f_N := 0.n_1 \cdots n_N$ is given by

$$P_N(2\pi f_N|\phi) = \sum_i |\alpha_i|^2 P_N(2\pi f_N|\theta_i).$$  \hspace{1cm} (A4)

The corresponding output state of the target system is

$$|\phi'_{f_N}\rangle = \sum_i \alpha_i (P_N(2\pi f_N|\theta_i))^{1/2} e^{iG(\theta_i, f_N)} |\theta_i\rangle,$$  \hspace{1cm} (A5)

where $G(\theta_i, f_N) = (2^N - 1)(\theta_i - 2\pi f_N)$. Since $2\pi f_N \to \theta_i$ is guaranteed for $N \to \infty$, the output state converges to $|\theta_i\rangle$ when the outcome is $f_N$. Thus the phase estimation algorithm implements a projective measurement in the eigenbasis of $U$ for $N \to \infty$.

If $U$ is given by a Hamiltonian evolution operator $U(t) = e^{-i\hat{H} t}$ (in the unit $\hbar = 1$) of a Hamiltonian $\hat{H}$, the phase estimation algorithm asymptotically implements a projective measurement of energy $\mathcal{M}$ on $\mathcal{H}_t$ that gives an outcome representing an energy eigenvalue $E_i$ of $\hat{H}$ up to periodicity of the phase on $\mathcal{H}_t$ and also gives a set of completely positive map (an instrument) $\{I_E\}$ to an eigenstate corresponding to the outcome $|E_i\rangle$. This property of the phase estimation algorithm is also used in thermalization algorithms $\mathbb{E}$.

Appendix B: Convergence of $p_N(y)$ to a delta function

The function $p_N(y)$ can be regarded as a probability distribution of a discrete random variable $y^N_n = 2\pi(n + \Delta)/2^N$ with $-2^{N-1} \leq n < 2^{N-1}$ satisfying

$$\sum_{n=-2^{N-1}}^{2^{N-1}-1} p_N(y_n^N) = 1$$  \hspace{1cm} (B1)

where $-1/2 \leq \Delta < 1/2$. We show that this discrete probability distribution converges to a delta function on the section $[-\pi, \pi]$ at $N \to \infty$.

To transform a discrete random variable to a continuous one, we construct a probability measure $\mu$ corresponding to the probability distribution $p_N^A$ at the limit $N \to \infty$. The measure is naturally defined by

$$\mu(A) = \lim_{N \to \infty} \sum_{y_n^N \in A} p_N(y_n^N),$$  \hspace{1cm} (B2)

where $A$ is a subset of $[-\pi, \pi]$. If $A = [a, b]$ for $a, b > 0$, we can bound $\mu(A)$ as

$$\mu(A) \leq \lim_{N \to \infty} N_A \left(\frac{1}{2^N \sin a}\right)^2.$$  \hspace{1cm} (B3)
Since $N_A \leq (b-a)(2^N+1)$,

$$0 \leq \mu(A) \leq \frac{b-a}{\sin^2 \frac{\pi}{2} a} \lim_{N \to \infty} \frac{2^N+1}{2^{2N}} = 0.$$  \hspace{1cm} (B4)

This implies $\mu(A) = 0$. If $A = [a,b]$ for $a, b < 0$, we can similarly obtain $\mu(A) = 0$. Thus $\mu(A) = 1$ should be satisfied in the case of $a < 0 < b$ due to Eq. (B1).

We define a set of functions $F$ of which elements are measurable and continuous at $x = 0$. Integral of $f \in F$ over the section $[-\pi, \pi]$ by the probability measure $\mu$ is represented by

$$\int_{-\pi}^{\pi} \mu(dx)f(x) = \lim_{n \to \infty} \sum_{k=-\infty}^{\infty} \frac{k}{n} \mu(A_k^n)$$  \hspace{1cm} (B5)

where $A_k^n = f^{-1}([k/n, (k+1)/n])$. This definition of integral is well defined due to the condition of $f$ to be measurable.

There is an integer $i$ such that $\{f(0)\} \subseteq A_i^0$ (or $\{f(0)\} = A_i \cap A_{i+1}$), where $A^i$ denotes the interior of the set $A$. ($A$ means the closure of $A$.) Due to the continuity, $A_i$ (or $A_i \cup A_{i+1}$) includes a section $[a, b]$ of $a < 0 < b$. Then $\mu(A_i) = 1$ should be satisfied. Thus the integral satisfies

$$\int_{-\pi}^{\pi} f(x)\mu(dx) = f(0).$$  \hspace{1cm} (B6)

Formally, we can represent this probability measure as a probability distribution $p(x)$ of a continuous random variable $x$,

$$\int_{-\pi}^{\pi} f(x)\mu(dx) = \int_{-\pi}^{\pi} f(x)p(x)dx = f(0).$$  \hspace{1cm} (B7)

Therefore the function $p$ can be regarded as a delta function for the set of functions $F$. The set of function $F$ includes smooth functions that are known as descriptions of physically natural values.

We note that the measure $\mu$ is not a measure called the “Dirac measure”. The set $A = \{2\pi \Delta/2^N \mid N \geq 1\}$ has non-zero value under the $\mu$ measure, but it gives zero under the Dirac measure.

Appendix C: Derivation of the diamond norm between $C_{U(t)}^{(m \varphi_U(t))}$ and $\Gamma_{U(t)}^m$

The diamond norm [22] is a norm of a liner superoperator, which takes account of the effect of the superoperator on the larger Hilbert space than the range of the superoperator. It is often used to evaluate the difference between two CPTP maps in quantum information.

A superoperator $S$ on a Hilbert space $H$ acting on an extended system $H \otimes H'$ satisfies

$$\|S\|_{op} < \|(|S \otimes I_{H'}|)\|_{op}$$  \hspace{1cm} (C1)

where the operator norm $\|S\|_{op}$ is the maximum of the trace norm of $S[A]$ for an operator $A$ under the condition $\|A\|_{tr} = 1$ and $I_{H'}$ denotes an identity superoperator on $H'$. Trace norm is defined as $\|A\|_{tr} = Tr |AA^\dagger|$. Since $\|S \otimes \mathcal{I}_{H'}\|_{op} \leq \|S \otimes \mathcal{I}_{H}\|_{op}$ holds for any Hilbert space $H'$, it is enough to consider $\|S \otimes \mathcal{I}_{H}\|_{op}$ to find the bound. Thus the diamond norm $\|\cdot\|_{\diamond}$ of a superoperator $S$ on the Hilbert space $H$ is defined by

$$\|S\|_{\diamond} := \|S \otimes \mathcal{I}_{H}\|_{op}.$$  \hspace{1cm} (C2)

The following lemma is convenient for calculating the diamond norm [22].

**Lemma 1** For any Hermitian preserving super operator $\Gamma$ on the Hilbert space $H$, the following equation is satisfied.

$$\|\Gamma\|_{\diamond} = \max_{P \in \Pi_1} \|(|\Gamma \otimes \mathcal{I}_{H}|)P\|_{tr}$$  \hspace{1cm} (C3)

where $\Pi_1$ is a set of rank-1 projectors on $H \otimes H$.

We evaluate the difference between the two CPTP maps $C_{U(t)}^{(m \varphi_U(t/m))}$ and $\Gamma_{U(t)}^m$ in terms of the diamond norm. Both maps act on the Hilbert space $\mathcal{H}_c \otimes \mathcal{H}_t = \mathcal{C}^2 \otimes \mathcal{C}^d$. Thus we search rank-1 projectors on $(\mathcal{H}_c \otimes \mathcal{H}_t)^{\otimes 2}$ to calculate the diamond norm. Any vector $|\Psi\rangle$ in $(\mathcal{H}_c \otimes \mathcal{H}_t)^{\otimes 2}$ can be represented by $|\Psi\rangle = \alpha |0\rangle \langle \psi| + \beta |1\rangle \langle \phi|$ where $\{|0\rangle, |1\rangle\}$ is the computational basis of the first control qubit system $\mathcal{H}_c$, by taking appropriate states $|\psi\rangle, |\phi\rangle \in \mathcal{H}_t \otimes \mathcal{H}_c \otimes \mathcal{H}_t$ and $\alpha, \beta$ satisfying $|\alpha|^2 + |\beta|^2 = 1$. Any rank-1 operator on $(\mathcal{H}_c \otimes \mathcal{H}_t)^{\otimes 2}$ can be written by $|\Psi\rangle \langle \Psi|$.

As a partitioned matrix, the projector is represented by

$$|\Psi\rangle \langle \Psi| = 
\begin{pmatrix}
|\alpha|^2 |\psi\rangle \langle \psi| & \alpha \beta^* |\psi\rangle \langle \phi| \\
\alpha^* \beta |\phi\rangle \langle \psi| & |\beta|^2 |\phi\rangle \langle \phi|
\end{pmatrix}. \hspace{1cm} (C4)$$

The left upper partition corresponds to the $|0\rangle |0\rangle$ element of the first system represented by $\mathcal{H}_c = \mathcal{C}^2$. The right upper partition is the $|0\rangle |1\rangle$ element, and so as the others. The projector $|\Psi\rangle \langle \Psi|$ is transformed by the maps $C_{U(t)}^{(m \varphi_U(t))}$ and $\Gamma_{U(t)}^m$ on $\mathcal{H}_c \otimes \mathcal{H}_t$ as

$$
\left(C_{U(t)}^{(m \varphi_U(t))} \otimes \mathcal{I}_{\mathcal{H}_c \otimes \mathcal{H}_t}\right)|\Psi\rangle \langle \Psi| \left(\Gamma_{U(t)}^m \otimes \mathcal{I}_{\mathcal{H}_c \otimes \mathcal{H}_t}\right) = 
\begin{pmatrix}
|\alpha|^2 |\psi\rangle \langle \psi| & \alpha \beta^* e^{-im \varphi_U(t/m)} |\psi\rangle \langle \phi| U(t) \langle \phi| U(t)^\dagger \\
\alpha^* \beta e^{im \varphi_U(t/m)} |\phi\rangle \langle \phi| U(t) \langle \phi| U(t)^\dagger & |\beta|^2 U(t) \langle \phi| \langle \phi| U(t)^\dagger
\end{pmatrix}. \hspace{1cm} (C5)$$
and

\[ \left( \Gamma^m_U(t/m) \otimes I_{H_e \otimes H_t} \right) |\Psi \rangle \langle \Psi| = \left( \begin{array}{c|c} |\alpha|^2 |\psi\rangle \langle \psi| & \alpha^* \beta \langle \phi| U(t) \phi \rangle \langle \phi| U(t) \dagger \rangle \\ \hline \langle \phi| U(t) \phi \rangle \langle \phi| U(t) \dagger | \beta|^2 \end{array} \right). \] (C6)

Here, \( U(t) \) acts on the first \( H_t \). Note that \( (\text{Tr} \left[ U \left( \frac{t}{m} \right) \right])^m = (a_U(t/m))^m e^{-i m \varphi_U(t/m)} \), where we shall assume \( e^{i m \varphi_U(t/m)} = 1 \) for simplicity. (The final consequence is the same for general cases.) Then the all energy eigenvalues of \( m \) by the following way. Since \( \text{Tr} \left[ U \left( \frac{t}{m} \right) \right] = 2 \), we first show the derivation of \( |\alpha|^2 |\psi\rangle \langle \psi| \) introduced in Eq. (8) of the main article. (We have omitted the degeneracy parameter of the energy eigenbasis for simplicity). This statement is shown by the following way. Since \( a_U(t/m) = 1 + O(1/m^2) \), the coherence factor can be sorted by the order of \( m \) as

\[ e^{-i m \varphi(t/m)} = \left( 1 - \frac{\text{Tr}[H]}{d} \right) t/m + O(1/m^2) \].

Since \( \text{Tr}[H]/d = \langle E \rangle \), we can conclude that

\[ e^{-i m \varphi(t/m)} = \left( 1 - \frac{\text{Tr}[H]}{d} \right) t/m + O(1/m) \]

\[ = e^{-i \langle E \rangle t} + O(1/m). \] (D1)

Appendix D: Concentration of the phase factor into the energy average value

The total relative phase factor \( e^{-i m \varphi(t/m)} \), which is obtained by \( m \) times the relative phase factor for each iteration \( e^{-i m \varphi(t/m)} \) introduced in Eq. (8) of the main article converges to \( \langle E \rangle t \) at \( m \to \infty \), where \( \langle E \rangle \) is the average of the all energy eigenvalues of \( H \), namely, \( \langle E \rangle = \sum E_i / d \) (We have omitted the degeneracy parameter of the energy eigenbasis for simplicity). This statement is shown by the following way. Since \( a_U(t/m) = 1 + O(1/m^2) \), the coherence factor can be sorted by the order of \( m \) as

\[ e^{-i m \varphi(t/m)} = \left( 1 - \frac{\text{Tr}[H]}{d} \right) t/m + O(1/m^2) \].

Since \( \text{Tr}[H]/d = \langle E \rangle \), we can conclude that

\[ e^{-i m \varphi(t/m)} = \left( 1 - \frac{\text{Tr}[H]}{d} \right) t/m + O(1/m) \]

\[ = e^{-i \langle E \rangle t} + O(1/m). \] (D1)

Appendix E: Probability of phase estimation algorithm

In this section, we show the derivation of the probability distributions \( P_N(2\pi f_N|\theta, t/m) \) and \( Q_N(2\pi f_N|\theta, U(t/m)) \) (Eqs. (1) and (10) in the main article) of the phase estimation algorithm. Since \( P_N(2\pi f_N|\theta, t/m) \) is a special case of \( Q_N(2\pi f_N|\theta, U(t/m)) \), we first show the derivation of \( Q_N(2\pi f_N|\theta, U(t/m)) \).

To estimate the phase in the N-digits prescession 2\pi f_N, we use a system consisting of a target system \( H_t \) and a control system consisting on \( N \)-qubit systems \( H_c^{\otimes N} \). The controlled-unitary operation of \( U^{2k} \) denoted by \( C_{U^{2k}} \) is applied on the \( k \)-th qubit of the control system.

The initial state of the control and target system is given by

\[ |0 \ldots 0 \rangle \otimes |\theta_i \rangle \langle \theta_i| \] (E1)

on \( H_c^{\otimes N} \otimes H_t \) where \( |0 \ldots 0 \rangle \otimes |0 \rangle \otimes \ldots \otimes |0 \rangle \in H_c^{\otimes N} \) is a state in the computational basis corresponding to a binary number \( 0 \ldots 0 \).

At the first step of the algorithm, the Hadamard gate is applied to each control qubit system. The state after this operation written by

\[ \frac{1}{2^N} \sum_{a_1,a_2,\ldots,a_N} |a_1 a_2 \ldots a_N \rangle \langle b_1 b_2 \ldots b_N | \otimes |\theta_i \rangle \langle \theta_i| \] (E2)

where \( a_k, b_k \in \{0, 1\} \).

At the second step, the universal controllization map of \( U^{2k} \), \( \Gamma^m_{U^{2k}(t/m)} \), is applied on the \( k \)-th control qubit and the target system for all \( 1 \leq k \leq N \). After this step, the state is transformed to

\[ \frac{1}{2^N} \sum_{a_1,a_2,\ldots,a_N} \prod_{k=1}^N (a_U(t/m))^{m2^k |a_k - b_k|} \times \exp \left( i 2^k - 1 (a_k - b_k) (\theta_i - m \varphi_U(t/m)) \right) \times |a_1 a_2 \ldots a_N \rangle \langle b_1 b_2 \ldots b_N | \otimes |\theta_i \rangle \langle \theta_i|. \] (E3)

At the last step, the inverse quantum Fourier transformation is applied and then the control qubits are measured in the computational basis. This is equivalent to perform a projective measurement in the Fourier basis \( \{ |f_N \rangle \langle f_N| \} \) on \( H_c^{\otimes N} \) where

\[ |f_N \rangle := \frac{1}{\sqrt{2^N}} \sum_{c_1,c_2,\ldots,c_N} e^{-i 2^k \pi f_N c_k} |c_1 c_2 \ldots c_N \rangle. \] (E4)

The probability of obtaining \( f_N \) by the measurement \( \{ |f_N \rangle \langle f_N| \} \) on a density operator \( \rho \) is given by \( \langle f_N | \rho |f_N \rangle \). Thus, the probability distribution \( Q_N(2\pi f_N|\theta, U(t/m)) \)
is
\[ Q_N(2\pi f_N|\theta, U(t/m)) = \]
\[ \frac{1}{2^N} \sum_{a_1, a_2, \ldots, a_N, b_1, b_2, \ldots, b_N} \prod_{k=1}^{N} (a_{U(t/m)})^{m2^k|a_k-b_k|} \]
\[ \exp \left( i2^{k-1}(a_k - b_k) \left( \theta_i + m\varphi_{U(t/m)} - 2\pi f_N \right) \right) \]
\[ = \frac{1}{2^N} \prod_{k=1}^{N} 1 + (a_{U(t/m)})^{m2^k} \]
\[ \times \cos 2^{k-1} \left( \theta_i + m\varphi_{U(t/m)} - 2\pi f_N \right). \]

The probability distribution \( P_N(2\pi f_N|\theta_i) \) corresponds to a special case of \( Q_N(2\pi f_N|\theta, U(t/m)) \) where \( a_{U(t/m)} = 1 \) and \( \varphi_{U(t/m)} = 0 \), namely,
\[ P_N(2\pi f_N|\theta_i) = \frac{1}{2^N} \prod_{k=1}^{N} \left( 1 + \cos \left( 2^{k-1} \left( \theta_i - 2\pi f_N \right) \right) \right). \]
\[ (E5) \]

This probability distribution is simplified to the form given by Eq. (2) in the main article,
\[ P_N(2\pi f_N|\theta_i) = \left( \frac{\sin 2^N (\theta_i - 2\pi f_N)/2}{2^N \sin (\theta_i - f_N)/2} \right)^2. \]
\[ (E6) \]
by using a formula
\[ \frac{1}{2^N} \prod_{k=1}^{N} (1 + \cos 2^{k-1}x) = \left( \frac{\sin 2^N x / 2}{2^N \sin x / 2} \right)^2, \]
\[ (E7) \]
which is obtained by repeatedly using
\[ 1 + \cos x = \frac{1}{2^N} \left( \frac{\sin x}{\sin \frac{x}{2}} \right)^2. \]
\[ (E8) \]
Eq. \[ (E8) \] is obtained by combining the following two formulae
\[ 1 + \cos x = 2 \cos^2 \frac{x}{2}, \] \[ \cos x = \frac{\sin 2x}{2 \sin x}. \]
\[ (E9) \]

In Fig. 4 we show the probability distributions for several different coherent factors. We can see that the sharpness of each distribution saturates at some number of the control qubits \( N \) when \( (a_{U(t/m)})^m \) is not equal to 1.

**Appendix F: Comparing runtime with brute force method**

In this section we compare the runtime of two methods to simulate the projective measurement of energy, our method based on phase estimation algorithm and universal controllization and the brute force method based on the process tomography and the diagonalization of the Hamiltonian. We regard the running time of the algorithm as the total calling time of the Hamiltonian dynamics under the assumption that the quantum computer can operate in a time scale much faster than that of the Hamiltonian dynamics of the target system.

We compare the calling times of the Hamiltonian dynamics of these two methods in compatible performance. Evaluating performance of a measurement as the projective measurement of energy is not as simple as evaluating
performance of a standard instrument \(\{I_i\}\) where outcome is simply given by an index \(i\) of an element of the instrument. The measurements we are considering here is described by \(\mathcal{M} = \{(x_i, I_i)\mid i \in X\}\) where \(x_i\) is a set of real valued outcomes, \(X\) is a countable set of indices, and \(\{I_i\}\) denotes an instrument. The ideal projective measurement of energy is described by \(\mathcal{M} = \{(E_i, I_i)\mid i \in X\}\) where a set of real valued outcomes \(\{E_i\}\) representing energy eigenvalues and \(\{I_i\}\) is an instrument that maps an input state \(\rho\) to an energy eigenstate \(|E_i\rangle\) (if the Hamiltonian is not degenerate) with probability \(\text{Tr}[I_i(\rho)]\). We use the following two functions \(R_1\) and \(R_2\) for evaluating how the measurements implemented by the two methods are “far” from the ideal projective measurement of energy. The larger values represents poorer performance. These functions are applicable even for the case that the Hamiltonian is degenerated, then in the following definitions we take account in the degeneracy of the Hamiltonian.

We define \(R_1\) that evaluates fluctuation of the outcome by

\[
R_1(\mathcal{M}|H) = \max_{\lambda,i} \sum_{j \in X} \text{Tr} [I_j(|E_i^\lambda\rangle \langle E_i^\lambda|)] (E_i - x_j)^2,
\]

where \(\lambda\) is a continuous variable which identifies a state from the eigenspace and \(|E_i^\lambda\rangle\) is the corresponding state. We define \(R_2\) that evaluates irrepeatablity of the output state of the measurement

\[
R_2(\mathcal{M}|H) = \sum_{i=1}^{n} \max_{\lambda} \left\| \sum_{j \in X} I_j(|E_i^\lambda\rangle \langle E_i^\lambda|) - |E_i^\lambda\rangle \langle E_i^\lambda| \right\|.
\]

We set \(m_{\mathcal{M}U(t/m)} = 0\) for brevity. If an ideal controlled-unitary operation \(C_{U(t)}\) is available, \(R_1\) of the projective measurement of energy based on phase estimation algorithm \(\mathcal{M}_{C_{U(t)}}\) is calculated according to Eq. (E6) by

\[
R_1(\mathcal{M}_{C_{U(t)}}|H) = \max_{k} \sum_{1 \leq f_N \geq 0} P(2\pi f_N|\theta_k) \times (E(f_N) - E_k)^2
\]

\[
= \max_{k} \sum_{1 \leq f_N \geq 0} \frac{P(2\pi f_N|\theta_k)}{t^2} (E(f_N) - E_k t)^2.
\]

We bound the two terms of the right-hand side of this equation. The first term is bounded by

\[
\sum_{f_N \in B_k} \frac{P(2\pi f_N|\theta_k)}{t^2} (E(f_N) t - E_k t)^2 \leq \varepsilon \sum_{f_N \in B_k} P(2\pi f_N|\theta_k) \leq \varepsilon.
\]

The first inequality is derived from the definition of \(B_k\) and the second inequality is derived from the property of the probability distribution. The other term is bounded as

\[
\sum_{f_N \notin B_k} \frac{P(2\pi f_N|\theta_k)}{t^2} (E(f_N) - E_k t)^2 \leq \sum_{f_N \notin B_k} \frac{4\pi^2}{t^2} \left( \frac{\sin 2^N (\theta_k - 2\pi f_N)^2}{2^{N} \sin (\theta_k - f_N)^2} \right)^2 \leq \frac{4\pi^2}{t^2} \frac{2^{N} \sin^2 (\theta_k - f_N)^2}{2^{N} \varepsilon t^4} \leq 16\pi^2 \frac{2^{N}}{2^{N} \varepsilon t^4},
\]

where the first inequality is derived from \(\Delta_{\max} t < 2\pi\) and the third from that the total number of outcomes \(f_N\) is \(2^N\). Combining these two results, we obtain

\[
R_1(\mathcal{M}_{C_{U(t)}}|H) \leq \varepsilon + \frac{16\pi^2}{2^{N} \varepsilon t^4}.
\]

On the other hand, \(R_2\) of \(\mathcal{M}_{PE}\) is calculated by

\[
R_2(\mathcal{M}_{C_{U(t)}}|H) = 0,
\]

since the total unitary matrix according to this algorithm commutes with the local Hamiltonian of the target system. For the same reason,

\[
R_2(\mathcal{M}_{PEA}|\tilde{H}) = 0.
\]

For given \(\Delta_{\max}\) and \(\varepsilon \in (0,1]\), we calculate the total running time \(T_{PE}\) (i.e., the total calling time of Hamiltonian dynamics) to obtain

\[
R_1(\mathcal{M}_{PEA}|\tilde{H}) \leq 3\varepsilon.
\]

First, choose \(t\) so that \(\Delta_{\max} t \leq \pi/2\). Next, choose \(N\) so that the second term of Eq. (F7) is equal to \(\varepsilon\), which implies that \(2^N = O((\Delta_{\max} t)^4/\varepsilon^2)\). Finally, choose \(m\) according to Eq. (13) in the main article, so that the second term in Eq. (18) of the main article is bounded as,

\[
2^N (2\delta) 4\pi^2 / t^2 \leq \varepsilon.
\]

The total running time is given by

\[
T_{PE} = (2^{N+1} - 1) \times m \times \frac{t}{m}.
\]
All in all, to achieve Eq. (10) requires
\[
T_{PE} = O \left( \frac{(\Delta_{\text{max}})^5}{\varepsilon^2} \right).
\] (F13)

Thus the running time of our method based on phase estimation and universal controllization does not depend on the dimension of the system \(d\), whereas the running time of the brute-force method depending on \(d^4\) to achieve same performance as the projectile measurement of energy.

Appendix G: Evaluation of \(\Delta_{\text{max}}\) for CUE

1. The sketch of proof

For the circular unitary ensemble (CUE), the expectation value of any power \(r\) of \((a_{U(i)})^2\) is given by using the correspondence between CUE and random permutations \(23\), i.e.,
\[
\langle (a_{U(i)})^{2r} \rangle_{\text{CUE}} \leq r!/d^{2r}.
\] (G1)

We show a proof of this statement in Subsection C2. The probability that \((a_{U(i)})^2 \geq 1/2\) is bounded as
\[
P \left\{ \langle a_{U(i)}^2 \rangle \geq 1/2 \right\} \leq \exp \left( -\frac{3}{4} \frac{(d-2)^2}{d+4} \right) + \left( \frac{\log d/2}{d} \right)^d \left( \frac{\log d/2}{d}\right)^d \left( \frac{2}{d} \right)^{d-1}
\] by using Bernstein’s inequality \(23\). Next, we consider the expectation value of \((a_{U(i/2)})^2\). If \(n_i\) is odd, the \(i\)-th eigenvalue of \(U(t/2)\) is \(e^{i\theta_{i/2}}\) and, if \(n_i\) is even, the \(i\)-th eigenvalue of \(U(t/2)\) is \(e^{-i\theta_{i/2}}\). Although the probability distribution of \(n_i\) is uniformly random, when \(\Delta_{\text{max}} t > 4\pi\) the probability depends on the value of \([\Delta_{\text{max}} t]/2\pi\). In the most biased case, they are 2/3 and 1/3 for even and odd cases, respectively. Then the average value of \(a_{U(1/2)}\) is less than or equal to 1/3. Substituting this average property to Bernstein’s inequality, the probability can be bounded as
\[
P \left\{ \langle (a_{U(1/2)})^2 \rangle \geq 1/2 \right\} \leq \exp \left( -\frac{34}{74 + 27\sqrt{2}}d/328 \right).
\] (G3)

Thus we can estimate \(\Delta_{\text{max}}\) by searching \(k'\) using the algorithm evaluating the coherence factor.

2. Calculating the expectation value of factor \((a_{U(i)})^2\) in CUE

The ensemble average of \(|\text{Tr} \{ U \}|^{2r}\) in terms of the CUE is identical to as the number of permutations on \(\{1, \ldots, r\}\) in which the length or the longest increasing subsequence is less than or equal to \(d\) \(20\). We denote the number as \(I(d, r)\) following to the reference \(20\).

By definition, \(I(d, r) = r!\) for \(r \leq d\). For \(r > d\), \(I(d, r)\) is smaller than \(r!\). Thus ensemble average of the quantity \((a_U)^{2r} = |\text{Tr} \{ U \}|^{2r}/d^2\) is calculated as
\[
\langle (a_U)^{2r} \rangle_{\text{CUE}} \begin{cases} 
= r!/d^{2r} & (r \leq d) \\
\leq r!/d^{2r} & (r > d).
\end{cases}
\] (G4)

The central moment of the \(r\)-th order of \(a_U^2\) is evaluated by
\[
\langle (a_U)^2 - \langle (a_U)^2 \rangle \rangle_{\text{CUE}} \leq \sum_{k=0}^{r} C_k \langle (a_U)^{2r-k} \rangle \langle (a_U)^k \rangle \leq \frac{r!}{d^{2r}} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = \frac{r!}{d^{2r}} e.
\] (G5)

3. Concentration inequality of factor \((a_{U(i)})^2\)

a. Extended Bernstein’s inequality

To evaluate the bound of the probabilities for obtaining \((a_{U(i)})^2 \geq 1/2\) given by \(22\) and \(23\), we use both original and extended version of Bernstein’s inequalities. Bernstein’s inequality gives a bound on the probability of the sum of independent random variables \(22\).

Theorem 1 (Bernstein’s inequality (original))
For a set of independent random variables \(\{X(i)\}_{i=1}^{n}\), let \(\sigma^2\) be the upper bound of the variances of the random variables \(\{X(i)\}_{i=1}^{n}\), namely, for all \(i\), \(\langle (X(i)^2) - \langle X(i) \rangle^2 \rangle \leq \sigma^2\). If the domain of the central moment of \(\{X(i)\}_{i=1}^{n}\) is bounded by a constant \(c\), the probability of \(\frac{1}{n} \sum_{i=1}^{n} (X(i) - \langle X(i) \rangle \geq t\) is bounded by
\[
P \left\{ \frac{1}{n} \sum_{i=1}^{n} (X(i) - \langle X(i) \rangle \geq t \right\} \leq \exp \left( -\frac{3nt^2}{2(3\sigma^2 + ct)} \right).
\] (G6)

Based on this inequality, we derive an extended version of the inequality for a single random variable for our problem.

Theorem 2 (Bernstein’s inequality (extended))
For a random variable \(X\), let \(\sigma^2\) be the variance of the random variable. If the \(r\)-th central moment of \(X\) satisfies the following inequality,
\[
\langle (X - \langle X \rangle)^r \rangle \leq e^{-2\sigma^2} (2 \leq r \leq d) \quad \text{(G7)}
\]
\[
\langle (X - \langle X \rangle)^r \rangle \leq e^{2r} (d < r), \quad \text{(G8)}
\]
where \( c \) is a constant satisfying \( 1 > c \log(1 + ct/\sigma^2) \), and the probability of \( X - \langle X \rangle \geq t \) is bounded by

\[
P \{(X - \langle X \rangle) \geq t \} \leq \exp\left( -\frac{3t^2}{2(3\sigma^2 + ct)} \right) + \frac{(c \log (1 + \frac{ct}{\sigma^2}))^{d+1}}{1 - c \log (1 + \frac{ct}{\sigma^2})} \left( \frac{1}{1 + \frac{ct}{\sigma^2}} \right) . \tag{G9}\]

In the following, we show the proof of the extended Bernstein’s inequality. We first introduce Chernoff’s inequality for a single random variable.

**Lemma 2 (Chernoff’s inequality)** For any random variable \( X \) and \( s, t > 0 \),

\[
P \{(X - \langle X \rangle) \geq t \} \leq e^{-st} \left( e^{ct} - 1 - sc \right) . \tag{G10}\]

Using Chernoff’s inequality, we prove an extended version of Bennett’s inequality, which gives a bound on the probability of a sum of independent random variables in terms of \( \sigma \) and \( c \).

**Lemma 3 (Bennett’s inequality (extended))** For a random variable \( X \), let \( \sigma^2 \) be the variances of the random variable. If the \( r \)-th central moment of \( X \) satisfies the following inequality,

\[
\langle (X - \langle X \rangle)^r \rangle \leq c^{-2} \sigma^r \quad (2 \leq r \leq d) \quad \tag{G11} \leq c^2 r! \quad (d < r), \tag{G12}\]

where \( c \) is a constant satisfies \( 1 > c \log(1 + ct/\sigma^2) \), the probability of \( X - \langle X \rangle \geq t \) is bounded by

\[
P \{(X - \langle X \rangle) \geq t \} \leq \exp\left( -\frac{\sigma^2}{c^2} h(t) \right) + \frac{(c \log (1 + \frac{ct}{\sigma^2}))^{d+1}}{1 - c \log (1 + \frac{ct}{\sigma^2})} \left( \frac{1}{1 + \frac{ct}{\sigma^2}} \right) , \tag{G13}\]

where \( h(x) = (1 + x) \log(1 + x) - x \).

To prove the extended version of Bennett’s inequality, we apply Chernoff’s inequality for each random variable \( X^{(i)} \). Then we have

\[
P \{(X - \langle X \rangle) \geq t \} \leq e^{-st} \left( e^{ct} - 1 - sc \right) . \tag{G14}\]

Then the right hand side of the inequality is bounded in terms of \( c \) and \( \sigma \) by

\[
e^{-st} \left( 1 + \sum_{r=2}^{d} \frac{\sigma^2 r! e^{-2r}}{r!} \right) + e^{-st} (sc)^d \sum_{k=1}^{\infty} (sc)^k . \tag{G15}\]

Using the identity relation

\[
\sum_{r=2}^{d} \frac{\sigma^2 r! e^{-2r}}{r!} \leq \sum_{r=2}^{\infty} \frac{\sigma^2 r! e^{-2r}}{r!} = \frac{\sigma^2}{c^2} (e^{2c} - 1 - sc) , \tag{G16}\]

and \( e^x \geq 1 + x \) for \( x \geq 0 \), we obtain

\[
P \{(X - \langle X \rangle) \geq t \} \leq \exp\left( \frac{\sigma^2}{c^2} (e^{2c} - 1 - sc) - st \right) + e^{-st} \left( \frac{(sc)^d}{1 - sc^2} \right) . \tag{G17}\]

To find the value of \( s \) that gives a proper bound, we define a function which is in the argument of the exponential of the first term

\[
f(s) := \frac{\sigma^2}{c^2} (e^{2c} - 1 - sc) - st . \tag{G18}\]

The minimum value of \( f \) is achieved when \( f'(s) = 0 \). Thus the minimum \( s \) satisfies

\[
\frac{\sigma^2}{c} (e^{2c} - 1) - t = 0 \tag{G19}\]

and we obtain

\[
s = \frac{1}{c} \log \left( 1 + \frac{ct}{\sigma^2} \right) . \tag{G20}\]

Substituting this value to the inequality \( \text{[G17]} \), we achieve Bennett’s inequality.

The extended version of Bernstein’s inequality is obtained from the extended version of Bennett’s inequality by using the inequality for the binary entropy

\[
h(x) \geq \frac{3ax^2}{2(3 + x)} . \tag{G21}\]

b. **Concentration inequality of factor \((a_U(t))^2\)**

We evaluate the bound of the probabilities for obtaining \((a_U(t))^2 \geq 1/2\) using the Bernstein inequalities. For CUE, by substituting \( \sigma^2 = 1/d^2 \), \( c = 1/d \) and \( t = 1/2 - 1/d \) into the extended Bernstein’s inequality, we obtain the bound for \( P \{(a_U(t))^2 \geq 1/2\} \) represented by \( \text{[G22]} \) due to the formula \( \text{[G26]} \).

The bound for \( P \{(a_U(t/2))^2 \geq 1/2\} \) given by \( \text{[G3]} \) is obtained by the following manner. We consider that the periodicity factor \( n_i \) of \( \theta_i + 2\pi n_i \) is uniformly independently distributed on \( n_i \in \{0, 1, \cdots, \lceil \Delta_{\text{max}} E t/(2\pi) \rceil\} \). If \( n_i \) is odd, \( i \)-th eigenvalues of \( U(t/2) \) is \( e^{i\theta_i/2} \) and if \( n_i \) is even \( i \)-th eigenvalues of \( U(t/2) \) is \( e^{-i\theta_i/2} \). Although the probability distribution of \( n_i \) is uniformly random, when \( \Delta_{\text{max}} > 4\pi \), the probabilities of the plus case and the minus case are \( 2/3 \) and \( 1/3 \), respectively, in the most biased case. The average value of \( a_U(t/2) \) is largest when the probability of plus and minus are most biased.

We define \( e^{i\beta} \) by \( \sum_i e^{i\theta_i/2}/d = e^{i\beta} a_U(t/2) \), \( \theta_i = \cos(\theta_i/2 - \varphi_i + \pi n_i) \) and \( X = \sum_i x_i/d \). Since \( a_U(t/2) = |X| \), we can evaluate the bound of the probability is by
Using $\langle X \rangle = 1/3$, we achieve
\[
P \{ (a U_{1/2})^2 \geq 1/2 \} \leq P \{ X \geq 1/\sqrt{2} \} + P \{ -X \geq 1/\sqrt{2} \}.
\] (G22)

We use Bernstein’s inequality for each term, by regarding $x_i$ is a random variable. By substituting $\sigma^2 = 8/9$ and $c = 2$, Bernstein’s inequality leads the bound given by Eq. (G23). This bound is exponentially small in terms of the dimension of the system, and becomes less than 0.02 when the system is composed of 6 qubits.

[1] S. Mukamel, *Phys. Rev. Lett.* 90, 170604 (2003).
[2] C. Jarzynski, and D. K. Wojcik, *Phys. Rev. Lett.* 92, 230602 (2004).
[3] M. Campisi, P. Talkner, and P. Hanggi, *Phys. Rev. Lett.* 105, 140601 (2010).
[4] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, *Nature* 471, 87 (2011).
[5] A. Riera, C. Gogolin, and J. Eisert, *Phys. Rev. Lett.* 108, 080402 (2012).
[6] A. Y. Kitaev, *Electr. Coll. Comput. Complex* 3, 3 (1996).
[7] Y. Aharonov, S. Massar, and S. Popescu *Phys. Rev. A* 66, 052107 (2002).
[8] I. L. Chuang and M. A. Nielsen. *J. Mod. Phys.* 44, 2455, (1997).
[9] X.-Q. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O’Brien *Nat. Commun.* 2, 413 (2011).
[10] X-Q Zhou, P. Kalasuwan, T.C. Ralph, and J.L. O’Brien, *Nature Photon.* 7, 223 (2013).
[11] Supplemental material I.
[12] Supplemental material II.
[13] L. Viola, E. Knill, and S. Lloyd, *Phys. Rev. Lett.* 82, 2417 (1999).
[14] Supplemental material III.
[15] Supplemental material IV.
[16] Supplemental material V.
[17] Supplemental material VI.
[18] E. Knill and R. Laflamme, *Phys. Rev. Lett.* 81, 5672 (1998).
[19] Supplemental material VII.
[20] M. L. Metha, *Random Matrices* (Academic, Amsterdam, 2004), Third ed., Vol. 142.
[21] J. Thompson, M. Gu, K. Modi, and V. Vedral, *arXiv:1310.2927*.
[22] John Watrous’s lecture note, Theory of Quantum Information (notes).
[23] Gabor Lugosi’s lecture note, “Concentration-of-measure inequalities”, (2009).