PHYSICAL REVIEW A, VOLUME 61, 022301

Problem of equilibration and the computation of correlation functions on a quantum computer

Barbara M. Terhal[1] and David P. DiVincenzo[2]
[1]ITF, Universiteit van Amsterdam,Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands
and CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
[2]IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

(Received 2 November 1998; revised manuscript received 16 August 1999; published 4 January 2000)

We address the question of how a quantum computer can be used to simulate experiments on quantum systems in thermal equilibrium. We present two approaches for the preparation of the equilibrium state on a quantum computer. For both approaches, we show that the output state of the algorithm, after long enough time, is the desired equilibrium. We present a numerical analysis of one of these approaches for small systems. We show how equilibrium (time-)correlation functions can be efficiently estimated on a quantum computer, given a preparation of the equilibrium state. The quantum algorithms that we present are hard to simulate on a classical computer. This indicates that they could provide an exponential speedup over what can be achieved with a classical device.

PACS number(s): 03.67.Lx, 05.30.−d, 89.80.+h, 02.70.Lq

I. LIMITS OF CLASSICAL COMPUTATION

The power of quantum computers has been demonstrated in several algorithms, of which the most striking have been Shor's factoring algorithm [1,2] and Grover's search algorithm [3]. From the very start, however, the quantum computer has also held the promise of being a simulator of physical systems. This is the content of the physical version of the Church-Turing principle proposed by Deutsch [4], which says that every finitely realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means. Thus we might expect that a universal quantum computer can be used to simulate any experiment that we could do on a real physical system. If such a simulation can be done efficiently (that is, without exponential slowdown), it is clear that this could be one of the major applications of a quantum computer. This promise seems to have been only partly fulfilled until now; it has been shown by several researchers [5,6] that a simulation of the unitary time evolution of a physical system that possesses some degree of locality (which realistic physical systems do) can be accomplished efficiently on a quantum computer. However, many quantities of interest that are determined by experiment, or by the use of classical simulation techniques, relate to open quantum systems, in particular to systems in thermal equilibrium. The thermal equilibrium (Gibbs) state (in the canonical ensemble) of a Hamiltonian H is given by

$$\rho_\beta=\sum_{m=1}^{N}\frac{e^{-\beta E_m}}{Z}|m\rangle\langle m|, \tag{1.1}$$

where $|m\rangle$ (E_m) are the eigenvectors (eigenvalues) of H. Z is the partition function,

$$Z=\sum_{m=1}^{N} e^{-\beta E_m}, \tag{1.2}$$

and $\beta=1/kT$ where k is Boltzmann's constant and T the absolute temperature in degrees Kelvin. The physical systems that concern us in this paper will have a finite-dimensional Hilbert space $\mathcal{H}$ that can be decomposed as

$$\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2\otimes\cdots\otimes\mathcal{H}_n, \tag{1.3}$$

where each $\mathcal{H}_i$ represents a small, constant Hilbert space, typically associated with some (generalized) spin or other local degree of freedom. The Hamiltonian couples these local Hilbert spaces, for example in correspondence with a d-dimensional spatial lattice, so that there is only coupling between adjacent ''spins'' on this lattice. The quantities of interest, computed in experiment or in a classical computation, are of the form

$$\mathrm{Tr}\, O_1(t_1)O_2(t_2)\cdots O_k(t_k)\rho_\beta, \tag{1.4}$$

where $O_i(t_i)$ are (possibly time-dependent) observables. Both for classical systems as well as for quantum systems, computational Monte Carlo methods have been developed to estimate correlation functions as in Eq. (1.4) [7–9]. The quantum Monte Carlo method for systems at finite temperature relies on a transformation introduced by Suzuki [10] that maps an initial quantum system on a d-dimensional lattice onto a $(d+1)$-dimensional classical system. This conversion then makes it possible to use classical computational sampling techniques to estimate correlation functions as in Eq. (1.4). There seem to be (at least) two situations when this approach runs into trouble and no good computational alternatives are available [9]: (1) the correlation functions depend explicitly on time t, and (2) the quantum system is of a fermionic nature. We will give a short explanation of why these problems are encountered.

The transformation from a classical to a quantum system is based on the generalized Trotter formula. Let $H=\sum_{i=1}^{k}H_i$ where each H_i is a Hamiltonian on a small constant Hilbert space. The Trotter formula reads

$$e^{\sigma H}=\lim_{n\to\infty}(e^{\sigma H_1/n}e^{\sigma H_2/n}\cdots e^{\sigma H_k/n})^n. \tag{1.5}$$

The partition function, Eq. (1.2) [and similarly correlation functions as in Eq. (1.4)], can be rewritten, using the Trotter formula and the identity $\Sigma_m |a^m_{i,j}\rangle\langle a^m_{i,j}| = \mathbf{1}$, where the pair of indices (i,j) labels a choice of basis, as

$$Z = \operatorname{Tr} e^{-\beta H} = \sum_{\{a_{i,j}\}} p_{\{a_{i,j}\}}, \tag{1.6}$$

where $p_{\{a_{i,j}\}}$ is a distribution over the values of the collection of variables $\{a_{i,j}\}$ and j indexes the repetitions of the factors of Eq. (1.5) from 1 to n. If the distribution is non-negative, we can write $p_{\{a_{i,j}\}} = e^{H_{eff}(\{a_{i,j}\})}$ where H_{eff} is now a classical Hamiltonian given by

$$H_{eff}(\{a_{i,j}\}) = \lim_{n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{k} \tilde{H}_i(a_{i,j}, a_{i+1,j}), \tag{1.7}$$

with $a_{k+1,j} = a_{1,j+1}$, $a_{k+1,n} = a_{1,1}$ and

$$\tilde{H}_i(a,b) = \ln[\langle a|\exp(-\beta H_i/n)|b\rangle]. \tag{1.8}$$

The distribution $p_{\{a_{i,j}\}}$ will only be non-negative when the matrix elements $\langle a|\exp(-\beta H_i/n)|b\rangle$ are positive. Thus it is important to choose the right sets of basis states $|a^m_{ij}\rangle$ to make the conversion to a classical sampling problem with a non-negative distribution. There are fermionic systems such as certain Hubbard models [9] in which it does not seem to be possible to choose such a good basis. For these systems it has turned out to be very hard to get good estimates of correlation functions by using classical Monte Carlo techniques. This problem is usually referred to as the ‘‘sign’’ problem.

When we are to compute time-dependent quantities, for example the function $f(it) = \operatorname{Tr} e^{iHt} O_1 e^{-iHt} O_2 \rho_\beta$, we need to use an imaginary time $\tau = it$ to perform the conversion of Eq. (1.5) to a classical system (we expand e^{iHt} with the Trotter formula). From the classical Monte Carlo sampling of the function $f(\tau)$ for real τ, we estimate $f(\tau)$ and then we could in principle analytically continue this function. However, we only have a finite number of samples of the function and each sample point has some inaccuracy. The errors that are introduced in estimating the Fourier components $\tilde{f}(\omega)$ from this data give rise to large fluctuations when we reconstruct $f(it)$ with the Laplace transform,

$$f(it) = \int_{-\infty}^{\infty} d\omega\, e^{-\omega t} \tilde{f}(\omega), \tag{1.9}$$

resulting in a bad approximation for the time correlation function $f(it)$.

The relevance of estimating a simple time correlation function [an example of Eq. (1.4)] such as

$$\operatorname{Tr}[A_t, B_{t'}]\rho_\beta = \langle [A_t, B_{t'}]\rangle_s, \tag{1.10}$$

where A_t and $B_{t'}$ are some Hermitian Heisenberg operators of the system, cannot be overestimated. Let us recall the many contexts in which Eq. (1.10) is used in describing experimental properties of many-particle quantum systems [11].

When $A = B = u$, where u is the displacement field of a crystal, Eq. (1.10) describes the phonon dynamics of solids as probed by inelastic neutron scattering. When A and B are the number-density operator, the dielectric susceptibility is represented; this correlation function describes a variety of other experiments, including x-ray photoemission and the so-called x-ray edge singularity. When we study the current-current response function, we obtain the electrical conductivity as described by the Kubo formula. (The density-density and current-current response functions are intimately related via the continuity equation.) Spin-dependent quantities are also of interest: with the spin-spin correlation function, information is obtained about the magnetic susceptibility and, thus, the magnon dynamics of ferromagnets and antiferromagnets, the Kondo effect, and the magnetic-dipole channel in neutron scattering. And finally, if A and B involve anomalous pair amplitudes which involve Fermion operators like $a_\downarrow(k) a_\uparrow(-k)$, the presence and dynamics of a superconducting phase can be probed.

In short, the dynamic pair correlation functions provide a window on many of the interesting quantities in experimental physics, and it would be highly desirable to have a method of obtaining estimates for these quantities by simulation on a quantum computer.

In this paper we develop an approach to tackle these problems on a quantum computer. We break the problem into two parts: First, we present an approach to prepare our quantum computer in the equilibrium state ρ_β of a given Hamiltonian (Secs. II and III). We will give two alternative routes to prepare an equilibrium state. For the first quantum algorithm we can prove that in the limit of large space and time, the algorithm will successfully produce the equilibrium state as its output. In any realistic situation we are faced with finite resources in space and time. In Secs. II G and II H we therefore present some numerical studies of the performance of the algorithm for small systems. In Sec. III we present an alternative quantum equilibration algorithm that is based on eigenvalue estimation. For this algorithm we prove as well that in the limit of large space and time equilibrium is achieved. In Sec. IV we describe a procedure for efficiently estimating quantities as in Eq. (1.4), given that the equilibrium state has been prepared.

We will not attempt to prove that our algorithms run in polynomial time even for a restricted class of quantum systems H and/or for restricted ranges of β. The equilibration problem, in its full generality, is expected to be a computationally hard problem. Even classically there is a large class of systems that exhibit a feature called frustration, for which calculating the partition function Z as in Eq. (1.2) is a $P^{\sharp}$-complete problem [12]. Also, for these systems, deciding whether the energy of the ground state is lower than some constant K is an NP-complete problem [13]. The quantum problem has an added difficulty: We cannot assume that we know the eigenvectors (and eigenvalues) of the Hamiltonian of the system that we would like to equilibrate. There is no evidence (yet) that a quantum computer can exponentially outperform a classical computer in estimating the partition function for certain *classical* systems, which would enable us to sample efficiently from the classical Gibbs distribution [14].

The quantum algorithms that we present are hard to simulate on a classical computer. In both of our equilibration algorithms we use the fact that one can implement the unitary time evolution of a local Hamiltonian on n qubits in a polynomial number of steps in n on a quantum computer [5]. A direct simulation of this procedure on a classical computer would cost exponential (in n) space and time and is therefore unrealistic. As we will show in Sec. IV, given a preparation of an equilibrium state, there exists an efficient procedure on a quantum computer to calculate (time-dependent) correlation functions. As we discussed above, there is no general efficient classical algorithm with which one can estimate time-dependent correlation functions. Our quantum algorithm provides such an algorithm for a quantum computer. Abrams and Lloyd [16] have shown that the unitary simulation of a fermionic system such as the Hubbard model, either in first or second quantization, can be performed efficiently on a quantum computer. The quantum algorithms that we will present will use this unitary evolution as a building block. Therefore these algorithms can be used to compute correlation functions for the Hubbard model on a quantum computer. This is a task for which we do not have a good classical algorithm, due to the ‘‘sign’’ problem, as we pointed out above.

We focus our efforts on quantum equilibration algorithms for Hamiltonians of which the eigenvalues and eigenvectors are not known beforehand. These are the Hamiltonians of, for example, Heisenberg models (in more than two dimensions), Hubbard models, t-J models, XY models, or many-electron Hamiltonians in quantum chemistry. On the other hand, knowing the eigenvectors and eigenvalues of a Hamiltonian, such as in the Ising model, is no guarantee that there exists an efficient (polynomial time) classical algorithm that produces the equilibrium distribution. The situation is similar for quantum algorithms; we do not know in what cases the equilibration algorithms presented in Secs. II and III give rise to a polynomial time algorithm (see also [15] for quantum algorithms for Ising-type models).

The process of equilibration is also essential in the actual realization of a quantum computer. One of the assumptions underlying the construction of a quantum computer [17] is the ability to put a physical system initially into a known state (or a thermal equilibrium state in a NMR quantum computer [18]), the computational $|00\ldots0\rangle\langle00\ldots0|$ state. The way this is done in an experimental setup is to let this state be the ground state of a natural Hamiltonian and subsequently to cool to low temperature such that the probability of being in this ground state is some constant. This natural Hamiltonian must be sufficiently simple for this equilibration to be achievable efficiently and also be sufficiently weak or tunable not to disturb the computation later on.

Markov chains in the quantum domain

Our two quantum equilibration algorithms are examples of the use of quantum Markov chains on a quantum computer. We will call a trace-preserving completely positive linear map $\mathcal{S}$ a TCP map. The algorithms that we present are described as the repeated application of a TCP map $\mathcal{S}$ on some initial state. Such a TCP map can be viewed as a generalization of a Markov matrix to the quantum domain. A classical Markov chain corresponds to the following process (cf. [7]). Let $i=1,\ldots,k$ be a set of states. We take time to be a discrete variable, taking the values $t=0,1,\ldots$. At some point in time $t=0$, we start with a probability distribution $\{p_i\geqslant 0\}_{i=1}^k$ over the states i such that $\Sigma_i p_i=1$. Through a stochastic process, which we describe with a Markov matrix M, this probability distribution is mapped onto a new probability distribution $\{p_i'\}_{i=1}^k$ at time $t=1$; i.e.,

$$p^T M = p'^T, \tag{1.11}$$

where p is the vector of probabilities at $t=0$. A homogeneous Markov chain corresponds to a chain in which M is the same matrix during all time steps. In the theory of Markov chains we study the properties of the matrix M determined by its eigenvalues and eigenvectors. Such a Markov process characterized by a matrix M can be viewed as a special kind of TCP map. The states $i=1,\ldots,k$ now correspond to a set of orthonormal states $|i\rangle$: $\langle i|j\rangle=\delta_{ij}$. At $t=0$ we start with a density matrix $\rho=\Sigma_i p_i|i\rangle\langle i|$. The TCP map corresponding to a classical Markov process maps

$$\mathcal{S}_M(\rho)=\rho', \tag{1.12}$$

where $\rho'=\Sigma_i p_i'|i\rangle\langle i|$. To give a full specification of $\mathcal{S}$ in terms of M we write

$$\begin{gathered}\mathcal{S}_M(|i\rangle\langle i|)=\sum_{j=1}^k M_{ij}|j\rangle\langle j|,\\ \mathcal{S}_M(|i\rangle\langle j|)=0.\end{gathered} \tag{1.13}$$

In this classical chain the density matrices that result from this stochastic process are all diagonal in the same basis $\{|i\rangle\}_{i=1}^k$. For a general quantum Markov chain this will not be the case. In Sec. II C we establish several basic properties of TCP maps that can form the starting point for developing a theory of quantum Markov chains in a quantum computational setting.

II. EQUILIBRATION I

A. Introduction

The canonical ensemble is the ensemble of states $\{p_i,|\psi_i\rangle\}$ or a density matrix $\rho=\Sigma_i p_i|\psi_i\rangle\langle\psi_i|$, such that ρ has a given energy-expectation value

$$\mathrm{Tr}\, H\rho=\langle E\rangle. \tag{2.1}$$

The equilibrium state in this ensemble [Eq. (1.1)] can be obtained by maximizing the von Neumann entropy of ρ under this energy constraint. Another way in which the canonical ensemble is defined is by considering the possible states of a system that is in contact with an infinite heat bath at a certain temperature T. The total energy of system and bath is constant, but bath and system exchange energy, so that the system equilibrates. This directly suggests that the way to prepare the equilibrium state on a quantum computer is to mimic this process. In considering the computational complexity of such a procedure in a straightforward simulation without optimization or shortcuts, we will have to include the space and time cost of the bath, which may be large. Also, the intuitive picture of equilibration between a weakly coupled large bath and system does not tell us anything about the rate at which this equilibration occurs. Furthermore, the equilibration process assumes a bath that is already in its equilibrium state. Can we make the bath simple enough that this bath state can be prepared efficiently? In this section we study this process of equilibration. We present an algorithm and we derive expressions that completely characterize the equilibration process in an idealized case: the coupling between the bath and the system is very small, the bath is very large, and the time of interaction is large. In order to treat this problem analytically we develop a perturbation theory in the strength of the coupling between bath and system in Sec. II D. In this perturbative regime we will see that the dynamics of our quantum Markov chain can be described by a classical Markov chain plus an additional ‘‘dephasing’’ process. Only in the idealized regime, using the perturbation theory approach, are we able to show that the algorithm gives the equilibrium state as output. We then proceed by a numerical study of the algorithm in realistic cases where the bath is of finite dimension, the strength of the interaction is nonzero, and the interaction time is limited.

B. Algorithm

Definition 1. Equilibration algorithm I.

Input parameters.

H_s	Hamiltonian of a $N=2^n$-dimensional quantum system
β	Inverse temperature
H_b	Hamiltonian of a $K=2^k$-dimensional ‘‘bath’’ quantum system
λH_{sb}	H_{sb} is the *NK*-dimensional ‘‘bath-system’’ interaction Hamiltonian; λ is the parameter that measures the strength of the interaction between bath and system
t	Interaction time between bath and system
r	Number of times the bath is refreshed in the algorithm

Define the total Hamiltonian of system and bath as

$$H=H_s\otimes\mathbf{1}_K+\mathbf{1}_N\otimes H_b+\lambda H_{sb}, \tag{2.2}$$

and the trace-preserving completely positive (TCP) map $\mathcal{S}_{\lambda,t}$ as

$$\mathcal{S}_{\lambda,t}(\rho)\equiv\mathrm{Tr}_b\, e^{iHt}\rho\otimes\rho_{b,\beta}\, e^{-iHt}. \tag{2.3}$$

(1) *Prepare system*. We prepare the n qubits in the state $|000\dots00\rangle\langle000\dots00|$.

(2) *Prepare bath*. We prepare the k qubits of the bath in their equilibrium state $\rho_{b,\beta}$ of H_b.

(3) *Evolve* system and bath for time t and discard bath, that is, perform the superoperator $\mathcal{S}_{\lambda,t}$ of Eq. (2.3).

(4) *Repeat* steps 2 and 3 r times such that

$$\|\mathcal{S}_{\lambda,t}^{r+1}(|000\dots00\rangle\langle000\dots00|) - \mathcal{S}_{\lambda,t}^{r}(|000\dots00\rangle\langle000\dots00|)\|_{tr}\leqslant\epsilon, \tag{2.4}$$

for all $r\geqslant r_0$; ϵ is some accuracy. See Appendix A for the definition of $\|\cdot\|_{tr}$. ■

We put several constraints on H_s, H_b, and H_{sb}. We will use local Hilbert spaces as in Eq. (1.3) of dimension 2 (qubits). H_s must be a ‘‘local’’ Hamiltonian. We define a *d*-local Hamiltonian on *n* qubits as one that can be expressed as

$$H_s=\sum_{i=1}^{\mathrm{poly}(n)}\mathbf{1}_{N/d}\otimes h_i, \tag{2.5}$$

where each h_i operates on a tensor product of several small qubit Hilbert spaces, whose total dimension is *d*. We will also assume that the eigenvalues of H_s are all distinct; the spectrum is nondegenerate. This will simplify the upcoming analysis. In order to treat Hamiltonians with degenerate spectra a change in the perturbation theory of Sec. II D will have to be made. We expect, however, that with that change the main result of Sec. II E, namely successful equilibration in the idealized case, will still hold. H_{sb} has the linear coupling form

$$H_{sb}=S\otimes B, \tag{2.6}$$

where both $S\in B(\mathcal{H}_s)$ and $B\in B(\mathcal{H}_b)$ are local Hamiltonians. H_b is the Hamiltonian of a system of noninteracting qubits; i.e., it is a sum of single-qubit Hamiltonians:

$$H_b=\sum_{i=1}^{k}\mathbf{1}_{K/2}\otimes h_i. \tag{2.7}$$

The bath's equilibrium state factorizes into a tensor product of qubit equilibrium states associated with each h_i:

$$\rho_{b,\beta}=\rho_{b,\beta}^1\otimes\cdots\otimes\rho_{b,\beta}^k. \tag{2.8}$$

This enables us to prepare the bath (step 2) efficiently. Appendix B shows that it will cost $2k$ elementary qubit operations to perform step 2. The locality of H_s, H_b, and H_{sb} is required in order to be able to simulate the unitary time evolution e^{iHt} in time $O(t^2/\delta)$ where δ is the accuracy with which the evolution is implemented [5,19].

We also choose

$$\langle B\rangle_b\equiv\mathrm{Tr}\, B\rho_{b,\beta}=0. \tag{2.9}$$

To understand the effect of a nonzero $\langle B\rangle_b$ we rewrite H as

$$H=(H_s+\lambda\langle B\rangle_b S)\otimes\mathbf{1}_K+\mathbf{1}_N\otimes H_b+\lambda S\otimes B', \tag{2.10}$$

with $B'=B-\langle B\rangle_b\mathbf{1}_K=0$, and thus choosing a nonzero $\langle B\rangle_b$ effectively corresponds to a change in the Hamiltonian of the system. We now discuss the last step of the algorithm, step 4. When the superoperator $\mathcal{S}_{\lambda,t}$ has the equilibrium state $\rho_{s,\beta}$ as its unique fixed point, then Eq. (2.4) for all $r \geqslant r_0$ implies that

$$\|\mathcal{S}^r_{\lambda,t}(|000\ldots 00\rangle\langle 000\ldots 00|)-\rho_{s,\beta}\|_{tr} \leqslant C_N f(\epsilon,\kappa), \tag{2.11}$$

for all $r \geqslant r_0$, where κ is the absolute value of the second largest eigenvalue of $\mathcal{S}_{\lambda,t}$. C_N is some constant polynomial in N. The function $f(\epsilon,\kappa)\to 0$ for $\epsilon \to 0$ and $\kappa \neq 1$. The functional dependence of f on κ is such that when κ increases (the equilibration slows down), f increases. For sufficiently small ϵ the equilibration process will lead to successful convergence to the equilibrium state.

There does not, however, exist a straightforward implementation of step 4. The first problem is that we would have to check the closeness of the rth and the $(r+1)$th iteration of $\mathcal{S}_{\lambda,t}$ for all $r \geqslant r_0$. In practice this has to be replaced with choosing a finite set of iterations r for which the invariance of $\mathcal{S}^r(|00\ldots 0\rangle\langle 00\ldots 0|)$ is tested. This problem is also encountered in classical Monte Carlo simulations. The second problem, which is a purely quantum phenomenon, is that by measuring $\rho_r \equiv \mathcal{S}^r_{\lambda,t}(\rho)$ we might disturb ρ_r. Thus to compare ρ_r with ρ_{r+1} we would have to run $\mathcal{S}$ again for $r+1$ times. To assemble some statistics on the difference between ρ_r and ρ_{r+1} we have to run r iterations of $\mathcal{S}$ several times. These considerations about the verification of the convergence of the equilibration process are of course not special to the use of a quantum computer; they are the same as in the equilibration of a quantum physical system in an experimental setup. Furthermore, it would be an impractical task to try to measure all the matrix elements of ρ_r; ρ_r contains an exponential amount of data of which we can extract only a polynomial amount by measurement in polynomial time. The best way to proceed is the same as what one does in classical Monte Carlo simulations [9]. If the goal of the computation is to estimate $\mathrm{Tr} O\rho_{s,\beta}$, then one computes the data points

$$\mathcal{O}_r = \mathrm{Tr}\, O\rho_r, \tag{2.12}$$

until $|\mathcal{O}_r - \mathcal{O}_{r+1}| \leqslant \epsilon$ for a sufficiently large set of iterations $r \geqslant r_0$. The same procedure can be carried out when the goal of the equilibration is to compute a time-dependent correlation function such as Eq. (1.4).

In the remainder of this section we will analyze this algorithm. In Sec. II C we give some general properties of TCP maps. In Sec. II D we discuss the non-Hermitian perturbation theory that will be the basis of the analysis of $\mathcal{S}_{\lambda,t}$ in the idealized case. In Sec. II E we derive explicit expressions for the idealized case. The idealized case is the case obtained by taking the limits $\lambda\to 0$, $k\to\infty$ and $t\to\infty$, but $\lambda^2 t$ as constant.[1] Then we can show that in this idealized case the process has a unique fixed point which is the equilibrium state. Many parameters in the quantum equilibration algorithm are not yet fixed. These are the time t, the number of repetitions r, the size of the bath k, the Hamiltonian of the bath H_b, the interaction Hamiltonian $S\otimes B$, and the interaction strength λ. In an actual realization of the quantum algorithm on a quantum computer with finite resources of space and time, we should try to choose the optimal set of values for these parameters. In Secs. II G and II H we present results from numerical simulations for small systems that tell us how the algorithm depends on this set of parameters. In these sections the following questions will be addressed:

(1) What is the influence of different choices for H_b, S, and B [Eqs. (2.6) and (2.7)]?

(2) How do the parameters r, λ, and t required for successful equilibration depend on n generically? How does k, the number of bath qubits, depend on n, the number of system qubits, for successful equilibration? Are they polynomially related?

The dynamics of open quantum systems, like the system in our algorithm that interacts with a bath, is most often studied with the use of a generalized master equation; see Fick *et al.* [21]. The exact master equation in integral form describes the time evolution of $\rho(t)=\mathcal{S}_{\lambda,t}(\rho)$ of Eq. (2.3):

$$\rho(t)=e^{-i\mathcal{L}_s t}\rho(0) - \lambda^2\int_0^t dt' \int_0^{t'} dt'' e^{-i\mathcal{L}_s(t-t')}\mathcal{M}(t',t'')\rho(t''), \tag{2.13}$$

where $\mathcal{L}$, the Liouvillian, is defined as

$$\mathcal{L}(\rho)=[H,\rho]. \tag{2.14}$$

so that $\mathcal{L}_s(\rho)=[H_s,\rho]$, etc. The operator $\mathcal{M}(t',t'')$ is the ''memory kernel,''

$$\mathcal{M}(t',t'')=\mathrm{Tr}_b\, \mathcal{L}_{sb}\, e^{-i(1-\rho_b\,\mathrm{Tr}_b)\mathcal{L}(t'-t'')}\mathcal{L}_{sb}\,\rho_b\,. \tag{2.15}$$

The form in which the master equation is most often used, however, is one in which two simplifying approximations are made: (1) the Born approximation (this relates to the weakness of the interaction parameter λ) and (2) the Markov approximation. The process described by $\mathcal{S}_{\lambda,t}$ is Markovian if we can write

$$\mathcal{S}_{\lambda,t}(\mathcal{S}_{\lambda,s}(\rho))=\mathcal{S}_{\lambda,t+s}(\rho), \tag{2.16}$$

for all $t\geqslant 0$ and $s\geqslant 0$. Note that the difference between the left- and right-hand sides of this equation is the following. On the left-hand side the environment $\rho_{b,\beta}$ [see the definition of $\mathcal{S}_{\lambda,t}$ in Eq. (2.3)] is refreshed after time t, whereas on the right-hand side the environment is kept for the whole evolution time $t+s$. Markovian behavior typically occurs when the rate at which the effect of the system on the bath is erased in the bath (in the sense of being spread throughout the bath) is much faster than the rate at which the system evolves; the system sees a ''fresh'' bath every time. In our algorithm this loss of correlations in the bath is enforced when after some time t the bath is replaced by a new bath (step 4). We would not be able to truly equilibrate a finite

---

[1] This limit is sometimes referred to as the van Hove limit [20].

system with a finite-dimensional bath if we would not keep refreshing it. Since the global evolution of bath and system is unitary, eventually we will get back to the initial unentangled state and, after tracing over the bath, to the initial state of the system (a so-called Poincaré recurrence). Whether Markovian dynamics is justified will depend on the size of the bath, the strength of the interaction, and the length of the interaction time. There are ways to make a simple but naive Markov approximation in Eq. (2.13) that lead to a master equation that fails to describe TCP dynamics [21,22]. The form of the master equation that does incorporate both the approximations and yields a physical completely positive map is the master equation in Lindblad form [23]:

$$\frac{\partial \rho}{\partial t}=-i[H_s,\rho(t)]+L\rho(t), \tag{2.17}$$

where L [24,22] can be expressed with a basis of operators F_i as

$$L\rho(t)=\frac{1}{2}\sum_{k,l=1}^{N^2-1} a_{kl}\{[F_k\rho(t),F_l^\dagger]+[F_k,\rho(t)F_l^\dagger]\}, \tag{2.18}$$

where a_{kl} is a positive semidefinite matrix. In a Lindblad equation describing the equilibration process, we expect L to depend on the system Hamiltonian H_s. The equilibrium state $\rho_{s,\beta}$ — if the algorithm is successful — should be a stationary state of the process, which implies that, since $[H_s,\rho_{s,\beta}]=0$, we must have that

$$L\rho_{s,\beta}=0. \tag{2.19}$$

Davies [24–26] has demonstrated that a process described by $\mathcal{S}_{\lambda,t}$ where the bath is an infinite-dimensional quantum system (for example, a quantum field) *does* equilibrate any quantum system in the limit where $\lambda\to 0$, $t\to\infty$, but $\lambda^2 t$ stays constant. By carefully taking a Born and Markov approximation, he derives a Lindblad equation of the form such that Eq. (2.19) is obeyed. We will perform a similar analysis here. The main point of difference is that we use a perturbative analysis of the dynamics which is only valid for small $\lambda^2 t$, but coincides in this regime with Davies' result. We furthermore obtain more explicit expressions for the dynamics in this limit.

One can write down the most general form of an operator L that obeys a quantum detailed balance condition [27], a stronger requirement that the stationarity of Eq. (2.19). Now, one might ask the following question: Could we implement this corresponding superoperator directly, without the use of a weakly coupled large bath, so as to save us time and space? We believe the answer is no, as L will depend on the eigenvectors and eigenvalues of H_s, which we do not know beforehand.

C. Some useful properties of TCP maps

In this section, we study some essential properties of the superoperator $\mathcal{S}_{\lambda,t}$ defined as in Eq. (2.3). This superoperator is a TCP map,

$$\mathcal{S}_{\lambda,t}:B(\mathcal{H}_N)\to B(\mathcal{H}_N), \tag{2.20}$$

where B is the algebra of linear operators on the Hilbert space $\mathcal{H}_N$. The set $\mathbf{TCP}[N,N]$ is the set of TCP maps $\mathcal{S}$: $B(\mathcal{H}_N)\to B(\mathcal{H}_N)$.

The elements of $B(\mathcal{H}_N)$ can be represented as $N\times N$ matrices. An alternative and convenient way to represent $B(\mathcal{H}_N)$ is as a N^2-dimensional complex vector space $\mathbf{C}^{N^2}$:

$$I:\chi\in B(\mathcal{H}_N)\to(\chi)_{ij}\in\mathbf{C}^{N^2}. \tag{2.21}$$

This representation leads to a matrix representation of a TCP map $\mathcal{S}$ on $\mathbf{C}^{N^2}$. Let A_i be the operation elements of $\mathcal{S}$, i.e.,

$$\mathcal{S}(\chi)=\sum_i A_i\chi A_i^\dagger,\quad \sum_i A_i^\dagger A_i=\mathbf{1}_N. \tag{2.22}$$

Then

$$\begin{aligned}(\chi')_{mn}=(\mathcal{S}(\chi))_{mn}&=\sum_i\sum_{k,l}(A_i)_{mk}(\chi)_{kl}(A_i^\dagger)_{ln}\\&=\sum_{k,l}\mathcal{S}_{mn,kl}(\chi)_{kl},\end{aligned} \tag{2.23}$$

with

$$\mathcal{S}_{mn,kl}=\sum_i (A_i)_{mk}(A_i^\dagger)_{ln}. \tag{2.24}$$

One can then study the eigenvectors and eigenvalues of the matrix representation of a TCP map. First, we will give three useful properties of TCP maps that follow directly from their definition.

Property 1. Let $B_{\rm pos}\in B$ be the set of positive semidefinite matrices. Let $\mathcal{S}\in\mathbf{TCP}[N,N]$. Then

$$\rho\in B_{\rm pos}\Rightarrow\mathcal{S}(\rho)\in B_{\rm pos}, \tag{2.25}$$

as $\mathcal{S}$ is (completely) positive. Let χ be an eigenvector of $\mathcal{S}$ with eigenvalue μ, $\mathcal{S}(\chi)=\mu\chi$. We have

$$\mathrm{Tr}\,\chi\neq 0\Rightarrow\mu=1, \tag{2.26}$$

as $\mathcal{S}$ is trace preserving. Let A_i be the operation elements in the decomposition of $\mathcal{S}$ as in Eq. (2.22). If χ is an eigenvector of $\mathcal{S}$ with eigenvalue μ, then $\chi^\dagger$ is also an eigenvector of $\mathcal{S}$ with eigenvalue μ^*. This follows from

$$(\mathcal{S}(\chi))^\dagger=\sum_i (A_i\chi A_i^\dagger)^\dagger=\mathcal{S}(\chi^\dagger). \tag{2.27}$$

Let $B_{\rm pos,1}$ be the set of positive semidefinite matrices that have trace 1, i.e., the density matrices. Thus property 1 implies that if a *density matrix* ρ is an eigenvector of the superoperator, it must have eigenvalue 1; that is, it is a fixed point of the map. On the basis of the TCP property of a map $\mathcal{S}$, we can also show the following.

Proposition 1. Let $\mathcal{S}\in\mathbf{TCP}[N,N]$. All eigenvalues μ of $\mathcal{S}$ have $|\mu|\leq 1$.

Proof (by contradiction). Assume χ is an eigenvector of $\mathcal{S}$ with eigenvalue $|\mu|>1$. Note that property 1 implies that χ has $\mathrm{Tr}\,\chi=0$. If χ is Hermitian, μ will be real. As χ is traceless, it must have at least one negative eigenvalue. One can always find a density matrix ρ and a small enough ϵ such that $\rho'=\rho+\epsilon\chi$ is still a density matrix. Let $\mathcal{S}$ operate r times on this density matrix. For large enough r the result $\mathcal{S}^r(\rho+\epsilon\chi)=\mathcal{S}^r(\rho)+\epsilon\mu^r\chi$ will no longer be a positive semidefinite matrix: take the eigenvector $|\psi\rangle$ of χ corresponding to the lowest (negative) eigenvalue $\lambda_{\min}$. Then

$$\langle\psi|\mathcal{S}^r(\rho)|\psi\rangle+\epsilon\mu^r\langle\psi|\chi|\psi\rangle\leqslant 1+\epsilon\mu^r\lambda_{\min} \tag{2.28}$$

will become negative for large enough r. But property 1 implies that $\mathcal{S}^r(\rho')$ is a density matrix; thus $|\mu|$ cannot be larger than 1. When χ is non-Hermitian, we reason similarly. One can find a density matrix ρ and a small enough ϵ such that $\rho'=\rho+\epsilon(\chi+\chi^\dagger)$ is a density matrix. Let $\mathcal{S}(\chi)=\mu\chi=|\mu|e^{i\phi}\chi$. Let $\lambda_{\mathrm{min,r}}$ be the smallest (and negative) eigenvalue of the traceless Hermitian matrix $e^{i\phi r}\chi+e^{-i\phi r}\chi^\dagger$. Then

$$\begin{aligned}\langle\psi|\mathcal{S}^r(\rho')|\psi\rangle&=\langle\psi|\mathcal{S}^r(\rho)|\psi\rangle\\&\quad+\epsilon|\mu|^r\langle\psi|(e^{i\phi r}\chi+e^{-i\phi r}\chi^\dagger)|\psi\rangle\\&\leqslant 1+\epsilon|\mu|^r\lambda_{\mathrm{min,r}}\end{aligned} \tag{2.29}$$

will become negative for some large r ($\lambda_{\mathrm{min,r}}$ is a quasiperiodic function of r so it cannot be small for all large r). ■

Another property about the existence of fixed points can be derived.

Proposition 2. Let $\mathcal{S}\in\mathbf{TCP}[N,N]$. $\mathcal{S}$ has a fixed point (which is a density matrix).

Proof. The set of density matrices $B_{\mathrm{pos},1}\in B(\mathcal{H}_N)$ is convex and compact. $\mathcal{S}$ is a linear continuous map and $\mathcal{S}(\rho\in B_{\mathrm{pos},1})\in B_{\mathrm{pos},1}$. Then the Markov-Kakutani theorem V.10.6 of [28] applies. ■

The existence of a fixed point does not by itself guarantee that the process described by $\mathcal{S}$ is ''relaxing,'' that is, $\lim_{r\to\infty}\mathcal{S}^r(\rho)=\rho_0$ for all ρ where ρ_0 is the fixed point. The existence of such a limit depends on whether the fixed point is unique. This following proposition proves that when there is a unique fixed point, relaxation will occur and the relaxation rate is determined by the second largest eigenvalue of $\mathcal{S}$ [29].

Proposition 3. Let $\rho_0\in B_{\mathrm{pos},1}(\mathcal{H}_N)$ be the unique fixed point of a TCP map $\mathcal{S}$. Let $\kappa=\max_{m|\mu_m\neq1}|\mu_m|$, the absolute value of the second largest eigenvalue of $\mathcal{S}$. Then for all density matrices ρ we have

$$\|\mathcal{S}^r(\rho)-\rho_0\|_{tr}\leqslant C_N\mathrm{poly}(r)\kappa^r, \tag{2.30}$$

where C_N is a constant depending on the dimension N of the system and $\mathrm{poly}(r)$ denotes some polynomial in r. Thus for all density matrices ρ

$$\lim_{r\to\infty}\|\mathcal{S}^r(\rho)-\rho_0\|_{tr}=0. \tag{2.31}$$

Proof. Let μ_i be the eigenvalues of $\mathcal{S}$. Let s be the number of distinct eigenvalues. We can bring any matrix $\mathcal{S}$ into Jordan form J by a similarity transformation M [31]:

$$\mathcal{S}=MJM^{-1}, \tag{2.32}$$

where

$$J=\sum_{i=1}^{s}(\mu_iP_i+N_i). \tag{2.33}$$

P_i are orthogonal projectors and N_i is a matrix of 1*s* above the diagonal in the ith block or N_i is the 0 matrix. When the eigenvalue μ_i is nondegenerate N_i is the 0 matrix. We therefore have $N_iN_j=0$ for $i\neq j$ and $P_iN_j=0$ for $i\neq j$. Call the unique largest eigenvalue $\mu_0=1$ and the corresponding projection P_0. As in Eq. (2.32) one can write

$$\mathcal{S}^r=MJ^rM^{-1}, \tag{2.34}$$

where J^r equals

$$J^r=\sum_{i=1}^{s}(\mu_i^rP_i+N_i'), \tag{2.35}$$

where N_i' is a nilpotent matrix in the ith block. Note that N_0' is 0, as μ_0 is unique. Let $\mathcal{S}^0$ be MP_0M^{-1} or $\mathcal{S}^0(\rho)=\rho_0$. We use $\|A\|_{tr}\leqslant\sqrt{N}\|A\|_2$. Note that $\|A\|_2$ refers to the Euclidean norm of A represented as a vector. This follows from $(\Sigma_{i=1}^N|x_i|)^2\leqslant N\Sigma_{i=1}^N|x_i|^2$ for complex numbers x_i. We have first of all

$$\|\mathcal{S}^r(\rho)-\rho_0\|_{tr}\leqslant\sqrt{N}\|(\mathcal{S}^r-\mathcal{S}^0)(\rho)\|_2. \tag{2.36}$$

This expression can be bounded with the use of the similarity transformation M to

$$\begin{aligned}\|\mathcal{S}^r(\rho)-\rho_0\|_{tr}&\leqslant\sqrt{N}|||M(J^r-J^0)M^{-1}|||_2\\&\leqslant C_{1,N}|||J^r-P_0|||_2,\end{aligned} \tag{2.37}$$

where $|||\cdot|||_2$ is defined in Appendix A and we use $\|\rho\|_2=\mathrm{Tr}\,\rho^2\leqslant1$ for density matrices. Using the expression for J^r, Eq. (2.35), we can also bound

$$|||J^r-P_0|||_2\leqslant r^NC_{2,N}\kappa^r. \tag{2.38}$$

Combining Eq. (2.37) and Eq. (2.38) gives us the desired result, Eq. (2.30). Equation (2.31) then follows as $\kappa<1$ by proposition 1. If $\mathcal{S}$ is diagonalizable, the nilpotents N_i in expression Eq. (2.35) are not present. By going through the same steps, a bound as in Eq. (2.30) can be derived without the factor $\mathrm{poly}(r)$. ■

We refer the reader to [22] for discussions and references concerning the existence of a unique fixed point and other properties of relaxation for a process that is described by a Lindblad equation, Eq. (2.17).

Finally we give a result which relates members of $\mathbf{TCP}[N,N]$ to the stochastic matrices. A real matrix M is stochastic when the entries of its columns add up to 1, i.e., $\Sigma_iM_{ij}=1$.

Proposition 4. Let $\mathcal{S}\in \mathbf{TCP}[N,N]$. $\mathcal{S}_{mm,nn}\in \mathbf{R}^+$, and, $\forall \ n$, $\Sigma_m \mathcal{S}_{mm,nn}=1$; that is, the elements $\mathcal{S}_{mm,nn}$ form an $N\times N$ stochastic matrix in the indices m and n. Also, $\forall \ n,k$, $n\neq k$, $\Sigma_m \mathcal{S}_{mm,nk}=0$.

Proof. $\mathcal{S}_{mm,nn}\in \mathbf{R}^+$ follows directly from Eq. (2.24). For the rest, we impose the unit trace condition on Eq. (2.23):

$$1=\sum_{m,k,l} \mathcal{S}_{mm,kl}\rho_{kl}. \tag{2.39}$$

This must be true for all density matrices represented by ρ. Taking $\rho_{kl}=\delta_{k,l}\delta_{k,k_0}$ gives the desired result

$$1=\sum_{m} \mathcal{S}_{mm,k_0k_0}. \tag{2.40}$$

We now separate Eq. (2.39) into diagonal and off-diagonal parts, using the Hermiticity of the density matrix ρ:

$$1=\sum_{m,k} \mathcal{S}_{mm,kk}\rho_{kk}+\sum_{m,k,l}^{k>l} (\mathcal{S}_{mm,kl}+\mathcal{S}_{mm,lk})\mathrm{Re}(\rho_{kl}) + i\sum_{m,k,l}^{k>l} (\mathcal{S}_{mm,kl}-\mathcal{S}_{mm,lk})\mathrm{Im}(\rho_{kl}). \tag{2.41}$$

The first term of Eq. (2.41) is always 1 because of Eq. (2.40). If we require Eq. (2.41) when the off-diagonal terms in ρ are $\rho_{kl}=\delta_{k,k_0}\delta_{l,l_0}\,(k>l)$, we obtain

$$\sum_{m} (\mathcal{S}_{mm,k_0l_0}+\mathcal{S}_{mm,l_0k_0})=0, \tag{2.42}$$

and setting the off-diagonal terms in ρ to $\rho_{kl}=i\,\delta_{k,k_0}\delta_{l,l_0}$ $(k>l)$ gives

$$\sum_{m} (\mathcal{S}_{mm,k_0l_0}-\mathcal{S}_{mm,l_0k_0})=0. \tag{2.43}$$

Adding these equations, we obtain the desired result

$$\sum_{m} \mathcal{S}_{mm,k_0l_0}=0, \quad k_0\neq l_0. \tag{2.44}$$

■

D. Perturbation theory

In this section we develop a perturbation thoery in the coupling λ for the superoperator $\mathcal{S}_{\lambda,t}$. The calculation will assume the diagonalizability of $\mathcal{S}_{\lambda,t}$. If all the eigenvalues of a matrix M are distinct, M is diagonalizable [31]. Therefore in many cases of interest for equilibration, this assumption for $\mathcal{S}_{\lambda,t}$ will be correct. An example of a simple superoperator that is nondiagonalizable is the following. The superoperator $\mathcal{S}$ operates on $B(\mathcal{H}_3)$ and is given by

$$\mathcal{S}(|i\rangle\langle j|)=0, \quad \text{for } i\neq j, \quad \mathcal{S}(|1\rangle\langle 1|)=|2\rangle\langle 2|,$$
$$\mathcal{S}(|2\rangle\langle 2|)=|2\rangle\langle 2|, \quad \mathcal{S}(|3\rangle\langle 3|)=|1\rangle\langle 1|. \tag{2.45}$$

The eigenvectors of $\mathcal{S}$ are $|i\rangle\langle j|$ for all $i\neq j$, the state $|2\rangle\langle 2|$ and $|1\rangle\langle 1|-|2\rangle\langle 2|$. This example shows that nondiagonalizability is not a property particular to superoperators describing quantum operations but is also found in classical Markov processes.

One can formally expand the superoperator $\mathcal{S}_{\lambda,t}$ as a power series in the coupling parameter λ,

$$\mathcal{S}_{\lambda,t}=\mathcal{S}_t^{(0)}+\lambda\mathcal{S}_t^{(1)}+\lambda^2\mathcal{S}_t^{(2)}+\lambda^3\mathcal{S}_t^{(3)}+\cdots. \tag{2.46}$$

In Sec. II E we will explicitly calculate the expressions for these expansion operators. We will show [Eqs. (2.69)–(2.72)] that condition equation (2.9) implies that $\mathcal{S}_t^{(1)}$ is zero for all t. On the basis of this expansion, we will make a perturbative expansion of the eigenvalues and eigenvectors of $\mathcal{S}_{\lambda,t}$:

$$\mu=\mu^{(0)}+\lambda\mu^{(1)}+\lambda^2\mu^{(2)}+\cdots, \tag{2.47}$$

$$\chi=\chi^{(0)}+\lambda\chi^{(1)}+\lambda^2\chi^{(2)}+\cdots. \tag{2.48}$$

Assuming that the perturbation expansion exists for this non-Hermitian operator, it will have the same structure as in the well-established procedures familiar in quantum theory for bounded Hermitian operators (see textbooks on quantum mechanics such as [30] or [31] for a more mathematical background).

In the representation of Eq. (2.24), $\mathcal{S}_t^{(0)}$ reads

$$(\mathcal{S}_t^{(0)})_{mn,kl}=(U^t)_{mk}(U^{t\dagger})_{ln}, \tag{2.49}$$

where $U=e^{iH_s}$. Unitarity of $\mathcal{S}_t^{(0)}$, as a *matrix* operator on vectors in $\mathbf{C}^{N^2}$, follows from

$$\sum_{k,l} (\mathcal{S}_t^{(0)})_{mn,kl}(\mathcal{S}_t^{(0)\dagger})_{kl,ij}=\sum_{k,l} (U^t)_{mk}(U^{t\dagger})_{ln}(U^t)_{jl}(U^{t\dagger})_{ki} =\delta_{mi}\delta_{jn}. \tag{2.50}$$

The eigenbasis of $\mathcal{S}_t^{(0)}$ is formed by the set of matrices $|n\rangle\langle m|$ where $|n\rangle$ are the eigenvectors of H_s. These eigenvectors come with eigenvalues $\mu_{nm,t}^{(0)}$:

$$\{|n\rangle\langle m|, \mu_{nm,t}^{(0)}=e^{it(E_n-E_m)}\}_{n,m=1}^{N,N}, \tag{2.51}$$

where E_n are the eigenvalues of H_s. Thus all density matrices of the form $|n\rangle\langle n|$, and mixtures of these, have degenerate eigenvalues $\mu_{nn,t}^{(0)}=1$. If the spectrum of H_s is nondegenerate (we assumed this in Sec. II B), then all other eigenvectors $|n\rangle\langle m|$ for $n\neq m$ have nondegenerate eigenvalues. These eigenvectors $|n\rangle\langle m|$ form an orthonormal set with the vector inner product on $\mathbf{C}^{N^2}$:

$$\mathrm{Tr}\,(|n\rangle\langle m|)^\dagger|k\rangle\langle l|=\delta_{nk}\delta_{ml}. \tag{2.52}$$

To carry out the perturbation theory, we switch to a ket notation for the density operators and a matrix notation for the superoperators. This will make it easier for us to perform the necessary manipulations of degenerate perturbation theory, in which the degenerate sector is isolated and a diagonalization performed within it.

We first organize the diagonal, degenerate part of this vector space to be indexed. To be specific, we introduce an orthogonal basis in this vector space such that

$$|\phi_i^{(0)}\rangle = |i\rangle\langle i|, \quad 1 \leqslant i \leqslant N, \tag{2.53}$$

$$|\phi_{i(m,n)}^{(0)}\rangle = |m\rangle\langle n|, \quad 1 \leqslant m,n \leqslant N, \quad m \neq n. \tag{2.54}$$

In the second equation the indexing i can be made consecutive by choosing

$$i(m,n) = nN + m - \frac{1}{2}n(n+1), \quad m > n,$$

$$i(m,n) = \frac{1}{2}N(N-1) + mN + n - \frac{1}{2}m(m+1), \quad n > m. \tag{2.55}$$

This organizes this new vector space into a direct-sum form $\mathbf{C}^{N^2} = \mathbf{C}_{\mathrm{D}} \oplus \mathbf{C}_{\mathrm{ND}}$, where ''D'' and ''ND'' stand for diagonal and nondiagonal (or degenerate and nondegenerate). $\mathbf{C}_{\mathrm{D}}$ has dimension N and $\mathbf{C}_{\mathrm{ND}}$ has dimension $N^2 - N$.

From the discussion above, we note that the degeneracy is lifted in lowest order by the second-order part of the superoperator $\mathcal{S}$ in the D sector, which we will denote $\mathcal{S}^{(2)}_{\mathrm{D,D}}$. Assume that $\mathcal{S}^{(2)}_{\mathrm{D,D}}$ is diagonalizable via the similarity transformation

$$M\mathcal{S}^{(2)}_{\mathrm{D,D}}M^{-1} = \tilde{\mathcal{S}}^{(2)}_{\mathrm{D,D}}, \tag{2.56}$$

where $\tilde{\mathcal{S}}^{(2)}_{\mathrm{D,D}}$ is a diagonal matrix (the tilde will denote quantities expressed in the new basis $M_{\mathrm{D}} \oplus \mathbf{1}_{\mathrm{ND}}|\phi^{(0)}\rangle$, which is in general nonorthogonal). In this new basis, the degeneracy of the diagonal terms of $\mathcal{S}$ is lifted to second order in λ (the diagonal terms can be written to second order as $\mu_i = 1 + \lambda^2\tilde{\mathcal{S}}^{(2)}_{ii}$), and since the largest off-diagonal terms in the D sector are now third order, given by

$$\lambda^3 M\mathcal{S}^{(3)}_{\mathrm{D,D}}M^{-1} = \lambda^3\tilde{\mathcal{S}}^{(3)}_{\mathrm{D,D}}, \tag{2.57}$$

the condition for the successful application of nondegenerate perturbation theory is now satisfied, assuming that no additional, accidental degeneracy occurs. (The condition is satisfied from the start in the ND sector.) Its form is essentially no different from the conventional perturbation expansion [30] for Hermitian operators. This expansion for the eigenvalues is

$$\mu_i = \mu_i^{(0)} + \lambda^2\tilde{\mathcal{S}}^{(2)}_{ii} + O(\lambda^3). \tag{2.58}$$

The form of this expansion is different depending on whether $i \in \mathrm{D}$ or $i \in \mathrm{ND}$, but only at $O(\lambda^4)$. The perturbation expansions for the eigenvectors are

$$|\phi_i\rangle = |\tilde{\phi}_i^{(0)}\rangle + \lambda \sum_{j \in \mathrm{D},\ j \neq i} |\tilde{\phi}_j^{(0)}\rangle \frac{\tilde{\mathcal{S}}^{(3)}_{ji}}{\tilde{\mathcal{S}}^{(2)}_{ii} - \tilde{\mathcal{S}}^{(2)}_{jj}} + O(\lambda^2), \quad i \in \mathrm{D}, \tag{2.59}$$

$$|\phi_i\rangle = |\phi_i^{(0)}\rangle + \lambda^2 \sum_{j \neq i} |\tilde{\phi}_j^{(0)}\rangle \frac{\tilde{\mathcal{S}}^{(2)}_{ji}}{\mu_i^{(0)} - \mu_j^{(0)}} + O(\lambda^3), \quad i \in \mathrm{ND}. \tag{2.60}$$

This expansion indicates that there is no mixing between the D and ND sectors until second order in λ. This expansion strategy will be taken up again in the numerical simulations, Sec. II G [Eq. (2.116)].

This perturbation analysis shows that the superoperator of Eq. (2.46) can be approximated by a simple one, for which the approximate eignevectors are correct to zeroth order in λ, and the eigenvalues are correct to the next nonvanishing order (λ^2). In this approximation the *D* and *ND* sectors are completely decoupled. In the *D* sector the superoperator is written as

$$(\mathcal{S}_{\lambda,t}(\rho))_{nn} \approx \sum_m P_{nm,t}\,\rho_{mm}, \tag{2.61}$$

$$P_{nm,t} = \delta_{nm} + \lambda^2(\mathcal{S}_t^{(2)})_{nn,mm}. \tag{2.62}$$

Note from proposition 4 that $P_{nm,t}$ is exactly a stochastic matrix; therefore the approximate dynamics in the *D* sector is that of a classical Markov process. The approximate dynamics in the *ND* sector is diagonal in the eigenbasis:

$$[S_{\lambda,t}(\rho)]_{nm} \approx \mu_{nm,t}\rho_{nm} = \mu^{(0)}_{nm,t}\rho_{nm} + \lambda^2(S_t^{(2)})_{nm,nm}\rho_{nm},$$

$$n \neq m. \tag{2.63}$$

So, the full expression for the approximation superoperator is

$$(S_{\lambda,t})_{nm,kl} \approx P_{nk,t}\delta_{nm}\delta_{kl} + \mu_{nm,t}(1-\delta_{nm})\delta_{nk}\delta_{ml}. \tag{2.64}$$

The simplications of Eqs. (2.61) and (2.63) make it possible to answer questions about the uniqueness of the fixed point and, in principle, the mixing properties of a repeated application of $\mathcal{S}_{\lambda,t}$, using techniques from classical Markov processes [32]. The splitting in two sectors, each having its own relaxation times, is similar to the phenomenological description of a relaxation process by means of Bloch equations or the Redfield equation [33]. This description in terms of the longitudinal relaxation time T_1 (D sector) and transversal relaxation time T_2 (ND sector) is, for example, used in NMR [33].

Of course, the ''smallness'' of the operators $\lambda^2\mathcal{S}^{(2)}, \lambda^3\mathcal{S}^{(3)}, \ldots$ compared to $\mathcal{S}^{(0)}$ will determine how fast the perturbation series converges. We will calculate the eigenvectors of $\mathcal{S}_{\lambda,t}$ to zeroth order in λ and the eigenvalues to second order in λ. The stochastic matrix $P_{nm,t}$ is determined in this approximation. The justification of this approximation will be given when we explicitly determine the expressions for $\mathcal{S}_{\lambda,t}$ in Sec. II E, where we set bounds on λ and t such that indeed λ^2 and higher-order corrections are small within some norm (for example, the $\|\cdot\|_\diamond$ norm given in [34,35]).

E. Calculation of expressions

Here we will calculate the elements of the superoperator described in the last section to lowest nontrivial order in λ (λ^2). Taking the second-order expression for P in Eq. (2.62), $Q_{nm,t}$ is defined by the expression

$$P_{nm,t}=\delta_{nm}+\lambda^2 Q_{nm,t}. \tag{2.65}$$

And taking μ of Eqs. (2.63) and using Eq. (2.51), we define $\nu_{nm,t}$ by

$$\mu_{nm,t}=e^{it(E_n-E_m)}(1+\lambda^2\nu_{nm,t}). \tag{2.66}$$

In this section we will find expressions for $Q_{nm,t}$ and $\nu_{nm,t}$ and exhibit the regime in which they give a valid description of $\mathcal{S}_{\lambda,t}$. We also show that for a large enough bath, the equilibrium state is the fixed point of the map $\mathcal{S}_{\lambda,t}$. We discuss under what conditions this fixed point is unique.

We will use operators in the Heisenberg representation. We denote such operators (for example on the system) as

$$A_t=e^{iH_s t}A\,e^{-iH_s t}. \tag{2.67}$$

The total Liouvillian $\mathcal{L}$ is defined as

$$e^{-i\mathcal{L}t}(\rho\otimes\rho_{b,\beta})=U^t(\rho\otimes\rho_{b,\beta})U^{t\dagger}. \tag{2.68}$$

One can expand the operator $e^{-i\mathcal{L}t}$ in a perturbation series in λ [21], take a partial trace over the bath, and identify the operators $\mathcal{S}_t^{(0)}=e^{-i\mathcal{L}_s t}$, $\mathcal{S}_t^{(1)}$ and $\mathcal{S}_t^{(2)}$ in Eq. (2.46):

$$\mathcal{S}_t^{(1)}=-i\,\mathrm{Tr}_b\int_0^t dt'\,e^{-i(\mathcal{L}_s+\mathcal{L}_b)(t-t')}\mathcal{L}_{sb}\,e^{-i(\mathcal{L}_s+\mathcal{L}_b)t'} \tag{2.69}$$

and

$$\mathcal{S}_t^{(2)}=-\mathrm{Tr}_b\int_0^t dt'\int_0^{t'}dt''\,e^{-i(\mathcal{L}_s+\mathcal{L}_b)(t-t')}\times\mathcal{L}_{sb}\,e^{-i(\mathcal{L}_s+\mathcal{L}_b)(t'-t'')}\mathcal{L}_{sb}\,e^{-i(\mathcal{L}_s+\mathcal{L}_b)t''}. \tag{2.70}$$

First we consider $\mathcal{S}_t^{(1)}$. We use Eq. (2.68) and Eq. (2.14) to rewrite $\mathcal{S}_t^{(1)}$ acting on $\rho\otimes\rho_{b,\beta}$ as

$$\mathcal{S}_t^{(1)}(\rho\otimes\rho_{b,\beta})=-i\lambda\,\mathrm{Tr}_b\int_0^t dt'\,e^{iH_s(t-t')}\otimes e^{iH_b(t-t')}[H_{sb},\rho_{t'}\otimes\rho_{b,\beta_{t'}}]\,e^{-iH_s(t-t')}\otimes e^{-iH_b(t-t')}, \tag{2.71}$$

where $\rho_{t'}$ is the time-evolved (with H_s) ρ and $\rho_{b,\beta_{t'}}$ is the time-evolved (with H_b) $\rho_{b,\beta}$. The equilibrium state $\rho_{b,\beta}$ is invariant under unitary evolution with $e^{iH_b t'}$ and thus $\rho_{b,\beta_{t'}}=\rho_{b,\beta}$. We then use the cyclic permutation invariance of the trace and $H_{sb}=S\otimes B$ to rewrite Eq. (2.71) as a simpler sum of two terms:

$$\mathcal{S}_t^{(1)}(\rho\otimes\rho_{b,\beta})=-i\lambda\int_0^t dt'[e^{iH_s(t-t')}S\rho_{t'}e^{-iH_s(t-t')}-e^{iH_s(t-t')}\rho_{t'}Se^{-iH_s(t-t')}]\mathrm{Tr}_b\,B\rho_{b,\beta}. \tag{2.72}$$

Then the condition equation (2.9) implies that $\mathcal{S}_t^{(1)}(\rho\otimes\rho_{b,\beta})$ is 0 for any ρ.

Let us consider the second-order term. The expression for $\mathcal{S}_t^{(2)}$ reads

$$\mathcal{S}_t^{(2)}=-e^{-i\mathcal{L}_s t}\int_0^t dt'\int_0^{t'}dt''\,[h(t'-t'')S_{-t'}S_{-t''}\rho-h(t''-t')S_{-t'}\rho S_{-t''}-h(t'-t'')S_{-t''}\rho S_{-t'}+h(t''-t')\rho S_{-t''}S_{-t'}], \tag{2.73}$$

where $h(t)$ is defined as $\langle BB_t\rangle_b$. We write

$$h(t)=\int_{-\infty}^{\infty}d\omega e^{it\omega}\tilde{h}(\omega). \tag{2.74}$$

Let S_{nm} be the matrix elements of the interaction S in this eigenbasis of H_s, $S_{nm}=\langle n|S|m\rangle$. Now we can find the expression for $Q_{mn,t}=(\mathcal{S}_t^{(2)})_{mm,nn}$. From Eq. (2.73) after integration over the variables t' and t'' and with the use of Eq. (2.74), we find

$$Q_{nm,t}=2\int_{-\infty}^{\infty}d\omega\tilde{h}(\omega)\left[\frac{|S_{mn}|^2[1-\cos t(\omega-E_n+E_m)]}{(\omega-E_n+E_m)^2}-\sum_l\frac{\delta_{nm}|S_{nl}|^2[1-\cos t(\omega-E_n+E_l)]}{(\omega-E_n+E_l)^2}\right]. \tag{2.75}$$

For the ‘‘decay factor’’ $\nu_{nm,t}$ in the ND sector we find

$$\nu_{nm,t}=\int_{-\infty}^{\infty}d\omega\tilde{h}(\omega)\left[\frac{2S_{nn}S_{mm}(1-\cos t\omega)}{\omega^2}-f(t,\omega,E_n)-f^*(t,\omega,E_m)\right], \tag{2.76}$$

with f^* the complex conjugate of f. The function f is given by

$$\mathrm{Re}\,f(t,\omega,E_n)=\sum_l\frac{|S_{ln}|^2[1-\cos t(\omega-E_n+E_l)]}{(\omega-E_n+E_l)^2} \tag{2.77}$$

and

$$\mathrm{Im}\,f(t,\omega,E_n)=\sum_l\frac{|S_{ln}|^2}{\omega-E_n+E_l}\left[1-\frac{\sin t(\omega-E_n+E_l)}{t(\omega-E_n+E_l)}\right]. \tag{2.78}$$

We will now look at the idealized case; i.e., we take the limits (remember k is the number of qubits in the bath)

$$P_{nm}^{\lambda^2 t} \equiv \lim_{\substack{t\to\infty,\lambda\to 0\\ \text{constant } \lambda^2 t}} \lim_{k\to\infty} P_{nm,t},$$

$$\mu_{nm}^{\lambda^2 t} \equiv e^{it(E_n-E_m)} \lim_{\substack{t\to\infty,\lambda\to 0\\ \text{constant } \lambda^2 t}} \lim_{k\to\infty} (1+\lambda^2 \nu_{nm,t}). \tag{2.79}$$

When the bath is infinitely large, it will have a continuous spectrum; $\tilde{h}(\omega)$ will be a smooth function. The rate of interaction vanishes, but as we take the limit $t\to\infty$, there is an effective nonzero interaction that is proportional to $\lambda^2 t$. Recall that

$$\delta(x) = \lim_{t\to\infty} \frac{1-\cos(tx)}{t\pi x^2}, \tag{2.80}$$

where $\delta(x)$ is the Dirac delta function, which is defined as $\int_{-\infty}^{\infty} dx\ \delta(x) = 1$ and, $\forall\ \ x \neq 0,\ \delta(x) = 0$. With the use of the δ function we find

$$P_{mn}^{\lambda^2 t} = \delta_{nm}\left(1-\lambda^2 t 2\pi \sum_l |S_{nl}|^2 \tilde{h}(E_n-E_l)\right) + \lambda^2 t 2\pi |S_{mn}|^2 \tilde{h}(E_n-E_m) \tag{2.81}$$

and

$$\mu_{nm}^{\lambda^2 t} = e^{it(E_n-E_m)}[1+\lambda^2 t 2\pi S_{nn} S_{mm} \tilde{h}(0) - \lambda^2 t \pi g(E_n) - \lambda^2 t \pi g^*(E_m)], \tag{2.82}$$

with

$$\mathrm{Re}\, g(E_n) = \sum_l |S_{ln}|^2 \tilde{h}(E_n-E_l) \tag{2.83}$$

and

$$\mathrm{Im}\, g(E_n) = \mathrm{P}\int_{-\infty}^{\infty} d\omega \tilde{h}(\omega) \sum_l \frac{|S_{ln}|^2}{\omega - E_n + E_l}, \tag{2.84}$$

where P is the principal value of the integral. In order to see in what regime the perturbation theory is correct, we check whether the process described by Eq. (2.81) and Eq. (2.82) corresponds to that of a TCP map. First we verify property 1 in Eq. (2.82); the eigenvalues of $|n\rangle\langle m|$ and $|m\rangle\langle n|$ are related by complex conjugation or $\mu_{nm}^{\lambda^2 t *} = \mu_{mn}^{\lambda^2 t}$. The trace-preserving property (also in property 1) is also obeyed:

$$\sum_m P_{mn}^{\lambda^2 t} = 1. \tag{2.85}$$

Complete positivity of the map implies that $P_{mn}^{\lambda^2 t}$ must be a matrix of probabilities; that is, we must have $P_{mn}^{\lambda^2 t} \geqslant 0$. Thus the first necessary condition for the validity of the perturbative approximation is

$$\textit{Condition 1}:\quad \forall\ n:\ \lambda^2 t \ll \frac{1}{2\pi \sum_l |S_{ln}|^2 \tilde{h}(E_n-E_l)}. \tag{2.86}$$

Equations (2.85) and (2.86) together ensure that $P_{mn}^{\lambda^2 t}$ is a stochastic matrix. Complete positivity also implies via proposition 1 that $|\mu_{nm}^{\lambda^2 t}| \leqslant 1$. In order that $|1+\lambda^2 t a| \leqslant 1$, where a is some complex number, we must have that $\mathrm{Re}\, a \leqslant 0$ and $\lambda^2 t \leqslant 2/|\mathrm{Re}\, a|$. This real part in Eq. (2.82) is indeed negative as $\tilde{h}(\omega)$ is positive, and we obtain a new condition:

$$\textit{Condition 2}:\quad \forall\ m,n:\quad \lambda^2 t \ll \frac{1}{\pi\left|-S_{nn}S_{mm}\tilde{h}(0) + \frac{1}{2}\sum_l |S_{ln}|^2 \tilde{h}(E_n-E_l) + \frac{1}{2}\sum_l |S_{lm}|^2 \tilde{h}(E_m-E_l)\right|}. \tag{2.87}$$

Note that this condition is quite similar to the condition in Eq. (2.86).

It is not hard to see that the stochastic matrix $P_{mn}^{\lambda^2 t}$ obeys detailed balance for the equilibrium distribution:

$$P_{mn}^{\lambda^2 t} e^{-\beta E_n} = P_{nm}^{\lambda^2 t} e^{-\beta E_m}, \tag{2.88}$$

as the equilibrium condition of the bath implies that

$$\tilde{h}(-\omega) = e^{-\beta\omega} \tilde{h}(\omega). \tag{2.89}$$

Thus the equilibrium density matrix $\rho_{s,\beta}$ is a fixed point of the idealized equilibration process. To consider whether this fixed point is unique, we note the following: If a stochastic matrix M is such that all its matrix elements $M_{ij} > 0$, then M has a unique eigenvalue equal to 1 [7]. If condition 1 is obeyed, we indeed have $P_{mn}^{\lambda^2 t} > 0$ and therefore the absolute value of the second largest eigenvalue (in the diagonal sector) is smaller than 1. For the off-diagonal sector, condition 2 says that the largest eigenvalue in the off-diagonal sector is strictly smaller than 1 in absolute value. Thus under these conditions, with proposition 3, we can conclude that the process converges to the equilibrium state. The expression of $P_{mn}^{\lambda^2 t}$ coincides with the derivation given by Davies [25] for small $\lambda^2 t$.

One can help to speed up the process in the off-diagonal sector by ‘‘dephasing’’; that is, after having the system and the bath interact for some time t, we perform the operation

$$\mathcal{D}^{\nu}(\rho_s)=\frac{1}{\nu}\sum_{u=0}^{\nu-1} e^{iH_s u}\rho_s e^{-iH_s u}, \tag{2.90}$$

which can be implemented with the assistance of an extra register in the state $(1/\sqrt{\nu})\Sigma_{u=0}^{\nu-1}|u\rangle$ which is used to condition the evolution $U=e^{iH_s s}$ and subsequently traced out. The dephasing has the effect of canceling off-diagonal terms in the eigenbasis of the system, i.e.,

$$\lim_{\nu\to\infty}\mathcal{D}^{\nu}\left(\sum_{k,l=1}^{N}\alpha_{kl}|k\rangle\langle l|\right)=\sum_{k=1}^{N}\alpha_{kk}|k\rangle\langle k|\,. \tag{2.91}$$

A complete dephasing can in general not be achieved in polynomial time in *n* (see Sec. III), and thus must be understood as an extra aid but not a solution to the equilibration problem.

From the expressions for $P_{mn}^{\lambda^2 t}$ and $\mu_{nm}^{\lambda^2 t}$ we can understand the physical picture of the interaction between bath and system. The system can make a transition from (eigen)level n to level m $(n\neq m)$; i.e., $P_{mn}^{\lambda^2 t}$ is nonzero, when S_{nm} is nonzero and $\tilde{h}(E_n-E_m)$ is nonzero. The function $\tilde{h}$ that occurs in Eq. (2.81) can be expressed as

$$\tilde{h}(\Delta E)=\lim_{K\to\infty}\sum_{l,j}^{K}\delta((\Delta E)-(\omega_l-\omega_j))|B_{lj}|^2 e^{-\beta\omega_l}/Z. \tag{2.92}$$

Therefore in order that $\tilde{h}(\Delta E=E_n-E_m)$ is nonzero, there must be at least one matching energy difference in the bath; i.e., there is an l and an j such that $|\omega_l-\omega_j|=\Delta E$ and B_{lj} is nonzero. Furthermore, the more such transitions there are, the faster the off-diagonal matrix elements decay. This confirms the intuitive picture that one might have of equilibration. Note also the similarity with the Fermi golden rule [22,36] that describes the transition probability from eigenlevel n to m in a unitary evolution that is perturbed by a time-dependent Hamiltonian.

For a finite-dimensional bath, we can express $h(t)\equiv\langle BB_t\rangle$ as

$$h(t)=\sum_{k,l} e^{it(\omega_k-\omega_l)}|B_{kl}|^2 e^{-\beta\omega_k}/Z_b\,, \tag{2.93}$$

where $B_{kl}=\langle k_b|B|l_b\rangle$ with $|l_b\rangle$ being the eigenstates of the bath Hamiltonian H_b and Z_b the partition function of the bath. Taking the limits $t\to\infty$ and $\lambda\to 0$ before letting the bath grow large leads to divergent expressions for $P_{mn}^{\lambda^2 t}$ and $\mu_{nm}^{\lambda^2 t}$, suggesting that the perturbation theory fails in this regime. This is not surprising, as the finiteness of the bath together with the limit $t\to\infty$ will lead to Poincaré recurrences (only the interaction cycle time is long due to $\lambda\to 0$).

F. Inverse quantum Zeno effect

In our numerical studies (Secs. II G and II H) we have observed a phenomenon that one might call the inverse quantum Zeno effect. It is a way of mapping an arbitrary initial state onto the completely mixed state $\mathbf{1}_N$ by interacting repeatedly and strongly with the state for a very short time. Here we will give a theoretical analysis that explains this observation. Consider the weak coupling expansion $\mathcal{S}_{\lambda,t}=\mathcal{S}_t^{(0)}+\lambda^2\mathcal{S}_t^{(2)}+O(\lambda^3)$ with $\mathcal{S}_t^{(2)}$ given as in Eq. (2.70). We expand these operators around $t=0$:

$$\mathcal{S}_{\lambda,t}(\rho)=\rho-it[H_s,\rho]+\frac{t^2\lambda^2}{2}([S\rho,S]+[S,\rho S])\langle B^2\rangle_b + O(t^2,\lambda^3 t^3). \tag{2.94}$$

In the limit $\lambda\to\infty$, but $t\to 0$, and *constant* $\lambda^2 t$, the higher-order terms $O(t^2,\lambda^3 t^3)$ will vanish. Thus we see that the fixed point of $\mathcal{S}_{\lambda,t}$ in this limit (assuming nonzero $\langle B^2\rangle_b$) must obey

$$[H_s,\rho]=0 \quad \text{and} \quad [[S,\rho],S]=0. \tag{2.95}$$

Notice that if we take the differential form of Eq. (2.94) and the prescribed limit, the equation is of the Lindblad form, Eq. (2.17). The state $\mathbf{1}_N$ certainly meets the requirements of Eq. (2.95), but is it unique? If S and H_s are such that they have no eigenspaces (except for the full space) in common, and both have a nondegenerate spectrum, we can show that $\mathbf{1}_N$ is the unique eigenvector. Equation (2.95) requires that either $[S,\rho]=0$ or $[S,\rho]$ be diagonal in the same basis as S. If $[S,\rho]=0$ but also $[H_s,\rho]=0$, then ρ can only be the state $\mathbf{1}_N$. What happens if $[S,\rho]$ is just diagonal in the same basis as S? Let $|n\rangle$ be an eigenvector of S with eigenvalue λ_n. We have for $n\neq m$

$$\langle n|\,[S,\rho]\,|m\rangle=0. \tag{2.96}$$

Rewriting this expression gives

$$\forall\; n,\, m,\, n\neq m, \quad \langle n|\,\rho\,|m\rangle(\lambda_n-\lambda_m)=0. \tag{2.97}$$

Now, because ρ is diagonal in the basis of H_s as $[H_s,\rho]=0$ and H_s and S have no eigenvectors in common, there exist n and m such that $\langle n|\rho|m\rangle\neq 0$. But the eigenvalues of S were nondegenerate, thus we obtain a contradiction. ■

When $\mathbf{1}_N$ is the unique eigenvector of this process, then, with the use of proposition 3, the repeated application as in step 4 of the *equilibration algorithm* I will eventually bring the system to the state $\mathbf{1}_N$.

We showed that for this ‘‘inverse quantum Zeno’’ effect to occur, S and H_s have to be such that they have no partial eigenspace in common and both have a nondegenerate spectrum. If we assume that S and H_s are d-local with d larger or equal to 4, then this does not impose a very strong constraint on S and H_s; the effect will occur for a generic S and H_s.

G. Specifications of the numerical simulation

The main purpose of this study is to understand the effects of bath size and the choice of bath and interaction Hamiltonians for a specific system Hamiltonian. In Table I we list some of the choices that have been made in the numerical analysis. We have randomly generated the elementary

TABLE I. Some settings in the numerical simulation.

	H_s	H_b	S	B
Dimension	$N=2,\ldots,2^4$	$K=2^2,\ldots,2^6$	N	K
Locality	$d_s=4$	$d_b=2$	4	4
Sampling scale a	1	$f(n,k,d_s,d_b)$	1	1

Hamiltonians h_i that make up H_s, H_b, and H_{sb}, Eqs. (2.5), (2.6), and (2.7), with a measure $\mathcal{M}$. We choose the diagonal elements of each h_i uniformly in $[-a,a]$, where a is sampling scale in Table I. The absolute value of the above-the-diagonal elements of h_i are chosen uniformly in $[0,a]$ and its phase is chosen uniformly in $[0,2\pi]$. The below-the-diagonal elements of h_i follow from Hermiticity. This defines $\mathcal{M}$. Note that $\mathcal{M}$ is not a unitarily invariant measure.

We take the Hamiltonians S and B as sums of all possible local two-qubit interactions ($d_s=4$ in Table I). For the Hamiltonian of the system H_s we also take a sum of all possible local two-qubit interactions. Note that this includes a set of Hamiltonians that exhibit frustration, for which we do not expect equilibration to be particularly fast.

In Sec. II E we observed that matching energy differences between bath and system are an important ingredient in the equilibration of the system, which is consistent with the intuitive picture of equilibration that was sketched in Sec. II A. However, as we do not know the eigenvalues of the system, we can only pick our bath so as to optimize the chance for matching level differences. The sampling scale of the bath $f(n,k,d_s,d_b)$ is determined by roughly optimizing these coincidences, $\Delta E_b=\Delta E_s$. This is done as follows.

Consider the density of states $p_s(E,a_s)$ of the system (the distribution of eigenvalues generated by the measure $\mathcal{M}$) and the density of states $p_b(E,a_b)$ of the bath. Here a_s is the sampling scale of the system which we set to 1 (see Table I). The quantity $[\mathrm{Tr}\,H_s]_{\mathcal{M}}$ is the mean and $[\mathrm{Tr}\,H_s^2]_{\mathcal{M}}/N$ is the variance of the distribution $p_s(E,a_s)$. The choice for $\mathcal{M}$ ensures that the distributions are symmetric around $E=0$:

$$[\mathrm{Tr}\,H_s]_{\mathcal{M}}=[\mathrm{Tr}\,H_b]_{\mathcal{M}}=0. \tag{2.98}$$

To optimize for matching we choose the variances to be equal:

$$\frac{[\mathrm{Tr}\,H_s^2]_{\mathcal{M}}}{N}=\frac{[\mathrm{Tr}\,H_b^2]_{\mathcal{M}}}{K}. \tag{2.99}$$

For large K the bath distribution will be Gaussian (central limit theorem), whereas the system distribution will be similar to a Gaussian distribution for large N (see Fig. 1). Thus, setting the variances equal brings the distributions close together.

Consider first $[\mathrm{Tr}\,H_b^2]_{\mathcal{M}}$. It is straightforward to calculate the variance of the eigenvalues of a qubit bath. Given a 2×2 Hermitian matrix m_{ij}, the eigenvalues $e_\pm=\frac{1}{2}[m_{11}+m_{22}\pm\sqrt{(m_{11}-m_{22})^2+4|m_{12}|^2}]$ have the property

$$[e_\pm^2]_{\mathcal{M}}=\frac{1}{4a_b^3}\int_{-a_b}^{a_b}dm_{11}\int_{-a_b}^{a_b}dm_{22}\int_0^{a_b}d|m_{12}|\,e_\pm^2=\frac{2a_b^2}{3}. \tag{2.100}$$

Let v_i be some ± pattern i of length k, corresponding to selecting e_+ or e_- for each qubit bath. Let E_{v_i} be an eigenvalue of the full bath, i.e., $E_{v_i}=\Sigma_{m=1}^{k}e_{v_i[m]}$ where $v_i[m]$ indicates that we select the mth bit in v_i. Then

$$\frac{[\mathrm{Tr}\,H_b^2]_{\mathcal{M}}}{K}=\frac{1}{K}\sum_{i=1}^{K}[E_{v_i}^2]_{\mathcal{M}}=\frac{2ka_b^2}{3}, \tag{2.101}$$

using $[e_{v_i[m]}e_{v_i[n]}]_{\mathcal{M}}=0$ for $m\neq n$. We calculate $[\mathrm{Tr}\,H_s^2]_{\mathcal{M}}=\Sigma_{i,j}[|(H_s)_{ij}|^2]_{\mathcal{M}}$ for $n>2$. We can write

$$\sum_{i,j}[|(H_s)_{ij}|^2]_{\mathcal{M}}=\sum_{i,j}\sum_{m=1}^{\binom{n}{2}}[|(h_m)_{ij}|^2]_{\mathcal{M}}, \tag{2.102}$$

where h_m is the mth local interaction Hamiltonian. We have used $[(h_k^*)_{ij}(h_m)_{ij}]_{\mathcal{M}}=0$. Each row of h_m has only four nonzero entries as the dimension of the local Hamiltonians d_s was set to 4. Using the fact that $[|(h_m)_{ij}|^2]_{\mathcal{M}}=\frac{1}{3}$ for all interaction terms m, we obtain

$$\frac{[\mathrm{Tr}\,H_s^2]_{\mathcal{M}}}{N}=\frac{4}{3}\binom{n}{2}. \tag{2.103}$$

For $n=1$, we have $[\mathrm{Tr}\,H_s^2]_{\mathcal{M}}/N=\frac{2}{3}$. Comparing Eqs. (2.101) and (2.103) gives the expression for a_b:

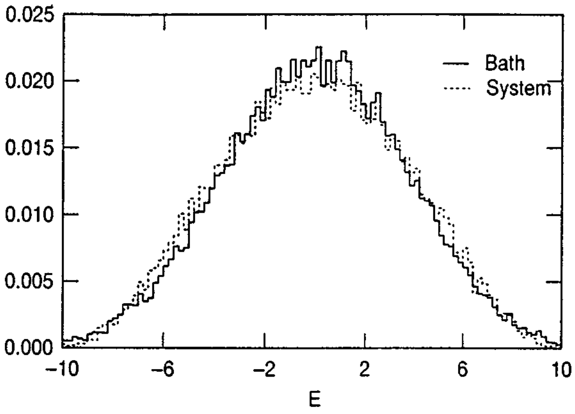

FIG. 1. A histogram (500 samples) of the density of states (unnormalized) for $N=32$ and $K=64$ with sampling scale set as Eq. (2.104).

$$a_b = f(n,k,4,2) = \sqrt{\frac{2}{k}\binom{n}{2}}. \qquad (2.104)$$

For $n=1$, $f(1,k,4,2)=\sqrt{1/k}$. Figure 1 illustrates how this setting determines the density of states of bath and system.

The numerical work consists of a calculation of the fixed point of $\mathcal{S}_{\lambda,t}$ as a function of t for a fixed λ and the second largest eigenvalue for different baths and different systems and temperatures. We follow a numerical procedure based on perturbation theory (Sec. II D) to perform a stable numerical evaluation of these quantities. Also, we expect from the analysis in Sec. II D that the small coupling regime, the realm where the perturbation theory is valid, is the regime in which we find good equilibration. When the coupling between bath and system is too strong the bath does not just exchange energy with the system, but higher-order (in the coupling) effects will bias the dynamics of the system in a way that depends on the bath. We will derive an effective coupling parameter $c(t)$ that depends on λ, but also on t, the strength of the interaction Hamiltonian H_{sb} and the energy spectrum of the bath. We can trust the answers from the numerical procedure only if we are in the regime in which perturbation theory is correct. This regime was heralded by the two conditions, Eq. (2.86) and Eq. (2.87), in Sec. II. Whether these conditions are obeyed depends on the specific choices of H_s, H_b, and S and B. We prefer to reformulate these conditions here such that they are obeyed for the average bath, system and interaction Hamiltonian obtained by sampling using $\mathcal{M}$ and the sampling scale. As the conditions are very similar, we take the first one, Eq. (2.86), and reformulate it as

$$c(t) \equiv \lambda^2 t\, 2\pi \frac{NK[S^2]_{\mathcal{M}}[B^2]_{\mathcal{M}}}{W_b} \ll 1, \qquad (2.105)$$

where $[S^2]_{\mathcal{M}}$, the average matrix element, is defined as

$$[S^2]_{\mathcal{M}} = \frac{1}{N^2}\sum_{i,j}[|S_{ij}|^2]_{\mathcal{M}} = \frac{1}{N^2}[\mathrm{Tr}_s S^2]_{\mathcal{M}}, \qquad (2.106)$$

and similarly for $[B^2]_{\mathcal{M}}$. W_b is the spectral width of the bath, i.e.,

$$W_b^2 = \frac{[\mathrm{Tr}\, H_b^2]_{\mathcal{M}}}{K}. \qquad (2.107)$$

Here we indicate the approximations made in obtaining Eq. (2.105) from condition 1 [Eq. (2.86)]:

$$\lambda^2 t\, 2\pi \sum_l |S_{ln}|^2 \tilde{h}(E_n - E_l) \ll 1. \qquad (2.108)$$

Using Eq. (2.93) and Eq. (2.74) we write the $\tilde{h}$ function as

$$\tilde{h}(E_n - E_l) = \sum_{k,m} \delta((E_n - E_l) - (\omega_k - \omega_m))|B_{km}|^2 e^{-\beta\omega_k}/Z. \qquad (2.109)$$

We will approximate the matrix elements $|B_{km}|^2$ as constants and replace them by their average $[B^2]_{\mathcal{M}}$. Then we can use density-of-states arguments to approximate the m sum over the δ functions by the inverse of the average spacing between the δ functions; this spacing is given by W_b/K:

$$\sum_m \delta((E_n - E_l) - (\omega_k - \omega_m)) \approx \frac{K}{W_b}. \qquad (2.110)$$

With these approximations, the partition-function sum over k in Eq. (2.109) becomes exactly one. So Eq. (2.109) becomes

$$\tilde{h}(E_n - E_l) \approx \frac{K[B^2]_{\mathcal{M}}}{W_b}. \qquad (2.111)$$

Now Eq. (2.108) is

$$\lambda^2 t\, 2\pi \frac{K[B^2]_{\mathcal{M}}}{W_b} \sum_l |S_{ln}|^2 \ll 1. \qquad (2.112)$$

If we again approximate the matrix elements $|S_{ln}|^2$ as constants and replace them by their average $[S^2]_{\mathcal{M}}$, and note that the l sum in Eq. (2.112) has N terms, we obtain Eq. (2.105).

For the simulations we have performed, we can find the values for $[S^2]_{\mathcal{M}}$ and $[B^2]_{\mathcal{M}}$ (note that these Hamiltonians have locality parameter $d=4$, as does the system Hamiltonian H_s) and obtain the expression

$$c(t) = \lambda^2 t \frac{16\pi}{3\sqrt{3}} \binom{k}{2} \sqrt{\binom{n}{2}} \ll 1 \qquad (2.113)$$

for $n>1$ and $k>1$. For a qubit system, $n=1$, and $k>1$ we obtain

$$c_1(t) \equiv \lambda^2 t \frac{8\pi\sqrt{2}}{3\sqrt{3}} \binom{k}{2} \ll 1. \qquad (2.114)$$

The quantity $c(t)$ in Eq. (2.105) will function as a rescaled time which depends on the strength of λ and the size of system and bath. In the regime where $c(t) \ll 1$ we expect a pertubative calculation of the eigenvectors and eigenvalues of the superoperator to be fairly accurate. The dimensionless parameter associated with the temperature is given by

$$\beta' = \beta W_s, \qquad (2.115)$$

where W_s is the spectral width of the system, Eq. (2.107) ($W_s = W_b$). From here on, β will refer to this scaled dimensionless parameter. Instead of expanding the superoperator $\mathcal{S}$ in a series in λ as in Eq. (2.46), we write

$$\lambda^2 \overline{\mathcal{S}}_t^{(2)} \equiv \mathcal{S}_{\lambda,t} - \mathcal{S}_t^{(0)}, \qquad (2.116)$$

where all higher-order terms are grouped in $\overline{\mathcal{S}}_t^{(2)}$. The calculation of eigenvalues and eigenvectors then follows the analysis of Sec. II D. We find that the choice for the bath and the interaction Hamiltonian influences whether the equilibration will succeed or not. Let

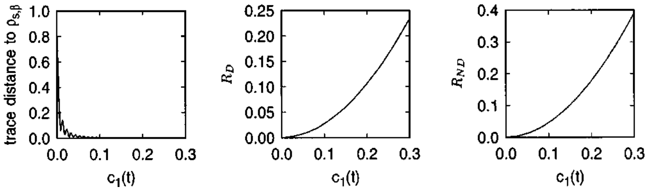

FIG. 2. An example of successful equilibration for $n=1$, $k=3$, and $\beta=3$.

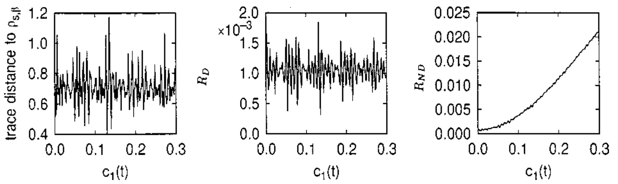

FIG. 3. An example of an unsuccesssful equilibration for $n=1$, $k=3$, and $\beta=3$.

$$\mathcal{D} \equiv \|\rho_{s,\beta} - \rho_0\|_{tr}, \tag{2.117}$$

where ρ_0 is the unit eigenvector obtained from the numerics. In Figs. 2 and 3 two extrema in dynamics are shown, each corresponding to a different choice for the system, bath, and interaction. In Fig. 2 the equilibration is successful, whereas in Fig. 3 the equilibration fails. R_{D} is defined as

$$R_{\mathrm{D}} = \frac{1-\kappa_{\mathrm{D}}}{\overline{c(t)}}, \tag{2.118}$$

where κ_{D} is the absolute value of the second largest eigenvalue in the diagonal sector and $\overline{c(t)}$ is the average coupling strength in the time interval that we consider, which is $c(t) \in [0,0.3]$ here. Similarly, we define

$$R_{\mathrm{ND}} = \frac{1-\kappa_{\mathrm{ND}}}{\overline{c(t)}} \tag{2.119}$$

for the nondiagonal sector. The qualitative difference in the behavior in Figs. 2 and 3 depends on the three requirements that we also found to be of importance in the idealized case that was treated in Sec. II E; these were the requirements for energy difference matching between bath and system, i.e., $|\Delta_{nm}| = |\Delta_{ij}|$ for system levels n and m and bath levels i and j and $|B_{ij}|^2 > 0$ and $|S_{mn}|^2 > 0$.

H. Numerical results for equilibration

We are interested in how well a randomly chosen bath and interaction equilibrate a system and how these averages are improved by choosing larger baths. As the mixing rates and the distance to the equilibrium state will in general be oscillating functions of the scaled time $c(t)$ (see Fig. 3) we will compute time averaged rates over a reasonable interval in $c(t)$,

$$[c(t_{init}) = 0, \quad c(t_{end}) = 0.5], \tag{2.120}$$

such that we are in the realm where perturbation theory is valid, Eq. (2.105). We denoted these time averages (not to be confused with bath averages) as $\overline{R_{\mathrm{D}}}$ and $\overline{\mathcal{D}}$ for the time averaged trace distance, Eq. (2.117), etc. In Fig. 4 we present histograms that show how, for a given fixed system *and* interaction, the equilibration process is different for a set of randomly chosen baths with fixed dimension. The insets show the distribution for the lowest bin. The vertical axis denotes the percentage of baths (the interval $[0\%, 100\%]$ is given as the interval $[0,1]$) for a certain distance and rate. We observe that the diagonal rate distribution is very broad, and therefore the mean of the distribution is not a very good (or a very stable) measure of the generic behavior. Furthermore, we find that the rate in the diagonal sector is much worse than in the nondiagonal sector and thus is the dominant factor in setting the mixing time. This conforms to the pattern in many quantum systems, for example for nuclear spins as observed by NMR, for which T_1 is generically larger than T_2 [33].

To study the dependence on β and on the dimension of the bath versus the dimension of the system, we compute the following data. We pick a system Hamiltonian H_s of n qubits that has some well spread out spectrum. We set the dimension of the bath and then we randomly pick both the bath Hamiltonian and the interaction Hamiltonian. Means are denoted as $[\cdot]_{\mathcal{M}b}$. For the rates we look both at the mean and the median. The median is denoted as $[[\cdot]]_{\mathcal{M}b}$; see Fig. 5.

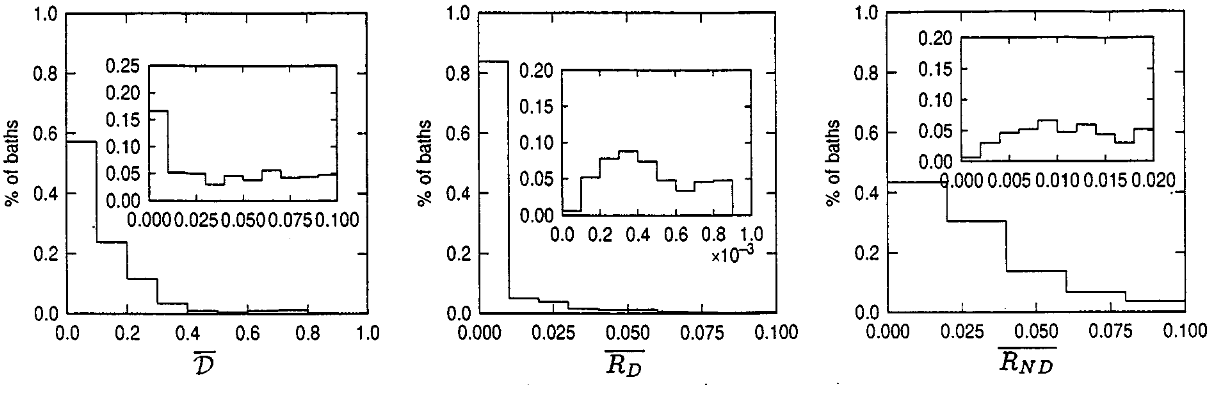

FIG. 4. An example of the distribution of baths (500 samples) for $n=2$, $k=3$, and $\beta=2$.

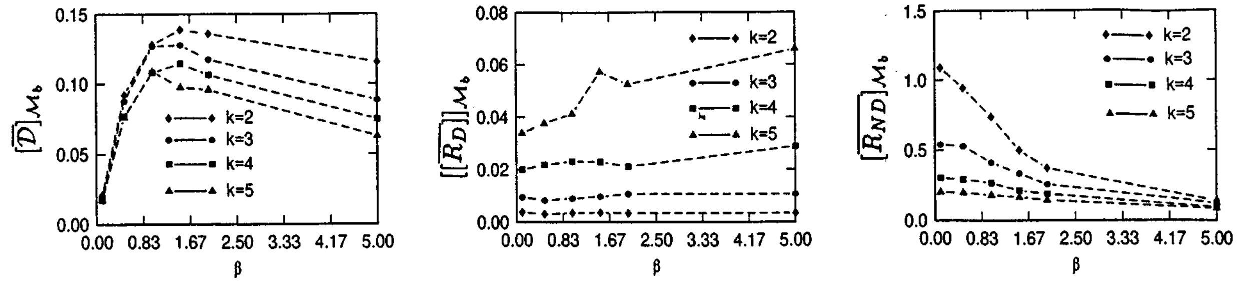

FIG. 5. Means and median for $n=1$ (500 samples).

The results for $n=1$, 2, 3, and 4 are shown in Figs. 5–8. We have given the median when the mean does not give a good representation of the distribution.

These data clearly indicate that larger baths improve the process of equilibration, both in the rates (D and ND) as well as in the closeness to the equilibrium state. The effects are the most pronounced at low temperature, where equilibration is in general harder as the system must relax to a single pure ground state. To understand the closeness scale, we show in Appendix A how far apart two arbitrarily chosen density matrices are; this number lies around 1 for the dimensions that were considered. For these estimates, we see a trend towards approximations getting worse for larger system sizes for low temperature. The scaled rates $\overline{[R_{\rm D}]_{\mathcal{M}_b}}$ and $\overline{[R_{\rm ND}]_{\mathcal{M}_b}}$ seem to be fairly constant; thus we see behavior that suggests that the rates are polynomially related to both system and bath number of qubits. We also observe that the nondiagonal rate (ND) is always higher than the diagonal rate (D). The data show a system Hamiltonian dependence; that is, the average equilibration for $n=4$ seems to be more successful than for $n=3$. We also observe that the difference between T_1 and T_2 becomes smaller with increasing β (lower temperature). Now we can try to give some answers to the question that were posed in Sec. II B [above Eq. (2.13)]. The parameters t and λ are grouped together in a single effective coupling $c(t)$. This effective coupling $c(t)$ should be small such that the perturbative approach is valid. How should one choose the value of parameter r, the number of iterations? This depends on the three quantities plotted in the figures, the trace distance, and the rate in the diagonal and the off-diagonal sector. An efficient equilibration would correspond to the process in which, while using a bath that is polynomially related in size to the system, both the rates are polynomially related to the number of system qubits and the trace distance is a small constant.

Now, if we extrapolate the data to large systems we can see that if we pick a bath size (in number of qubits) that is polynomially related to the system size (note that the number of eigenvalues is then *exponentially* related), the rates of relaxation are polynomially related to the system size (in qubits); in fact, we find that the scaled rates are constant. This behavior we do not find for the scaling of the trace distance. When we choose a polynomial relation between system and bath size, the data suggest that the relaxed state could still be fairly far away from the true equilibrium state for large system sizes.

In choosing the Hamiltonians H_b, S, and B one should try to optimize for energy matching. The numerical data show that choosing a large bath is beneficial for equilibration. The constraint that the bath consists of a set of uncoupled qubits does not seem to impose a serious restriction on the equilibration process. We believe, however, that, when H_b corresponds to a set of uncoupled qubits and the Hamiltonian B does not couple these qubits, i.e. $B=\Sigma_i B_i$ where each B_i acts on a single qubit, then the equilibration process might be somewhat impaired. The reason is that the pairs of energy levels in the bath for which $|B_{ij}|^2$ is nonzero are then restricted to the energy levels for each qubit separately. The number of matching energy-level pairs for an n-qubit bath is thus n. For a general interaction term B it will be $O(2^n)$. This can lead to quite different dynamics.

III. EQUILIBRATION II

We present an alternative to the algorithm in Sec. II. This algorithm relies on the technique for the estimation of eigen-

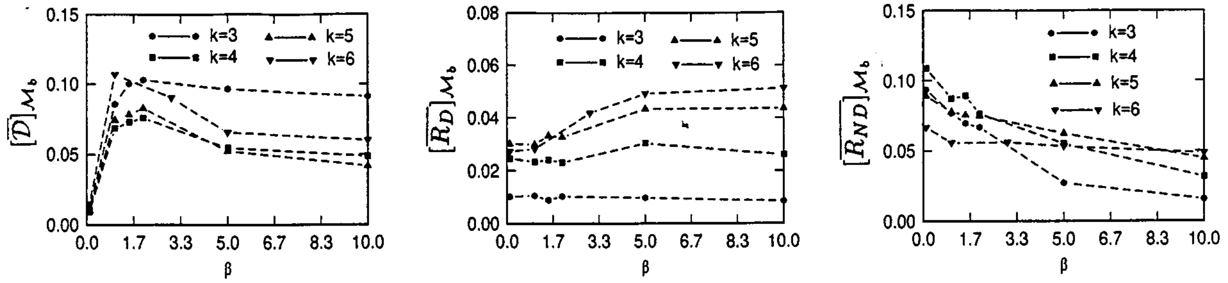

FIG. 6. Means for $n=2$ (200–500 samples).

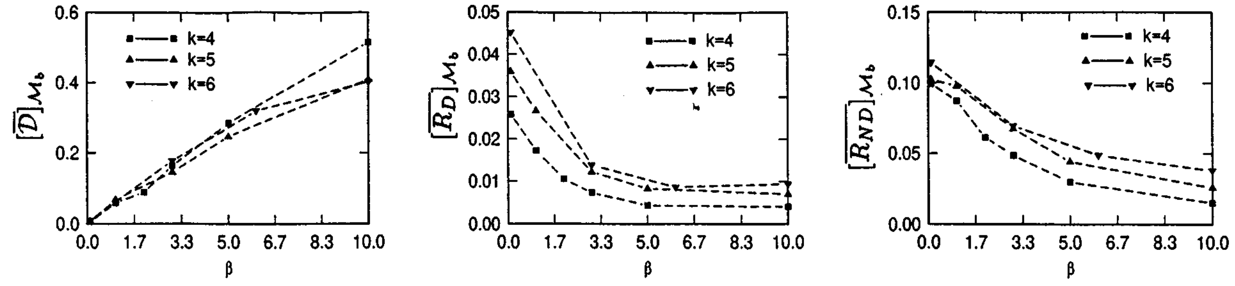

FIG. 7. Means for $n=3$ (50–100 samples).

values, originally given in [2] (see [37,38]). This eigenvalue estimation routine has also been used as a building block in an interesting quantum algorithm in [19] and [39].

Let H_s be a d-local Hamiltonian with nondegenerate eigenvalues as in Sec. II.

Definition 2. Equilibration algorithm II.

(1) *Initialize* the system in the (infinite temperature) completely mixed state $\mathbf{1}_N$. Also add one m-qubit register set to $|00\ldots00\rangle\langle 00\ldots00|$. This last register will be used to compute an m-bit estimate of an eigenvalue.

(2) *Compute eigenvalues* with the use of the Fourier transform and *dephase* in the computational eigenvalue basis.

Let U be the eigenvalue computation routine, i.e.,

$$U|n\rangle\otimes|\underbrace{00\ldots0}_{m}\rangle=|n\rangle\otimes|\underbrace{(00\ldots0)\oplus s_n}_{m}\rangle, \tag{3.1}$$

where $|s_n\rangle$ is an m-bit estimate of the eigenvalue E_n defined by $H|n\rangle=E_n|n\rangle$. The dephasing is a simple superoperator $\mathcal{D}$ on the eigenvalue register that operates as

$$\mathcal{D}(|s_i\rangle\langle s_i|)=|s_i\rangle\langle s_i|,\quad \mathcal{D}(|s_i\rangle\langle s_j|)=0. \tag{3.2}$$

The total transformation maps

$$\mathcal{D}(U(\mathbf{1}_N\otimes|00\ldots0\rangle\langle 00\ldots0|)U^\dagger)$$
$$=\sum_{n=0}^{N-1}\sum_{s=0}^{2^m-1}p(n,s)|n\rangle\langle n|\otimes|s\rangle\langle s|, \tag{3.3}$$

where $p(n,s)$ is a probability distribution, peaked at $s\sim 2^m E_n$ for large m.

(3) *Prepare* an additional N-dimensional quantum system, the bath, also in $\mathbf{1}_N$. Add a m-qubit register and one qubit register set to $|00\ldots00\rangle\langle 00\ldots00|$.

(4) *Compute eigenvalues* of the bath as for the system in step 2.

(5) *Interact* the system and bath according the following rule $\mathcal{R}$ (‘‘partial swap’’):

$$U_{\mathcal{R}}|n,m\rangle|0\rangle|s,t\rangle$$
$$=\begin{cases}|m,n\rangle|0\rangle|s,t\rangle & \text{if } t<s,\\ (p_{st}^{\beta/2}|m,n\rangle|0\rangle+\sqrt{1-p_{st}^{\beta}}|n,m\rangle|1\rangle)|s,t\rangle & \text{if } t\geqslant s,\end{cases} \tag{3.4}$$

where $p_{st}^{\beta}=e^{-\beta(t-s)}$. Here $|n\rangle$ and $|s\rangle$ are registers of the system and $|m\rangle$ and $|t\rangle$ are registers of the bath.

(6) *Trace* over the single-qubit register, all bath registers, and the eigenvalue register of the system. The system will be in some state

$$\rho_s=\sum_n \alpha_n|n\rangle\langle n|. \tag{3.5}$$

The steps 2–6 define a TCP map $\mathcal{S}$, $\mathcal{S}(\mathbf{1}_N)=\rho_s$.

(7) *Repeat* steps 2–6 r times such that

$$\|\mathcal{S}^{r+1}(|000\ldots00\rangle\langle 000\ldots00|)$$
$$-\mathcal{S}^{r}(|000\ldots00\rangle\langle 000\ldots00|)\|_{tr}\leqslant\epsilon, \tag{3.6}$$

for all $r\geqslant r_0$ and ϵ is some accuracy. ■

The advantage of this algorithm is its simplicity and its similarity to a classical algorithm; we create a Markov chain

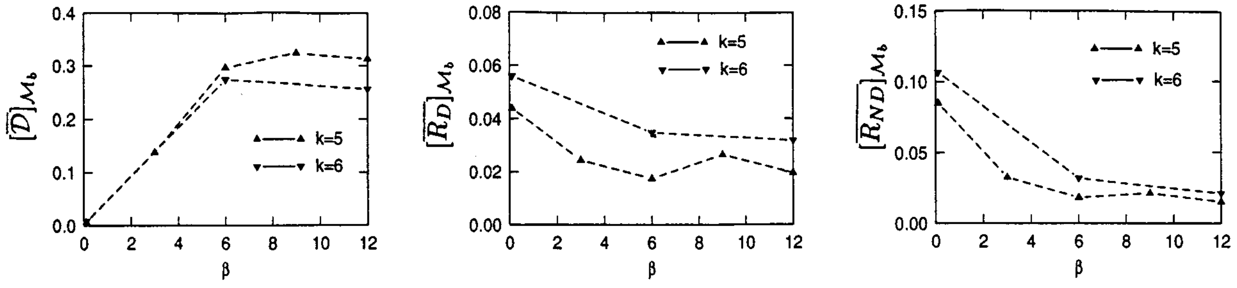

FIG. 8. Means for $n=4$ (15–20 samples).

in the eigenbasis of the system. The disadavantage of the algorithm is that it is very likely to be slow; the computation of the eigenvalues to high accuracy with the use of the Fourier transform can take exponential time in the number of qubits of the system. This routine has to be performed twice, for the system and bath, in each round of the chain. First, let us show that in the case when the eigenvalues are computed exactly in steps 2 and 4, i.e, $p(s,n)=\delta_{2^m E_n, s}/N$, the Markov chain equilibrates the system. Recall [37] that the routines of steps 2 and 4 compute rescaled eigenvalues

$$E_n' = f_1 E_n + f_2, \tag{3.7}$$

with f_1 and f_2 depending on the maximum and minimum eigenvalues (of which we assume that we can find an estimate) such that $E_n' \in [0,1)$. In the following we will drop these primes. The chain that is created can be represented as

$$\sum_n \alpha_n^{(k)} |n\rangle\langle n|, \tag{3.8}$$

where $\alpha_m^{(k)} = \Sigma_n \alpha_n^{(k-1)} P_{n\to m}$. We have

$$P_{n\to m} = \begin{cases} \dfrac{1}{N} & \text{if } E_m < E_n, \\ \dfrac{1}{N}\left[1 + \displaystyle\sum_{E_k \geq E_n} (1 - p_{nk}^{\beta})\right] & \text{if } E_m = E_n, \\ \dfrac{1}{N} p_{nm}^{\beta} & \text{if } E_m > E_n. \end{cases} \tag{3.9}$$

Note that $\Sigma_m P_{n\to m} = 1$ as required. The equilibrium state, Eq. (1.1), obeys the detailed balance condition

$$\forall \;\; n,m \;\; P_{n\to m} e^{-\beta E_n} = P_{m\to n} e^{-\beta E_m}. \tag{3.10}$$

All the matrix elements of the Markov matrix $P_{n\to m}$ are nonzero. Therefore the chain will have a unique fixed point which is equal to the equilibrium state due to detailed balance. Thus for all probability distributions α_n we have

$$\lim_{k\to\infty} \sum_n \alpha_n P_{n\to m}^{(k)} = \frac{e^{-\beta E_m}}{Z}. \tag{3.11}$$

Notice that it is not hard to prepare the initial states of the system and bath. One way to make the completely mixed state $\mathbf{1}_N$ is to make a maximally entangled state $(1/\sqrt{N})\Sigma_{i=0}^{N-1}|i\rangle|i\rangle$ and trace over the second register. This takes $O(n)$ steps. The partial swap in step 5 can be implemented with $O(n)$ elementary qubit steps. The dephasing in step 2 is introduced to keep the form of the algorithm clean, but it does not affect its output. This dephasing is implemented by measuring the eigenvalue register in the computational basis and discarding its answer. When using an m-bit eigenvalue register the joint probability $p(n,s)$ in the first round (after step 2) is equal to

$$p(n,s) = \frac{1}{N}\left|\frac{1}{2^m}\sum_{l=0}^{2^m-1} e^{2\pi i l(E_n - s/2^m)}\right|^2. \tag{3.12}$$

When $p(n,s)$ is not a δ function on the eigenvalue, the Markov chain will still be in the eigenbasis of the system; it will be a concatenation of chains. The transition probability of this new chain is

$$P'_{n\to m} = \sum_{s,t} p(s\,|\,n) P_{s\to t}\, p(m\,|\,t), \tag{3.13}$$

where $p(s|n)$ is a conditional probability, defined by $p(n,s)=p(s|n)p(n)$, and $P_{s\to t}$ is the exact chain [when $p(s|n)=\delta_{2^m E_n, s}$]. Note that $\Sigma_s p(s|n)=1$, so that $P'_{n\to m}$ is a stochastic matrix. Let us make a few remarks about the behavior of such an approximate equilibration process. If this new Markov chain is close to the exact Markov chain, we can bound the deviation from the exact fixed point with perturbation theory [40]. Let

$$P'_{n\to m} = P_{n\to m} + E_{nm}, \tag{3.14}$$

where E_{nm} is a deviation matrix defined by Eq. (3.14). Let $\rho_\Delta = \rho'_{s,\beta} - \rho_{s,\beta}$ where $\rho'_{s,\beta}$ is the fixed point of the Markov chain P'_{nm}. Assume that P is diagonalizable. Let Y be the matrix defined as

$$Y = (\mathbf{1} - P + P^{(\infty)})^{-1} - P^{(\infty)}, \tag{3.15}$$

where $P^{(\infty)}$ is the infinite iteration of P. We can write $P^{(\infty)} = \mathrm{diag}(1,0,\ldots,0)$ in the basis where the stationary state $\rho_{s,\beta}$ is an eigenvector. In this basis, with diagonalizability, P is of the form $\mathrm{diag}(1,\lambda_2,\ldots,\lambda_N)$. We can then write

$$Y = \mathrm{diag}\left(0, \frac{1}{1-\kappa}, \ldots, \frac{1}{1-\lambda_N}\right), \tag{3.16}$$

where κ is the the absolute value of second largest eigenvalue. For later use we note that the norm $|||Y|||_2 = 1/|1-\kappa|$. It is possible to write the deviation ρ_Δ in terms of Y and E:

$$\rho_\Delta = (\mathbf{1} - YE)^{-1} YE \rho_{s,\beta}, \tag{3.17}$$

when E is small enough such that $\mathbf{1} - YE$ is invertible. This expression can be derived from $P^{(\infty)}\rho_\Delta = 0$, which follows from the uniqueness of the stationary state ρ. We now use

$$\|\rho_\Delta\|_{tr} \leq \sqrt{N}\|\rho_\Delta\|_2, \tag{3.18}$$

as in proposition 3. Then using the expression for Y, Eq. (3.17) and Eq. (3.18) [see also below Eq. (2.37)] we can bound

$$\begin{aligned} \|\rho_\Delta\|_{tr} &\leq C_N \,\mathrm{Tr}\, \rho_{s,\beta}^2 \left(1 - \frac{|||E|||_2}{|1-\kappa|}\right)^{-1} \frac{|||E|||_2}{|1-\kappa|} \\ &\leq C_N \left(1 - \frac{|||E|||_2}{|1-\kappa|}\right)^{-1} \frac{|||E|||_2}{|1-\kappa|}. \end{aligned} \tag{3.19}$$

Thus the size of the correction ρ_Δ will be determined by the strength of the perturbation $|||E|||_2$ and the rate of convergence of the original Markov chain P.

For a general H_s, the computation of an m-bit approximation of the eigenvalues can cost an exponential (in m) number of elementary gates. As there are 2^n eigenvalues, knowing the $m=\log_2 \mathrm{poly}(n)$ bits of the values of E_n still leaves groups of $2^n/\mathrm{poly}(n)$ eigenvalues indistinguishable. Thus only in very special cases, if the gates $U_s^{2^m}$ can be implemented with a polynomial (in m) number of elementary steps (as in Shor's factoring algorithm [1]), is it possible to compute the eigenvalues to high accuracy efficiently.

We have demonstrated a way to set up a Markov chain on a quantum computer that will converge to the equilibrium state for long enough time. For special Hamiltonians, there might be more efficient ways to tune and modify this kind of algorithm. The rule $\mathcal{R}$ might be chosen to depend on other features of the eigenstates $|n\rangle$ and $|m\rangle$ as in the classical Metropolis algorithm where transitions are made between states that are related by local spin flips. There might be Hamiltonians for which the calculation of an eigenvalue, given the eigenvector, is efficient.

IV. (TIME-DEPENDENT) OBSERVABLES

Given that we have prepared n qubits in the equilibrium state corresponding to a certain Hamiltonian H_s, we can then proceed by experimenting and measuring. The simplest measurement that we could try to perform is the estimation of the expectation value of a d-local (Hermitian) observable O:

$$\langle O\rangle_s = \mathrm{Tr}\,\rho_{s,\beta}\,O. \tag{4.1}$$

As O is local, we write $O=\Sigma_{i=1}^{\mathrm{poly}(n)}O_i$ where each operator O_i operates on a Hilbert space of constant dimension d. We can calculate the eigenvectors and eigenvalues of each O_i rapidly on a (possibly) classical computer, which takes $\mathrm{poly}(n,c)$ operations. If O_i has eigenvalues μ_i that are both smaller as well as larger than zero, we define O_i^+ as

$$O_i^+ = \frac{1}{\max\limits_k \mu_k + |\min\limits_k \mu_k|}\,(O_i + |\min\limits_k \mu_k|\mathbf{1}) \tag{4.2}$$

such that O_i^+ is positive semidefinite and has eigenvalues smaller than or equal than 1. If O_i has only positive or zero eigenvalues, we just ''normalize'' the operator by dividing by $\max_k \mu_k$, and similarly if O has only negative eigenvalues. Let I be a positive operator valued measurement (POVM [41]) with operation elements $A_{1,i}$ and $A_{2,i}$ and corresponding outcomes 1 and 2 such that

$$E_{1,i} = A_{1,i}^\dagger A_{1,i} = O_i^+, \quad E_{2,i} = A_{2,i}^\dagger A_{2,i} = \mathbf{1} - O_i^+. \tag{4.3}$$

This measurement will give outcome 1 with probability

$$p_{1,i} = \mathrm{Tr}\,O_i^+\rho, \ \text{ etc.} \tag{4.4}$$

The operators $A_{1,i}$ and $A_{2,i}$ are given by

$$A_{1,i} = U_i(\mathrm{diag}_{O_i^+})^{1/2}\,U_i^\dagger \quad \text{and}$$

$$A_{2,i} = U_i(\mathbf{1} - \mathrm{diag}_{O_i^+})^{1/2}\,U_i^\dagger, \tag{4.5}$$

where $\mathrm{diag}_{O_i^+}$ is the diagonal form of O_i^+ and U_i the diagonalizing matrix. We summarize these results in a proposition.

Proposition 5. The estimation of $\mathrm{Tr}\,\rho O$ where O is a d-local observable with precision δ and error probability ϵ and $\rho \in B_{\mathrm{pos},1}(\mathcal{H}_N)$ $(N=2^n)$ takes $TO(\ln(1/\epsilon)/\delta^2)\mathrm{poly}(n,c)$ operations where T is the time to prepare the state ρ.

Proof. All commuting observables O_i can be measured once for a single preparation of ρ. To estimate a probability p with precision δ and error probability ϵ we need $O(\ln(1/\epsilon)/\delta^2)$ samples [42]. ■

More interesting is an algorithm to estimate time-dependent expectation values. Let O_1 and O_2 be two d-local observables. We consider how to estimate a time-dependent quantity [identical to Eq. (1.10)]

$$\mathrm{Tr}\,\rho_\beta[O_1, O_{2t}], \tag{4.6}$$

where O_{2t} is in the Heisenberg representation. Notice that O_{2t}, the time-evolved operator, will for general t *not* be local. Thus we cannot use proposition 5. The way these quantities come about in linear response theory [11] provides the key for how to estimate them on a quantum computer. One considers a system that is perturbed at some initial time $t=0$: its time evolution is generated by the perturbed Hamiltonian $H_s + \lambda O_1(t)$ $[O_1(t<0)=0]$ and λ is small. After time t we consider the response of the system to the perturbation by measuring another observable O_2. Notice that with proposition 5, it is simple to perform this experiment. Linear response means that we take into account corrections of order λ, but no higher order, in the estimation of

$$\delta\langle O_2\rangle_s = \mathrm{Tr}\,O_2\rho_t - \mathrm{Tr}\,O_2\rho_\beta, \tag{4.7}$$

where ρ_t is the time-evolved system density matrix. This first-order correction takes the form [36]

$$\delta\langle O_2\rangle_s \approx i\lambda \int_0^t dt'\ \mathrm{Tr}\,\rho_\beta[O_1(t'), O_{2t-t'}]. \tag{4.8}$$

If the disturbance $O_1(t) = O_1\delta(t=0)$, we find on the right-hand side the correlation function of Eq. (4.6). The quantity of Eq. (4.6) is interesting, because it can be used to compute the simplest reponse of the system, the linear response of Eq. (4.8), which we can directly estimate on our quantum computer, provided that both O_1 and O_2 are local. But we are of course not restricted to a linear response regime: λ is a parameter that we can tune freely. A sequence of measurements could determine higher response functions that will involve quantities such as

$$\langle O_{1t_1}O_{2t_2}O_{3t_3}\cdots O_{kt_k}\rangle_s. \tag{4.9}$$

V. CONCLUSION

It seems that by asking the question of how fast real quantum systems equilibrate, we have opened a Pandora's box of hard-to-answer questions. If there are many simple quantum systems in nature that equilibrate slowly (that is, *not* in polynomial time) by any dynamics that does not require extensive preknowledge of the system, then it would be unreasonable to ask our quantum computer to perform this task efficiently. By relaxation in polynomial time we mean the following: in polynomial time in n and $1/\epsilon$ we obtain a state that is within ϵ trace distance of the equilibrium state where ϵ is a small constant. It might be the case that leaving aside the classical phenomenon of frustration, relaxation does *not* take place in polynomial time. The idea here is that for a quantum system, the eigenbasis is not known beforehand, but must be singled out on the basis of an estimation of the eigenvalues, which is generically a hard problem.

This, however, is not in contradiction with physical and experimental reality as we know it, as the quantities that are measured in an experimental setup usually involve operators on a small number of qubits; these are the experiments that can be done efficiently (in polynomial time) and thus do not necessarily probe the system's complete state. For example, the outcomes of the set of measurements $\sigma_{i_1}\otimes\sigma_{i_2}\otimes\cdots\otimes\sigma_{i_n}$ where σ_{i_j} is one of the Pauli matrices or $\mathbf{1}$, completely determines the state, but there are 4^n measurements in this set. In an experimental setup, we might randomly select a polynomial subset of them and there is some small chance of order $\mathrm{poly}(n)/4^n$ that these are the measurements that distinguish the equilibrated state from the present state in the laboratiry that is supposed to approximate it. The estimates of time-dependent correlations could possibly be more sensitive to distance from equilibrium, as these involve time-evolved, nonlocal operations. The numerical study suggests that product baths whose size is polynomially related to the system can function as adequate baths in the sense of providing relaxation in polynomial time. The relaxed state could still be a rather rough approximation to the true equilibrium state, but as we argued above, it might be a good starting point for subsequent measurements.

We have taken the bath to be part of the (cost of) the quantum computer. In any experimental setup, there is a natural bath that is used to equilibrate and cool the quantum computer. Can we use this bath for a computational problem such as equilibration? Consider for example the NMR quantum computer [18] where computation takes place at room temperature. In the regime in which the heat bath has a non-Markovian character it has been shown to be possible to alter the Hamiltonian of the system and the coupling to the bath dynamically (see [43], but also standard books on NMR [33]). These techniques could make it possible to simulate the time evolution of a ''designer'' Hamiltonian and also to equilibrate the system to the equilibrium state of this designer Hamiltonian.

Finally, we have taken steps in developing a theory of quantum Markov chains for quantum computational purposes. It will be interesting to bring this theory to the next level. For example, we can try to define the notions of irreducibility and ergodicity for these quantum Markov chains. One of the essential question is then, can we find quantum Markov chains that are rapidly mixing for interesting computational problems for which no good classical algorithms are available? Such a problem could for example be a problem of equilibration for a specific Hamiltonian H. In this study we have laid the ground work for this future research.

ACKNOWLEDGMENTS

We would like to thank Charles Bennett, Daniel Loss, John Smolin, Ashish Thapliyal, Reinhard Werner, and Ronald de Wolf for stimulating discussions. D.P.D. thanks the U.S. Army Research Office for support under Contract No. DAAG55-98-C-0041.

APPENDIX A: NORMS

In this appendix we give the definitions of several norms and inner products. The inner product between vectors in $\mathbf{C}^{N^2}$ can be represented on $B(\mathcal{H}_N)$ as

$$\langle\chi_1|\chi_2\rangle=\mathrm{Tr}\,\chi_1^\dagger\chi_2. \tag{A1}$$

The trace norm [34,35] is defined as

$$\|A\|_{tr}=\mathrm{Tr}\sqrt{A^\dagger A}. \tag{A2}$$

What makes this norm attractive is that it captures a measurable closeness of two density matrices ρ_1 and ρ_2 [35]:

$$\|\rho_1-\rho_2\|_{\mathrm{tr}}=\max_A\sum_j\,|P_1^A(j)-P_2^A(j)|, \tag{A3}$$

where P_1^A and P_2^A are the probability distributions over outcomes j that are obtained by measuring observable A on ρ_1 and ρ_2. The matrix norm $|||\cdot|||_2$ is defined as

$$|||A|||_2=\max_{x:\|x\|_2=1}\|Ax\|_2. \tag{A4}$$

where $\|\cdot\|_2$ is the Euclidean norm on $\mathbf{C}^{N^2}$: $\sqrt{\langle v|v\rangle}$ for $|v\rangle\in\mathbf{C}^{N^2}$. We have

$$\|Ax\|_2\leqslant|||A|||_2\|x\|_2. \tag{A5}$$

In order to aid in the interpretation of the numerical results of Sec. II G, we present some numerical estimates for the average $\|\cdot\|_{tr}$ distance of two randomly chosen density matrices. We first have to choose a measure over $B_{\mathrm{pos},1}$. All density matrices can be written as $\rho=\Sigma_i\lambda_i|i\rangle\langle i|$ with $\Sigma_{i=1}^N\lambda_i=1$. The eigenvalues λ_k lie on a $(N-1)$-dimensional simplex S in $\mathbf{R}^N$. We use the Euclidean metric $\|\cdot\|_2$ induced on the simplex. The Haar measure on the group of unitary matrices $U(N)$ induces a uniform measure on the set of projectors $\{|i\rangle\langle i|\}_{i=1}^{N^2}$. Together this defines a measure $\mathcal{M}_{B_{\mathrm{pos},1}}$ [44]. Within this measure, one can express the average distance

TABLE II. The average distance between two randomly selected density matrices.

dim	Mean	Standard error $=\sqrt{\mathrm{var}/(n-1)},\ n=1000$
4	0.90388	0.00740588
8	0.96190	0.00514057
16	1.00294	0.00341226
32	1.01452	0.00220363
64	1.02617	0.00132233

between two density matrices ρ_1 and ρ_2, using the unitary invariance of $\|\cdot\|_{tr}$, as

$$[\|\rho_1-\rho_2\|_{tr}]_{\mathcal{M}_{B_{\mathrm{pos},1}}}$$

$$=\frac{1}{\mathrm{Vol}(S)^2 V(U(N))}\int dU\int_0^1 d\lambda_1\cdots d\lambda_k \times\int_0^1 d\mu_1\cdots d\mu_k\,\delta\left(\sum_i \lambda_i-1\right)\delta\left(\sum_i \mu_i-1\right) \times\mathrm{Tr}\left|\sum_j \lambda_j|j\rangle\langle j|-U\sum_j \mu_j|j\rangle\langle j|U^\dagger\right|. \tag{A6}$$

The values obtained by a numerical calculation of Eq. (A6) are tabulated in Table II.

APPENDIX B: PREPARATION OF THE BATH

To prepare the state

$$\rho_{b,\beta}=\rho_{b,\beta}^1\otimes\cdots\otimes\rho_{b,\beta}^k, \tag{B1}$$

given $H_b=\Sigma_{i=1}^k\mathbf{1}_{K/2}\otimes h_i$, we first calculate the eigenvalues and eigenvectors of each qubit Hamiltonian h_i. We prepare the state

$$\prod_{i=1}^k (e^{-\beta e_{i,0}}|0\rangle\langle 0|+e^{-\beta e_{i,1}}|1\rangle\langle 1|)/Z_i. \tag{B2}$$

with $\{e_{0,i},e_{1,i}\}$ the eigenvalues of qubit Hamiltonian h_i. This can be done by changing an initial state $|0\rangle\langle 0|$ with probability $e^{-\beta e_{i,1}}/Z_i$ into state $|1\rangle\langle 1|$ for each i. We then rotate each qubit to its eigenbasis $\{|b_{i0}\rangle,|b_{i1}\rangle\}$:

$$\otimes_{i=1}^k U_{b_i}=\otimes_{i=1}^k(|b_{i0}\rangle\langle 0|+|b_{i1}\rangle\langle 1|). \tag{B3}$$

In total we perform $2k$ elementary qubit operations plus some constant classical overhead.

---

[1] P. W. Shor, SIAM J. Computing **26**, 1484 (1997).
[2] A.Yu. Kitaev, e-print quant-ph/9511026; Russ. Math. Surveys **52**, 1191 (1999).
[3] L. K. Grover, in *Proceedings of the 28th Annual ACM Symposium on Theory of Computing* (ACM, New York, 1996), p. 212.
[4] D. Deutsch, Proc. R. Soc. London, Ser. A **400**, 97 (1985).
[5] S. Lloyd, Science **273**, 1073 (1996).
[6] C. Zalka, Proc. R. Soc. London, Ser. A **454**, 313 (1998); S. Wiesner, e-print quant-ph/9603028; B.M. Boghosian and W. Taylor, Physica D **120**, 30 (1998).
[7] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods* (Wiley, London, 1964).
[8] *Quantum Monte Carlo Methods in Equilibrium and Nonequilibrium Systems*, edited by M. Suzuki, Springer Series in Solid-State Sciences (Springer-Verlag, Berlin, 1986).
[9] H. De Raedt and W. von der Linden, *The Monte Carlo Method in Condensed Matter Physics*, Topics in Applied Physics Vol. 71 (Springer-Verlag, Berlin, 1991).
[10] M. Suzuki, Prog. Theor. Phys. **56**, 1454 (1976).
[11] S. Doniach and E. H. Sondheimer, *Green's Functions for Solid State Physicists* (Benjamin-Cummings, Reading, MA, 1974).
[12] M. Jerrum and A. Sinclair, SIAM J. Comput. **22**, 1087 (1993).
[13] F. Barahona, J. Phys. A: Math. Gen. **15**, 3241 (1982).
[14] In [15] quantum algorithms were presented to sample from the equilibrium distribution of arbitrary Ising spin glass models. The performance of these algorithms does not necessarily provide an exponential speedup over the best classical algorithms (see [12]) for the hard instances though.
[15] D. Lidar and O. Biham, Phys. Rev. E **56**, 3661 (1997).
[16] D.S. Abrams and S. Lloyd, Phys. Rev. Lett. **79**, 2586 (1997).
[17] D. P. DiVincenzo, in *Mesoscopic Electron Transport*, Vol. 345 of *NATO Advanced Study Institute, Series E: Applied Sciences*, edited by L. Sohn, L. Kouwenhoven, and G. Schoen (Kluwer, Dordrecht, 1997), p. 657; e-print cond-mat/9612126.
[18] N. Gershenfeld and I.L. Chuang, Science **275**, 350 (1997); D. Cory, A. Fahmy, and T. Havel, Proc. Natl. Acad. Sci. USA **94**, 1634 (1997).
[19] D. Abrams and S. Lloyd, e-print quant-ph/9807070.
[20] L. van Hove, Physica **23**, 441 (1957).
[21] E. Fick, G. Sauermann, and W. D. Brewer, *Quantum Statistics of Dynamic Processes*, Springer Series in Solid-State Sciences Vol. 86 (Springer-Verlag, Berlin, 1990); M. Celio and D. Loss, Physica A **158**, 769-783 (1989).
[22] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*, Lecture Notes in Physics (Springer-Verlag, Berlin, 1987).
[23] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[24] E. B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976).
[25] E.B. Davies, Commun. Math. Phys. **39**, 91 (1974).
[26] E.B. Davies, Math. Ann. **219**, 147 (1976).
[27] R. Alicki, Rep. Math. Phys. **10**, 249 (1976).
[28] N. Dunford and J. T. Schwartz, *Linear Operators* (Inter-

science, New York, 1958), pt. I.
[29] R. F. Werner (private communication).
[30] A. Messiah, *Quantum Mechanics* (Wiley, London, 1976), Vol. II, Chap. XVI.
[31] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, England, 1985).
[32] Cf. M. Jerrum and A. Sinclair, in *Approximation Algorithms for NP-hard Problems*, edited by D. S. Hochbaum (PWS Publishing, Boston, 1996), pp. 482–520.
[33] A. Abragam, *The Principles of Nuclear Magnetism* (Clarendon Press, Oxford, 1961); R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon Press, Oxford, 1987).
[34] S. Richter and R.F. Werner, J. Stat. Phys. **82**, 963 (1996).
[35] D. Aharonov, A. Kitaev, and N. Nissan, in *Proceedings of the 30th Annual ACM Symposium on Theory of Computation* (ACM, New York, 1997), pp. 20–30.
[36] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).
[37] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. London, Ser. A **454**, 339 (1998).
[38] R. Jozsa, Proc. R. Soc. London, Ser. A **454**, 323 (1998).
[39] D.A. Lidar and H. Wang, Phys. Rev. E **59**, 2429 (1999).
[40] Cf. G. W. Stewart and Ji-guang Sun, *Matrix Perturbation Theory*, *Computer Science and Scientific Computing* (Academic, San Diego, 1990); P.J. Schweitzer, J. Appl. Probab. **5**, 401 (1968); M. Haviv and L. van der Heyden, Adv. Appl. Probab. **16**, 804 (1984).
[41] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993).
[42] W. Feller, *An Introduction to Probability Theory and Its Application* (Wiley, London, 1968), Vol. I.
[43] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).
[44] K. Zyczkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).