Optimizing the rodeo projection algorithm

Thomas D. Cohen and Hyunwoo Oh
*Department of Physics and Maryland Center for Fundamental Physics, University of Maryland,
College Park, Maryland 20742, USA*

The rodeo algorithm has been proposed recently as an efficient method in quantum computing for projection of a given initial state onto a state of fixed energy for systems with discrete spectra. In the initial formulation of the rodeo algorithm, evolution times were chosen randomly via a Gaussian distribution with a fixed standard deviation. In this paper, it is shown that such a random approach for choosing times suffers from exponentially large fluctuations in the suppression of unwanted components: as the number of iterations gets large, the distribution of suppression factors obtained from random selection approaches a log-normal distribution, leading to remarkably large fluctuations. We note that by choosing times intentionally rather than randomly, such fluctuations can be avoided and strict upper bounds on the suppression can be obtained. Moreover, the average suppression using fixed computational cost can be reduced by many orders of magnitude relative to the random algorithm. A key to doing this is to choose times that vary over exponentially many timescales, starting from a modest maximum scale and going down to timescales exponentially smaller.

I. INTRODUCTION

Quantum computers, suggested by Richard Feynman in the 1980s [1], are expected to solve problems that are not accessible to classical computers. It is known that for a number of problems the asymptotic scaling of quantum algorithms are better than that of the best-known algorithms with classical computers. A notable example is the factoring of large numbers, where Shor's algorithm [2] has a computational cost that is a polynomial in the logarithm of the size of the system while the best-known classical algorithms have a cost that grows faster than any power law [3–6]. It is widely believed that many quantum mechanical problems, which suffer from exponentially bad sign problems [7] when computed classically via Monte Carlo methods, would only require polynomial resources on a quantum computer. Examples include real-time dynamics [8,9], where the time evolution induces complex phases that induce sign problems, and certain problems in which a chemical potential for fermions, such as QCD at a finite chemical potential [10], induces an complex effective action after integrating out the fermions. This latter example is of great importance in nuclear physics since QCD behaves this way. Thus, quantum computing is potentially useful for simulating a wide range of quantum systems with exponentially lower cost, spanning fields from condensed matter physics to nuclear physics and particle physics [11–16].

To simulate quantum physics, one needs to prepare initial states and evolve them in time according to the underlying Hamiltonian. It has long been recognized that the state preparation problem is a key to viable quantum computation, and research over nearly three decades has focused on developing methods for state preparation with a focus on using minimal resources [17–28].

The rodeo algorithm [29] has been suggested as a cost-effective method of preparing quantum states by projecting a given initial state onto an energy eigenstate for systems with discrete spectra; the success rate is proportional to the square of the overlap between the initial state and the target state of interest, and the time needed is logarithmic in the accuracy making it exponentially faster than phase estimation algorithm. It has been performed on a superconducting quantum computer [30] and variations of the algorithm have been considered [31]. An alternative approach, also based on a projection method via filtering processes has been suggested [32], which is designed to exploit particular knowledge about the character of the state.

The rodeo algorithm can be used either as a direct method of state preparation or in conjunction with other techniques for state preparation such as adiabatic quantum computing [33–39]—which could act to extend its utility in state preparation to situations in which the overlap between the initial and final states is small.

The rodeo algorithm is based on quantum interferometry and needs to be done for a relatively large number of iterations to ensure high accuracy; each iteration is associated with a running time and accurate projections require numerous distinct times. It is called the rodeo algorithm since during each iteration there is a probability that states other than the desired state are thrown out of the state in analogy to a hapless bull rider in a rodeo.

In Ref. [29] the emphasis was on the suppression factor per iteration. This might create an impression that the optimal approach is to create the maximum suppression per iteration and then perform a sufficient number of iterations to achieve the desired suppression. However, this need not be the case: large suppression factors per iteration require large times. It can be more efficient in terms of total time to use more iterations with shorter times per iteration (and less suppression per iteration).

The goal of this paper is to study how to optimize the algorithm to obtain the most accurate version of the algorithm

with minimal computational costs. As the dominant computational cost is running a Hamiltonian over time, we will use the net time expended as a proxy for computational costs. The algorithm involves multiple iterations of a basic computational structure, which is run for a given amount of time; optimization involves deciding what times should be used for each iteration and how many iterations should be run.

In this context, we note that if the computation is to be done on a digital quantum computer based on gates, then "time" could represent a mathematical time in a Trotterized version [40,41] of the time evolution. However, the rodeo algorithm would also be viable on analog quantum simulators—provided that one could use quantum gates to turn the simulator on and off in a controlled way. It is not implausible that such a scheme of gate-controlled analog quantum-simulation (GCAQS) might be the first setting in which rodeo projection proves to be decisive.

While the rodeo algorithm can be used to project onto any discrete state of the spectrum for any Hermitian operator, for simplicity of discussion, this paper will focus on the projection onto the ground state of a Hamiltonian of physical interest—which is likely to be a common application. Moreover, to further simplify discussion, we will assume that a constant has been added to the Hamiltonian of the system to make the ground-state energy zero. However, the discussion given here goes through *mutatis mutandis* for the more general case of an arbitrary discrete state with the any value of the ground-state energy.

In this paper we demonstrate that the version of the rodeo algorithm shown in Ref. [29], which is based on choosing the evolution times for the iterations randomly using a Gaussian distribution with a fixed standard deviation, has large fluctuations in efficiency that grow exponentially with the number of iterations. We show that these fluctuations can be suppressed by a very efficient algorithm designed with a set of "super-iterations" which include iterations of multiple timescales.

We begin in Sec. II with an introduction to the rodeo projection algorithm; additional details are given in Appendix A. Section III describes how the rodeo algorithm whose times for each iteration are chosen randomly via a Gaussian distribution can be optimized to minimize computational costs while maximizing suppression of unwanted states. This will serve as a baseline for comparison with other approaches. It is shown in Sec. IV that this random version of the rodeo algorithm has large fluctuations, which grow exponentially in the number of iteration. This implies that if one wishes to ensure a fixed level of accuracy, with a high level of confidence one would need to use substantially more resources than predicted by the average value. In order to suppress these fluctuations, in Sec. V the notion of a "super-iteration" algorithm is introduced. In Sec. VI a simple *ad hoc* prescription of choosing super-iterations is considered; additional details are given in Appendix B. Its results are compared with the random approach and shown to yield substantially stronger suppression of unwanted components than the random approach. Finally, a few practical issues associated with the implementation of the rodeo algorithm are considered in Sec. VII.

II. RODEO PROJECTION

The rodeo projection algorithm acts by suppressing components of the initial state depending on the energy of the state. Since the algorithm is only applicable to discrete states, there is necessarily an excited state with a minimum excitation, which we denote Δ. It is extremely helpful to know the value of Δ when attempting to optimize the performance of the algorithm: clearly it makes no sense to spend computational resources in order to improve the suppression of would-be components with energies less than Δ as such components do not exist.

The existence of a minimum excitation energy Δ, defines a natural timescale for our problem:

$$T_0 = \frac{2\pi\hbar}{\Delta}, \tag{1}$$

which is the time needed for a full period for the phase evolution of the lowest excited state.

Using time as our proxy for computational cost is slightly problematic in that it is a dimensionful quantity. One can rescale all times in the problem by a constant factor and rescale all energies by its inverse and so the computational difficulty is the same. Moreover, a large extent in the analysis of the rodeo algorithm involves products of energies and times. Accordingly we will often express quantities in terms of the following dimensionless combinations:

$$\zeta = \frac{ET}{2\pi\hbar} = \frac{E}{\Delta}\frac{T}{T_0}, \quad \zeta_{\text{tot}} = \frac{ET_{\text{tot}}}{2\pi\hbar} = \frac{E}{\Delta}\frac{T_{\text{tot}}}{T_0}, \tag{2}$$

where T is a time that occurs during one iteration and T_{tot} is the total time. The factor of 2π is included so that ζ measures the number of periods.

The rodeo algorithm involves dynamics in which the physical system of interest is entangled with an ancilla qubit. The scheme is based on iterations, with the concept of "successful" and "unsuccessful" iterations determined by a measurement of the ancilla qubit, indicating whether an iteration is successful or not. The input to the jth iteration is the physical state $|\psi\rangle_{\text{phys}}^{j\,\text{initial}}$ and the output of successful iterations is the physical state $|\psi\rangle_{\text{phys}}^{j\,\text{final}}$; the time used to implement the iteration is T_j.

Consider an initial physical state prior to the first iteration given by

$$|\psi\rangle_{\text{phys}}^{\text{initial}} = \alpha_g|\psi_g\rangle + \sum_c \alpha_c|\psi_c\rangle \quad \text{with}$$

$$P_g^i = |\alpha_g|^2 \quad \text{and} \quad |\alpha_g|^2 + \sum_c |\alpha_c|^2 = 1, \tag{3}$$

where the subscript g indicates the ground-state component and c indicates the excited states (where $\hat{H}_{\text{phys}}|\psi_a\rangle = E_a|\psi_a\rangle$); P_g^i is the initial probability that the system is in the ground state. The central features of the algorithm are the following:

(1) Following n consecutive successful iterations, the cth excited state is suppressed compared to the ground state by a factor of

$$s_c = \prod_{j=1}^n s_c^j \quad \text{with}$$

$$s_c^j = \cos^2\left(\frac{\omega_c T_j}{2}\right) = \cos^2\left(\pi\zeta_c^j\right), \quad \text{where}$$

$$\omega_c \equiv \frac{E_c}{\hbar} \quad \text{and} \quad \zeta_c^j \equiv \frac{\omega_c T^j}{2\pi}. \tag{4}$$

One can represent the overall suppression factor after n iterations for an arbitrary component in terms of ζ_{tot} [given in Eq. (2)] as

$$s(\zeta_{\text{tot}}) = s\left(\frac{E_c}{\Delta}\frac{T_{\text{tot}}}{T_0}\right) = \prod_{j=1}^{n}\cos^2\left(\frac{\omega_c T_j}{2}\right),$$

$$\text{where} \quad T_{\text{tot}} = \sum_{k=1}^{n} T_n. \tag{5}$$

(2) A consequence is that P_g^n, the probability that the physical system is in the ground state after n consecutive successful iterations, is

$$P_g^n = \frac{|\alpha_g|^2}{|\alpha_g|^2 + \sum_c\left[\prod_{j=1}^{n}\cos^2\left(\pi\zeta_c^j\right)\right]|\alpha_c|^2}$$

$$= \frac{P_g^i}{P_g^i + \int_\Delta^\infty dE\ s\left(\frac{E}{\Delta}\frac{T_{\text{tot}}}{T_0}\right)\rho(E)}$$

$$= \frac{P_g^i}{P_g^i + \left(1 - P_g^i\right)S_E}, \tag{6a}$$

where $s\left(\frac{E}{\Delta}\frac{T_{\text{tot}}}{T_0}\right)$ was given above. It is useful to consider the overall suppression for all excited states. This is naturally given in terms of the spectral density of the initial state, $\rho(E)$,

$$\rho(E) \equiv {}_{\text{phys}}^{\text{initial}}\langle\psi|\delta(E - \hat{H})|\psi\rangle_{\text{phys}}^{\text{initial}}. \tag{6b}$$

S_E, the overall suppression factor for excited states, is then given by

$$S_E \equiv \frac{\int_{\Delta-\epsilon}^\infty dE\ s\left(\frac{E}{\Delta}\frac{T_{\text{tot}}}{T_0}\right)\rho(E)}{\int_{\Delta-\epsilon}^\infty dE\ \rho(E)} \quad \text{with}\ \ 0 < \epsilon < \Delta. \tag{6c}$$

[The integral starts at $\Delta - \epsilon$ to ensure that the delta function for the first (discrete) excited state is included.] Clearly, S_E goes to zero as $n \to \infty$, which implies that as $n \to \infty$, $P_g^n \to 1$.

(3) Assuming the ground-state energy is known with very high accuracy and the algorithm is implemented without errors, the probability that n consecutive successful iterations occur starting from $|\psi\rangle_{\text{phys}}^{j\,\text{initial}}$, $P^{n\,\text{s}}$, is given by

$$P^{n\,\text{s}} = |\alpha_g|^2 + \sum_c\left(\prod_{j=1}^{n}\cos^2\left(\pi\zeta_c^j\right)\right)|\alpha_c|^2, \tag{7}$$

where $|\psi\rangle_{\text{phys}}$ is the physical state after n consecutive successful iterations. This implies that

$$\lim_{n\to\infty} P^{n\,\text{s}} \to |\alpha_g|^2 = \left|\langle\psi_g|\psi\rangle_{\text{phys}}^{\text{initial}}\right|^2. \tag{8}$$

(4) If an iteration is unsuccessful prior to obtaining a physical state with sufficiently high probability to be in the ground state (for the purposes of the problem one is studying), one needs to recreate the initial state and start from the beginning. The expected number of times one needs to create (or recreate) $|\psi\rangle_{\text{phys}}^{\text{initial}}$ is $\frac{1}{P^{n\,\text{s}}}$ where n is the number of successful iteration needed to achieve the desired accuracy.

(5) The expected net time to obtain a successful projected state with the accuracy associated with n iterations

$$T^{\text{expected}} = \frac{T^{n\,\text{s}}}{\left|\langle\psi_g|\psi\rangle_{\text{phys}}^{\text{initial}}\right|^2}, \tag{9}$$

where $T^{n\,\text{s}}$ is the time needed for n successful iterations; in Eq. (9) it is assumed that $\sum_c(\prod_{j=1}^{n}\cos^2(\pi\zeta_c^j))|\alpha_c|^2$ is negligibly small.

To see how these features come about, it is useful to focus on the details of how a single iteration works; this is given in Appendix A. Prior to the beginning of the jth iteration, the the physical system is in pure state given in the energy eigenbasis by

$$|\psi\rangle_{\text{phys}}^{j\,\text{initial}} = \alpha_g^j|\psi_g\rangle + \sum_c\alpha_c^j|\psi_c\rangle$$

$$\text{with}\quad \left|\alpha_g^j\right|^2 + \sum_c\left|\alpha_c^j\right|^2 = 1. \tag{10}$$

Thus, $P_g^{j\,i}$, the initial probability for the system to be in the ground state at the onset of iteration j, is given by $P_g^{j\,i} = |\alpha_g^j|^2$.

As shown in Appendix A the effect of a successful iteration on the state is to affect the following transformation:

$$|\psi\rangle_{\text{phys}}^{j\,\text{initial}} \xrightarrow[\text{iteration}]{\text{successful}} |\psi\rangle_{\text{phys}}^{j\,\text{final}} \quad \text{with}$$

$$|\psi\rangle_{\text{phys}}^{j\,\text{final}} = \frac{\alpha_g|\psi_g\rangle + \sum_c\frac{1}{2}(1 + e^{-i2\pi\zeta_c^j})\alpha_c^j|\psi_c\rangle}{\sqrt{\left|\alpha_g^j\right|^2 + \sum_c\left|\alpha_c^j\right|^2\cos^2\left(\pi\zeta_c^j\right)}}$$

$$\text{where}\quad \zeta_c^j \equiv \frac{E_c T^j}{2\pi\hbar}. \tag{11a}$$

Thus

$$|\psi\rangle_{\text{phys}}^{j\,\text{final}} = \alpha_g^{\prime j}|\psi_g\rangle + \sum_c\alpha_c^{\prime j}|\psi_c\rangle$$

$$\text{with}\quad \alpha_g^{\prime j} = \frac{\alpha_g^j}{\sqrt{\left|\alpha_g^j\right|^2 + \sum_c\left|\alpha_c^j\right|^2\cos^2\left(\pi\zeta_c^j\right)}},$$

$$\alpha_c^{\prime j} = \frac{\alpha_c^j\cos\left(\pi\zeta_c^j\right)e^{-i\pi\zeta_c^j}}{\sqrt{\left|\alpha_g^j\right|^2 + \sum_c\left|\alpha_c^j\right|^2\cos^2\left(\pi\zeta_c^j\right)}}. \tag{11b}$$

A successful iteration of the algorithm increases the probability that the system is in the ground state:

$$P_g^{j\,1\,\text{s}} = \left|\alpha_g^{\prime j}\right|^2 = \frac{P_g^{j\,i}}{\left|\alpha_g^j\right|^2 + \sum_c\left|\alpha_c^j\right|^2\cos^2\left(\pi\zeta_c^j\right)} \geqslant P_g^{j\,i}, \tag{12}$$

where the fact that $|\alpha_g^j|^2 + \sum_c|\alpha_c^j|^2\cos^2(\pi\zeta_c^j)$ is necessarily less than or equal to unity gives rise to the inequality. Moreover this inequality is only saturated if every component c either has $\alpha_c^j = 0$ or has $\cos^2(\pi\zeta_c^j) = 1$, an exceedingly unlikely possibility. Thus a successful iteration will enhance the ground-state component of the physical system.

From Eq. (11b), it is clear that the enhancement of the ground-state probability comes through the relative suppression of components other than the ground state, which is given

by

$$s_c^j \equiv \left(\frac{|\alpha_c'^j|}{|\alpha_g'^j|} \Big/ \frac{|\alpha_c^j|}{|\alpha_g^j|} \right)^2 = \cos^2 \left(\pi \zeta_c^j \right). \qquad (13)$$

Since this suppression happens for each iteration, the net suppression for component c is the product $\prod_{j=1}^n s_c^j$, which establishes feature (i) above; feature (ii) follows logically from it.

The reduced density matrix for the physical system is given in Eq. (A3a) of Appendix A. Since the reduced density matrix sums over the auxiliary qubit, it sums over both successful and unsuccessful iterations. From the form of Eq. (A3a), it is apparent that the total probability that the physical system is in the ground state—including both successful and unsuccessful iterations—is given by $|a_g^j|^2$, which as noted above is the initial probability of the system being in the ground state at the start of the jth iteration. Thus, if one includes the contributions of both successful and unsuccessful iterations, the total probability that the system is in the ground state is unchanged by the iteration. Since this holds for all iterations, it follows that the net probability that the system is in the ground state is unchanged from the initial configuration prior to any iterations. Feature (iii) follows immediately; features (iv) and (v) are natural consequences of (iii).

Once the number of iterations and the times for each iteration are fixed, the total suppression for components with any given energy is determined. However, the algorithm as described so far has not been fully specified, since criteria for choosing the times for each iteration and the number of iterations have not been given.

It is useful to recall that the suppression factor for any iteration only depend on $\zeta_c^j = E_c T^j / (2\pi \hbar)$. Since our interest is to suppress contributions with various energies contained in the initial state, the goal is to pick times that usefully suppress these. Clearly, it is suboptimal to use the same time for each iteration of algorithm: if this were done energies for which $w_c T^j$ is an odd integer (where $\omega = E/\hbar$) would have a suppression of unity (i.e., no suppression) no matter how many iterations were done and energies near these points would only be weakly suppressed. The remainder of the paper is dedicated to considering various schemes for picking the times.

III. OPTIMIZING THE RANDOM RODEO ALGORITHM

The original formulation of the algorithm chose times randomly from a Gaussian distribution with a predetermined standard deviation for each iteration [29]. We will refer to this incarnation of the rodeo algorithm as the random rodeo algorithm (RRA). The logic of RRA simple: there is no correlation between the times apart from that fixed by the overall standard deviation. Given this one might reasonably expect that all excited states will have large suppression after many iterations since one does not have coherent regions that are largely unsuppressed.

This section will discuss how to optimize RRA to maximize suppression for minimal computational cost. As will be discussed in subsequent sections, there exist other ways to pick times which generically produce more suppression at less computational cost than the optimized RRA. But to see that, it is important to know just how well RRA can do.

The optimization problem is not yet well posed. Different components of the initial state are suppressed differently and depending on the information one has about the initial state one might sensibly alter what one is attempting to optimize. For example, if one knew that the excited state components in the initial state was dominated almost entirely by components with energies between Δ (the minimum excitation energy) and 2Δ, one would attempt to configure the algorithm to suppress those configurations optimally; the strategy could be qualitatively different if the initial state contained substantial contributions with excited energies covering many orders of magnitude.

Let us consider the case where we know nothing about the initial state beyond the fact that it is normalized and all excited components have energy greater than or equal to Δ. In that case, it seems reasonable to adopt a conservative strategy to optimize the suppression of the components with energies equal to or greater than Δ which on average are the least suppressed (i.e., have the largest expectation value of the suppression factor s) assuming fixed average total time. Since this variant of the algorithm picks times randomly all one can specify is the average total time; similarly one cannot ensure that there will not be components that are less suppressed than average, but at least one can be sure that in a statistical sense the suppression should be at least as large as this for any component. Equivalently, one can ask how much time one needs to spend on average to ensure that the expected average suppression for all components with energy greater than Δ are suppressed on average by at least fixed predetermined factor.

This is straightforward to do. From Eq. (4), the suppression for a component with energy $E = \hbar\omega$ after n iterations of the algorithm is $s(\omega) = \prod_{j=1}^n \cos^2(\frac{\omega T_j}{2})$ and since the T_j are chosen randomly and independently of each other, the ensemble average of the suppression is given by

$$s^{\mathrm{Avg}} = \int \left(\prod_{j=1}^n dT_j \; p(T_j, T) \right) \cos^2 \left(\frac{\omega T_1}{2} \right) \cos^2 \left(\frac{\omega T_2}{2} \right) \cdots \cos^2 \left(\frac{\omega T_n}{2} \right) = \left[\int dT' p(T', T) \cos^2 \left(\frac{\omega T'}{2} \right) \right]^n$$

$$= \left(\frac{1}{2} \right)^n \left[1 + \exp \left(-\frac{\pi \omega^2 T^2}{4} \right) \right]^n = \left(\frac{1}{2} \right)^n [1 + \exp(-\pi^3 \zeta^2)]^n, \qquad (14)$$

where n is the number of iterations, $p(T_j, T)$ is given from the positive half of a normal distribution whose average value is T (so that $T = \sqrt{\frac{2}{\pi}} T_{\mathrm{RMS}}$) and the substitution $\zeta \equiv \frac{\omega T}{2\pi}$ is made in the final form of s.
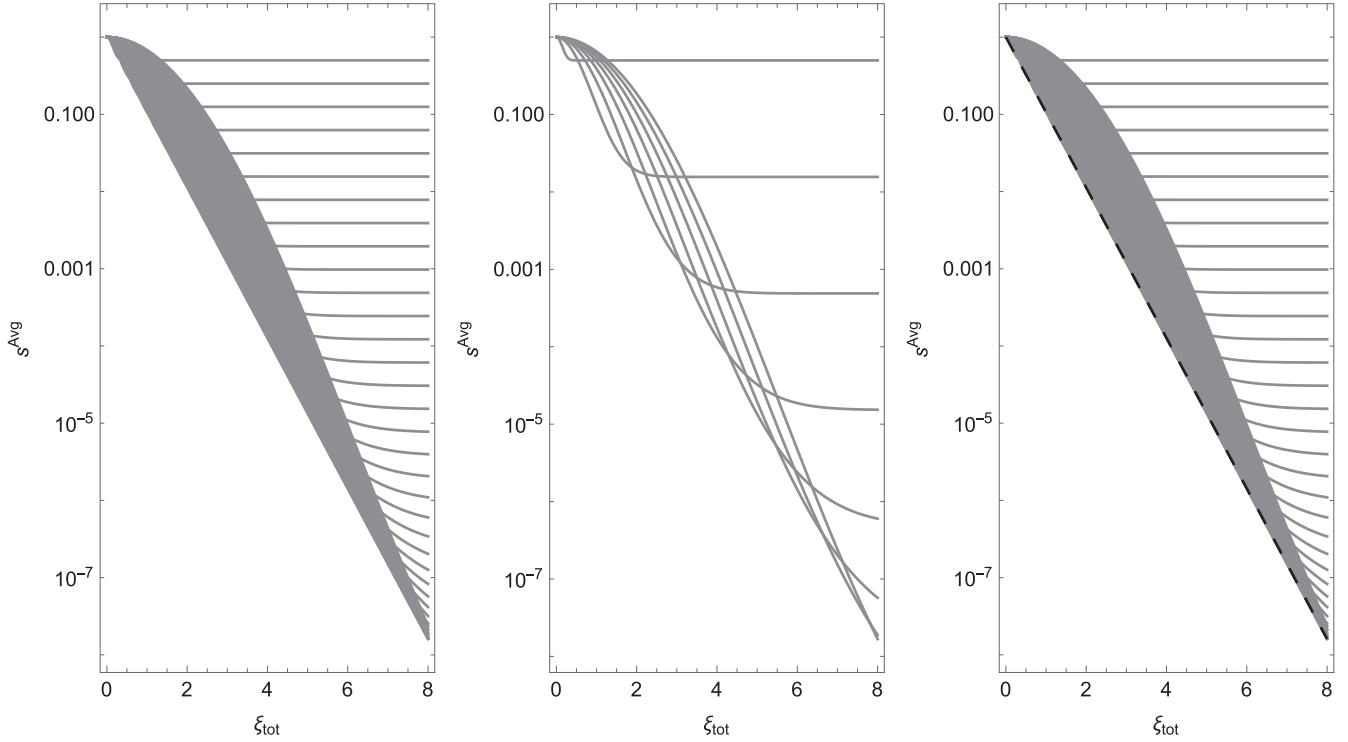
FIG. 1. Average suppression factor as a function of ζ_{tot} for various values of n. The left panel shows all values of n from 1 to 40. As the curves become dense near the separatrix, obscuring the features, the middle panel shows every fifth value of n. The right panel includes a dashed line corresponding to the exponential fit given in Eq. (16d).

It should be clear that as $\zeta \to 0$, $s^{\text{Avg}} \to 1$. This simply reflects the fact the algorithm is designed so that the ground state remains unsuppressed. On the other hand, for large ζ, s^{avg} asymptotes[1] to $1/2^n$. Moreover, it approaches its asymptotic value quite quickly since $\exp(-\pi^3 \zeta^2)$ becomes very small for quite modest values of ζ.

Note that the total average time, T_{tot}, is simply the number of iterations times the average time per iteration given by $T_{\text{tot}} = nT$. It is useful to rewrite Eq. (14) in terms of T_{tot} or equivalently $\zeta_{\text{tot}} = \omega T_{\text{tot}}/2\pi$:

$$s^{\text{Avg}} = \left(\frac{1}{2}\right)^n \left[1 + \exp\left(\frac{-\pi^3 \zeta_{\text{tot}}^2}{n^2}\right)\right]^n. \qquad (15)$$

In Fig. 1, s^{Avg} is plotted as a function of ζ_{tot} for various values of n. The leftmost panel includes all n's from 1 to 40. It is clear that the $\zeta_{\text{tot}} - s^{\text{Avg}}$ plane separates into two regions. Situations in which there are no values of n for which a given value of ζ_{tot} has a corresponding s^{Avg} are in the lower left. In this region it is impossible to use RRA to obtain an average level of suppression, s^{Avg}, for the given ζ_{tot}. Conversely, the top right region corresponds to situations in which it is possible to use RRA to obtain an average level of suppression, s^{Avg}, for the given ζ_{tot}. There is a clear separatrix between these two regions.

---

[1]Note this asymptotic behavior differs from the $1/4^n$ discussed in Ref. [29] that $1/4^n$ behavior corresponds to the geometric mean while the present result is for the arithmetic mean. The qualitative difference between these behaviors reflects very large fluctuations and will be discussed in the next section.

The way the separatrix forms is a bit hard to discern from the left most panel of Fig. 1 since the curves for the various n's are so dense. To clarify how the nearly exponential separatrix forms, the middle panel plots only every fifth n. For values of ζ_{tot} greater then 1, the separatrix is approximated extremely well by a simple exponential—a line in the log plot.

The most natural fit to a simple exponential is the one that exactly matches the exact separatrix once for each value of n. This is straightforward to do: if one takes the functional form of Eq. (15) but takes n to be a continuous real variable rather than having discrete integer values, one can minimize S^{Avg} with respect to n for fixed ζ_{tot}. This yields

$$n = \alpha \zeta_{\text{tot}}, \qquad (16a)$$

where α satisfies

$$\frac{2\pi^2}{\alpha^2 \left[1 + \exp\left(\frac{\pi^3}{\alpha^2}\right)\right]} = \ln\left[\frac{2}{1 + \exp\left(-\frac{\pi^3}{\alpha^2}\right)}\right], \qquad (16b)$$

so that

$$\alpha \approx 4.271 \qquad (16c)$$

and

$$S^{\text{Avg}} = \exp\left(-\beta \zeta_{\text{tot}}\right) \quad \text{with}$$

$$\beta = -\alpha \ln\left[\frac{1 + \exp\left(-\frac{\pi^3}{\alpha^2}\right)}{2}\right] \approx 2.244. \qquad (16d)$$

Equations (19)–(22) give the minimum value of S^{Avg} for any value of n, integer or discrete. However, RRA is restricted to integer values. Thus, Eq. (22) gives a lower bound for S^{Avg}:

$$S^{\text{Avg}} \geqslant \exp\left(-\beta\,\zeta_{\text{tot}}\right) \text{ with } \beta \approx 2.244. \tag{16e}$$

However, when ζ_{tot} is such that n [as given in Eq. (19)–(21)] happens to be an integer, then the simple exponential form precisely matches the separatrix and Eq. (23) becomes an equality. Moreover, since the curves for the various n's near the separatrix become quite dense by the time ζ_{tot} reaches one, one expects the simple exponential form to be an excellent approximation. This is seen in the right panel of Fig. 1.

With this result in hand let us return to the optimization problem of determining how much time one needs to spend to ensure that all components with energy greater than Δ (the gap or minimum excitation energy) are suppressed on average by at least fixed predetermined factor S. Recalling that $\zeta_{\text{tot}} = \frac{E T_{\text{tot}}}{2\pi\hbar}$, Eq. (16e) becomes $T_{\text{tot}} \geqslant -\frac{2\pi\hbar\ln(S)}{\beta E}$. To ensure that the time is long enough so that all components of the state are suppressed by at least S, one needs to use Δ for the energy. Thus,

$$T_{\text{tot}} \geqslant -\frac{2\pi\hbar\ln(S)}{\beta\Delta} \tag{17}$$

If one uses $T_0 \equiv 2\pi\hbar/\Delta$ as a natural unit of time then

$$\frac{T_{\text{tot}}}{T_0} \geqslant -\frac{\ln(S)}{\beta} \approx 0.4456\ln(S). \tag{18}$$

IV. FLUCTUATIONS IN THE RANDOM RODEO ALGORITHM

Ensuring that the ensemble-averaged suppression factor is smaller than some fixed value S for all excited state energies might seem to be a sensible way to obtain a controlled level of suppression. This would certainly be true if the statistical fluctuations in the suppression factor were themselves under control. For example, if the suppression factors were distributed normally so that the ratio of the distributions width to the average suppression scaling like $n^{-\frac{1}{2}}$, then at large n one could be quite confident that the suppression factor in various runs did not exceed S much. This would allow a reliable and predictable projection onto the ground state.

Unfortunately, this is not the case. Rather, there are exceptionally large fluctuations, which make it very difficult (and computationally expensive) to produce a given level of suppression in a reliable way.

A. Fluctuations in statistical measures over the random ensemble

A hint that exceptionally large fluctuations are present can be seen in the fact that there are multiple qualitatively different "natural" expectations for the amount of suppression after n iterations.

Since the average of $\cos^2$ is $\frac{1}{2}$, it seems reasonable that well away from $\zeta = 0$ (where there is no suppression) on average each iteration should randomly sample from $\cos^2$ leading to expectation of an average suppression of $\frac{1}{2}$ per iteration. On the other hand, it also seems reasonable that well away from $\zeta = 0$, where a random assignment of times will lead to an approximately uniform distribution of phases in the cosine and as noted in Ref. [29], the geometric mean of $\cos^2(\theta)$ leads to an expectation of a suppression of $\frac{1}{4}$ per iteration.

Which one of these "natural" expectations is correct? An apparently puzzling feature is that there is strong evidence that each of these is at least approximately correct: Ref. [29] gives evidence for a suppression of $\frac{1}{4}$ per iteration: it plots the logarithm of the suppression for four values of ζ for one random run of up to 150 and shows that for $\zeta = 3.0$ the logarithm suppression closely follows $-N\ln 4$ (where N is the number of iterations (with notable fluctuations) while for $\zeta = 2.0$ it nearly does so. This strongly supports the notion that the suppression per iteration is at least close to $\frac{1}{4}$ for sufficiently large ζ.

However, the evidence that the (ensemble) average suppression per iteration is $\frac{1}{2}$ for large enough ζ is also compelling. From Eq. (4), the suppression for a component with energy $E = \hbar\omega$ for n iterations of the algorithm is $s(\omega) = \prod_{j=1}^{n}\cos^2(\frac{\omega T_j}{2})$ and since the T_j are chosen randomly and independently of each other, the ensemble average of the suppression is given by Eq. (15). Clearly as ζ gets large, s, the ensemble average suppression, approaches 2^{-n}, not 4^{-n}.

These two asymptotic scaling behaviors may seem incompatible, but they need not be. It is perfectly possible that "typical" suppression factors scale at least roughly as 4^{-n} for sufficiently large ζ, while the average is dominated by atypical suppression factors that occur relatively infrequently but are orders of magnitude larger than the typical ones.

It is easy to verify that this is happening. In Table I the geometric mean and root-mean-square (r.m.s.) values for the suppression for the ensemble used in RRA are given along with the arithmetic mean computed earlier. The expressions are derived analogously to the arithmetic mean.

From the expressions for r.m.s. and arithmetic mean one can easily construct the ratio of the standard deviation σ to the (arithmetic) mean s^{Avg}:

$$\frac{\sigma}{s^{\text{Avg}}} = \left[\left(\frac{3}{2}\right)^n \left(\frac{\left(1 + \frac{\exp(-4\pi^3\zeta^2)}{3} + \frac{4\exp(-\pi^3\zeta^2)}{3}\right)}{(1 + \exp(-\pi^3\zeta^2))^2}\right)^n - 1\right]^{\frac{1}{2}} \xrightarrow{\text{large } n}$$

$$\times \left(\frac{3}{2}\right)^{\frac{n}{2}} \left(\frac{\left(1 + \frac{\exp(-4\pi^3\zeta^2)}{3} + \frac{4\exp(-\pi^3\zeta^2)}{3}\right)}{(1 + \exp(-\pi^3\zeta^2))^2}\right)^{\frac{n}{2}} \xrightarrow{\text{large } \zeta} \left(\frac{3}{2}\right)^{\frac{n}{2}}. \tag{19}$$

TABLE I. Three statistical quantities characterizing the suppression factor as a function pf ζ_{tot} given by the ensemble used in the random rodeo algorithm. The second column gives an expression for the quantity as a function of ζ_{tot} for fixed n. The third column gives the asymptotic value of these forms at large ζ_{tot}. In all cases the asymptotic form is reached to very good approximation whenever $\zeta_{tot} > 1$. The fourth column gives the best exponential fit analogous to the separatrix in the ζ_{tot}-s plane which were calculated analogously to the method described in the discussion surrounding Eqs. (16a)–(16e). As there, the exponential fit to the separatrix determines the minimal time needed to ensure that the statistically weighted suppression factor achieves a fixed level or better for all energies.

Statistical quantity	$s(\zeta_{tot})$ for fixed total average time and fixed n	Asymptotic form for large ζ_{tot}	Best exponential fit for separatrix in the ζ_{tot}-s plane
Geometric mean	$\exp\{n\int_{-\infty}^{\infty}dT\,\ln[\cos^2(\pi T\zeta_{tot})]p(T,1)\}$	4^{-n}	$\exp(-\beta\zeta_{tot})$ with $\beta\approx 4.46$
Arithmetic mean	$(\frac{1}{2})^n[1+\exp(-\pi^3\zeta_{tot}^2)]^n$	2^{-n}	$\exp(-\beta\zeta_{tot})$ with $\beta\approx 2.244$
Root-mean-square	$(\frac{3}{8})^{n/2}(1+\frac{\exp(-4\pi^3\zeta_{tot}^2)}{3}+\frac{4\exp(-\pi^3\zeta_{tot}^2)}{3})^{n/2}$	$(\frac{3}{8})^{n/2}$	$\exp(-\beta\zeta_{tot})$ with $\beta\approx 1.637$

The behavior for large n follows because the quantity

$$\left(\frac{3}{2}\right)\left(\frac{\left(1+\frac{\exp(-4\pi^3\zeta^2)}{3}+\frac{4\exp(-\pi^3\zeta^2)}{3}\right)}{[1+\exp(-\pi^3\zeta^2)]^2}\right)$$

is larger than unity for nonzero values of ζ_{tot} and becomes much large than one for large n. Empirically, both the r.m.s. and arithmetic mean approach their asymptotic quite rapidly at even modest values ζ_{tot}.

Clearly, at large n, the fluctuations in the distribution of s values become exponentially larger as a function of n.

The three statistical quantities which are all sample from the same distribution are different. If the suppression factors were dominated by one characteristic scale, all of these quantities would be at the same scale. But as seen explicitly in the third column, three differ and the differences between them are exponentially large as a function of n.

To see how the large scale of the fluctuations affects the ability of the scheme to reliably accomplish a projection, the exponential fits in the fourth column of the Table I are shown in Fig. 2. The implication of this is quite striking. If, for

example, one considers the case of $\zeta_{tot} = 5$, RRA can be used to ensure that the geometric mean of suppression is below 2.07×10^{-10}, but the arithmetic mean of the suppression factor is at best 1.34×10^{-5}, nearly 5 orders of magnitude worse than the geometric mean; the best that can be achieved for the r.m.s. is 2.79×10^{-4} which is almost 6 orders of magnitude worse than the achievable value from the geometric mean.

It was suggested earlier that "typical" values of the suppression factors scale at least roughly as 4^{-n} for sufficiently large ζ, explaining the results seen in Ref. [29]. It is easy to understand why this occurs. The suppression factor is the product of many variables each of which is randomly chosen from the same distribution. In such a situation one expects a log-normal distribution to emerge when the number of samples, n, becomes large: the central limit theorem can be applied to the logarithms of the variables which are combined additively.[2] The geometric mean of a log-normal distribution is known to be its median. This implies that at large n one expects that the median of the suppression factor—a reasonable proxy for a "typical" value—should be the geometric mean; in this case the geometric mean is easily shown to be 4^{-n}.

B. Fluctuations in a single run

So far, we have focused on the fluctuations of statistical quantities for fixed values of ζ_{tot} evaluated over the statistical ensemble of RRA. However, intuitively, one might also expect that the large fluctuations seen in the ensemble average would be reflected as large fluctuations as a function of ζ for any given run. This subsection provides numerical evidence that this occurs by looking at a representative example. Indeed, one sees that typical suppression factors are of order 4^{-n}, the average suppression is of order 2^{-n}, and the average suppression
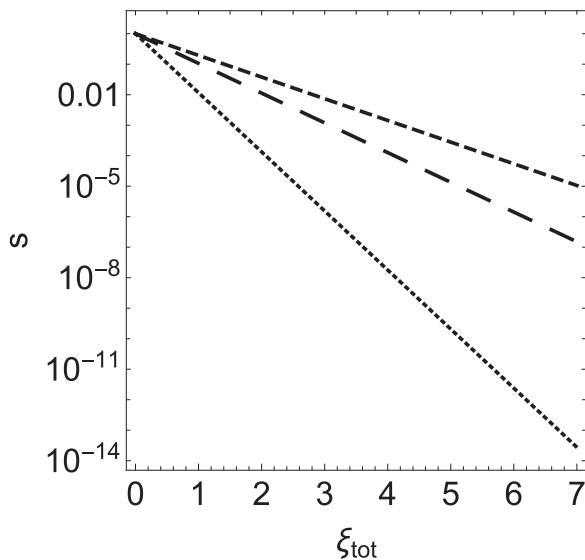

FIG. 2. Exponential fits to average (arithmetic mean) suppression factor (long dashed), root-mean-square suppression factor (short-dashed), and geometric mean (dotted) as a function of ζ_{tot}.

[2]Analogously with the central limit theorem for normal distributions, when the underlying distribution has a limited range, the range of the underlying distribution can put constraints on the ability of the tail of the log-normal distribution to correctly describe the true distribution for n samples. In the present context, the suppression factor is bounded from above by unity. However, the log-normal distribution has support for $S > 1$. Thus, quantities which evaluated with the log-normal distribution receive significant contributions from regions of $S > 1$ will not be correctly described by the log-normal distribution. It turns out that the r.m.s. is such a quantity.
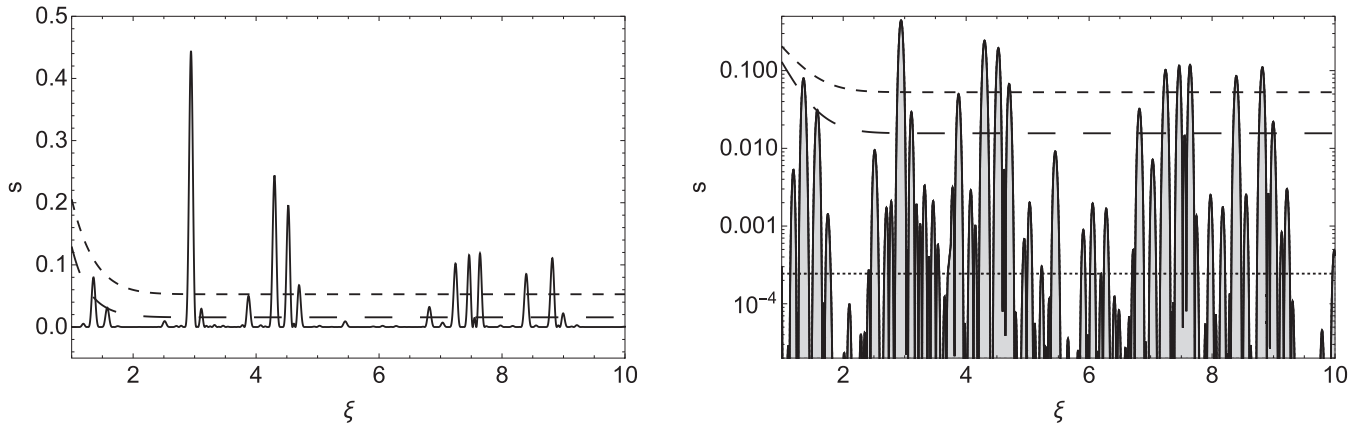
FIG. 3. The solid line represents the suppression factor for a random run with six iterations with time chosen from the positive side of a normal distribution and an average time of unity. The long-dashed curve is the ensemble average for six iterations with that distribution for the suppression, the short-dashed curve is the r.m.s., and the dotted line in the log plot is the geometric mean.

is dominated by narrow peaks in ζ with suppression factors that are much less suppressed than the average.

In Fig. 3 a representative sample is given for the suppression factor obtained after six iterations with times picked from the half-normal distribution with an average of unity and plotted over a range of $1 < \zeta < 10$. (Six-iteration was chosen so that the ensemble average $1/64$ was still large enough to be easily discerned from zero by eye.) The existence of very large fluctuations compared to the average is readily apparent. The highest peak is 13 times larger than the average, while typical values between the peaks are so exceptionally small that they can only be discerned in a log plot.

The minimum value showed in the log plot gives an indication of just how small it is. It is clear that the fraction of value of ζ in the regime from 2 to 10 (where the average function has essentially asymptoted to 2^{-6}) that has $s < 4^{-6}$ is close to half: a numerical integral of $\Theta(4^{-6} - s[\zeta])$ where Θ is the Heaviside step function yields an estimate that approximately 58% of the points in this range have $s(\zeta) < 4^{-6}$ for this set of iterations. This is consistent with the notion that a "typical" point has $S(\zeta)$ of order 4^{-n}.

Similarly, the average value of s in the region from 2 to 10 has an average value of $s(\zeta)$ of order 2^{-n}. Numerically it is approximately 0.014 while $2^{-6} \approx 0.0156$, so for this particular run it is about 90% of what one would expect the average to be, assuming that individuals reflect the ensemble average at large n and that $n = 6$ is sufficiently large. Given the paucity of large peaks in this range of ζ and the fact that only six iterations were used, this is remarkably close.

The existence of these large fluctuations represents a serious challenge for using RRA in an efficient way. In order to ensure that the suppression factor has a high probability of reaching a needed level of suppression for all relevant components of the initial state, one must use computational resources that are several times larger than the times needed for a typical component to be suppressed to that level. If computational resources are fixed (as is likely when the first quantum computer becomes available for which the rodeo algorithm is viable for a particular problem) then the fluctuations imply that one does not even have a reliable estimate of the order of magnitude of

the suppression, making one quite unsure of just how well an initial state has been projected onto the ground state.

V. SUPER-ITERATIONS

There are two reasons why nonrandom variants of rodeo projection might be preferable to RRA. The first is related to the previous section: the random algorithm produces exponentially large fluctuations (in n) which greatly restrict the efficiency of the approach if one wishes to reliably know the level of suppression. Ideally one could intentionally choose parameters that control these fluctuations. The second is simply that for any particular well-defined quantity associated with suppression there exists an optimal choice (in terms of n and the various values of ζ—or equivalently T—for each iteration) that maximizes (or minimizes as the case may be) the quantity. At best, a random algorithm can approximate this optimal choice. Thus, as a matter of principle an appropriate intentional choice of the parameters should always perform better than the random algorithm—regardless of the quantity one is attempting to optimize.

Note that large fluctuations arise when multiple iterations have suppression factors near unity. Suppression factors of unity occur when $\pi \zeta_c^j$ is an integer while suppression factors of zero occur when $\pi \zeta_c^j$ is an half-integer. One can restrict the scale of such large fluctuations by requiring that suppression factors of unity are always multiplied by suppression factors of zero for all nonzero values of ζ. A key insight is that this can be arranged to occur with a finite total time (or ζ) (and indeed only twice the time of the underlying distribution).

This can be done using "super-iterations," rather than the single iterations considered heretofore. A super-iteration consists of an infinite set[3] of iterations each of which takes half the time (or ζ_c) of the previous one. Since the times form a geometric series the net time for a super-iteration is exactly twice of the underlying iteration.

---

[3]Of course, in practice this will need to be large but finite; the effects of keeping it finite will be discussed in Sec. VII A.
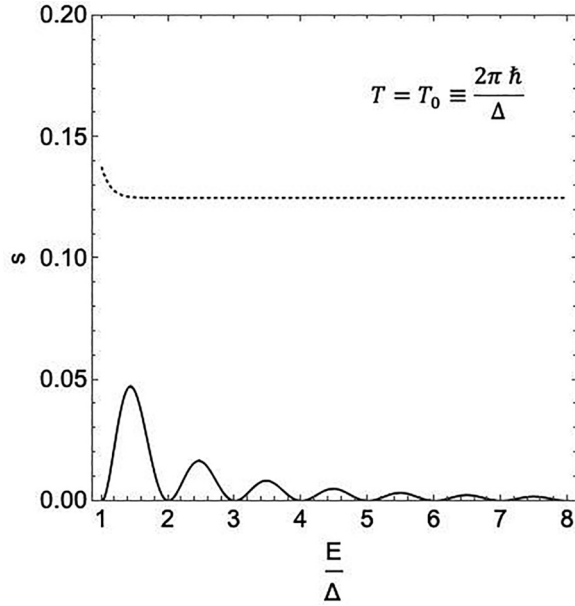
FIG. 4. The solid line represents the suppression factor for a single "super-iteration" as a function of the excitation energy in units of Δ, the minimum excitation energy. The time was fixed to be $T_0 = 2\pi\hbar/\Delta$. The dotted line uses the same total (average) time and gives the ensemble average suppression factor for three iterations—the number of iteration that ensures that the maximum value of the average suppression factor is as small as possible.

Note from Eq. (4) that the second stage of the super-iteration with a time given by $T^j/2$ has a suppression factor of zero whenever $\cos^2(\pi\zeta_c^j/2)$ is zero, which occurs when ζ_c^j is twice an half-integer (odd integer). Thus it coincides with places where the first stage has a suppression factor of unity and exactly cancels them. It does not cancel all of the place where the initial iteration was unity. However, the next stage—where the a time is given by $T^j/4$—will cancel when ζ_c^j is four times an half-integer, the stage after that when ζ_c^j is eight times an half-integer and so forth. Thus, as a whole the super-iteration will cancel every place where original iteration was unity at a cost of spending twice the time.[4]

The suppression factor for the jth super-iteration at given $\zeta_{c\,\mathrm{sup}}^j$ (which is twice ζ_c^j, the base iteration, since a super-iteration takes twice as long) is given by

$$s_{c\,\mathrm{sup}}^j = \prod_{k=0}^{\infty} \cos^2\left(\frac{\pi\zeta_c^j}{2^k}\right) = \prod_{k=1}^{\infty} \cos^2\left(\frac{\pi\zeta_{c\,\mathrm{sup}}^j}{2^k}\right)$$

$$= j_0^2\left(\pi\zeta_{c\,\mathrm{sup}}^j\right) = \left(\frac{\sin(\pi\,\zeta_{c\,\mathrm{sup}})}{\pi\,\zeta_{c\,\mathrm{sup}}}\right)^2, \qquad (20)$$

where $j_0(x)$ is the 0th spherical Bessel function.

From Eq. (2), $\zeta = (E/\Delta)(T/T_0)$. Thus if one allocates a time of $T_0 = 2\pi\hbar/\Delta$—one period of phase evolution for the lowest excited state—and does just a single super-iteration,

_____

[4]It is analogous to the method suggested in [32] to remove states with discrete quantum numbers.

the suppression factor has a function of energy will be given by $j_0(\pi E/\Delta)^2$ and is shown in Fig. 4.

Figure 4 plots s for $E/\Delta \geqslant 1$, since Δ is the minimum excitation so that all excited state components in the initial state have $E \geqslant \Delta$. The choice of $T = T_0$ ensures that for the lowest excited state $s = 0$, the components with a minimal excitation energy are completely suppressed. There are two other important salient features. The obvious one is evident from the scale of the suppression factor. Note that the maximum suppression factor is less than 0.05 for all energies above the lowest excited state. The other is that s falls off rapidly with E/Δ or equivalently ζ.

The overall small scale of s for all relevant energy is quite striking given that the net time is very modest—only a single period of the lightest excitation. A useful comparison is with RRA: if one restricted the (average) total time of a single period of the lightest excitation, and optimized the average for $E > \Delta$, then one should use three iterations (see Fig. 1). The average suppression factor is plotted as a dotted line. Note that for all values of $E/\Delta > 1$ the super-iteration is well below the average value of the random algorithm; it is always at a factor of 2.6 smaller. Moreover, for larger $E/\Delta > 1$ the single super-iteration rapidly becomes orders of magnitude better than three iterations of the random algorithm. The fact that super-iterations drop off so rapidly at large ζ is a clear benefit in attempting to construct optimized versions of the rodeo algorithms—if a scheme using multiple super-iterations concentrates on improving the suppression at values of ζ between one and two, it will be essentially guaranteed to have excellent suppression everywhere for $\zeta > 1$. This is not true of the random algorithm.

VI. OPTIMIZING RODEO PROJECTION

The demonstration that a single super-iteration achieves substantially more suppression at all energies than the random algorithm with the same fixed time—and does so without suffering from large uncontrolled fluctuations—makes clear the advantages of intentionally choosing the parameters of the rodeo algorithm rather than relying on randomness to achieve optimal suppression.

As noted in Sec. III, in order to optimize the rodeo projection algorithm one needs to know the quantity that one is optimizing. Even if one poses a problem as one fixes the total computational resources to be used (or more precisely its proxy, the total time in units of $T_0 = \frac{2\pi\hbar}{\Delta}$) and asks to optimize the suppression factor, one has not fully specified the optimization problem since the suppression factor is a function of excited energies.

If one knew a significant amount about the spectral density of the initial state, then one might choose to optimize the algorithm in a manner which suppresses components that have high probabilities more heavily. Merely knowing information about the spectrum of the theory, without any addition information about the initial state could influence the choice of what we wish to optimize. For example, if one knew that the spectrum contained no states with energies between Δ and 2Δ, then it would not make sense to expend computational resources suppressing components with energies in that region since one knows that they do not appear in the initial state.

The most straightforward approach to optimization has two steps: the first is to choose some quantity $Q(s, T_{tot})$ that is a functional of the suppression factor $s(\zeta_{tot})$, and the total time, and has the property that when minimized yields strong suppression given the state of interest. The second step is to seek the s produced by the rodeo algorithm subject to constraints on the total time that minimizes Q as best one can.

Note that whatever strategy one adopts in choosing Q, there exists an optimal choice of parameters that minimizes Q once T_{tot} is specified. That said, it may not be practical to find the optimal choice or even come close. One can explore a range of possibilities and ultimately use the one which provides the smallest value of Q.

We start by considering the situation where nothing is known about the initial state beyond the fact that all unwanted components have energies greater than or equal to Δ, and one wishes to adopt a conservative strategy in which one has a rigorous bound on the total suppression of excited states. There is a clear strategy: one can choose Q to be the largest value as a function of E in the physically relevant regime $E \geqslant \Delta$:

$$Q = \max_{E \geqslant \Delta} \; s\left(\frac{E}{\Delta} \frac{T_{tot}}{T_0}\right), \tag{21}$$

which, in effect, implies choosing the worst possible value of the suppression in the possible physical range. By choosing Q to be the worst case, one ensures that the suppression factors for all components of the state are not larger than Q, which in turn puts an upper bound for S_E, and the total amount by which the excited state components are suppressed:

$$S_E \equiv \frac{\int_{\Delta-\epsilon}^{\infty} dE \; s\left(\frac{E}{\Delta} \frac{T_{tot}}{T_0}\right) \rho(E)}{\int_{\Delta-\epsilon}^{\infty} dE \; \rho(E)} \leqslant Q. \tag{22}$$

Moreover, it does so without using any information about the initial state or the spectrum (beyond the fact that lowest excitation is Δ). Also, Q depends solely on $s(\zeta_{tot})$ and the total time, but not on the initial state. Thus one merely needs to seek to optimize Q once for each T_{tot} and not undergo the daunting task of repeating the optimization for each initial state.

A potential downside of strategy [Eq. (21)] is its very conservative nature. As noted it will provide an upper bound for S_E, but it may be a very large upper bound compared to the least upper bound. Note that if we construct s from super-iterations with energies which are a few times Δ, the suppression factor becomes extremely small near chosen energies even if only a few super-iterations are used. If the initial state is heavily weighted towards such energies, the actual overall suppression of excited components, S_E, will be much smaller than the bound produced given by Q, and if the bound on S_E is fixed at the level of projection needed for the problem at hand, substantially more computational resources would be expended than needed. After considering this conservative bound, we will consider ways to improve upon it.

Unfortunately, even for this most conservative approach, the mathematical problem of finding s such that it yields the minimum value of $\max_{E \geqslant \Delta} s\left(\frac{E}{\Delta} \frac{T_{tot}}{T_0}\right)$ subject to the condition that s is obtained through some number of rodeo iterations with fixed total time is an open one. However, one can adopt
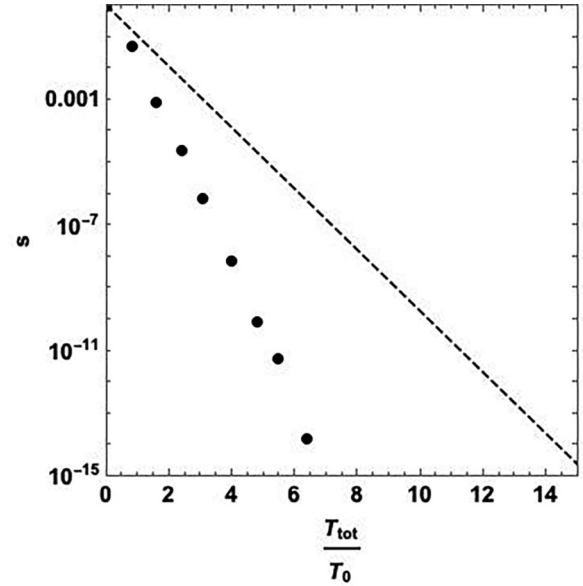

FIG. 5. The solid points are an absolute upper bound on the suppression factor using intentionally fixed times the scheme outlined in Appendix B for various number of super-iterations as a function of total time. For comparison, the dashed line represents a lower bound for the time for the ensemble average of S to be greater than a fixed amount using the random rodeo algorithm of Ref. [29].

ad hoc prescriptions and test how well they do. It is natural to consider prescriptions based on super-iterations as these seem very efficient.

We illustrate the existence of prescriptions for which the times are chosen intentionally rather than randomly and that yield exceptionally small suppression factors in quite modest times using a scheme outlined in Appendix B. The prescription entails generating suppression factors based on a fixed number of super-iterations of various times (all of order T_0), with each super-iteration generated from the previous one. The times for each super-iteration and the *maximum* suppression factors (for $E > \Delta$) are listed in Table II. Note that the maximum suppression factors are remarkably small: for $T_{tot}/T_0 \sim 3$. the suppression factor is $\sim 10^{-6}$ while for $T_{tot}/T_0 \sim 6.4$ the suppression factor is $\sim 1.5 \times 10^{-14}$.

It is instructive to compare the results using suppression factors obtained from the scheme in Appendix B with the random approach. In Fig. 5 the absolute *upper* bound of the suppression factor is compared with the *lower* bound for the time for the ensemble average of S to be greater than a fixed amount. It is clear that for fixed total times the intentional approach gives suppression factors that are orders of magnitude better than the random approach, when T is a couple of time T_0 or more. For example, when $T/T_0 = 6.4$, the intentional approach yields an upper bound on the total suppression that is a factor of 3.77×10^7 smaller than the lower bound ensemble average. If the calculation is restricted to short times (which is likely to be true when quantum computers first become large enough to implement the rodeo algorithm with sufficient coherence times), the existence of achieving order of magnitude better performance at fixed times is striking.

TABLE II. The times for each super-iteration and the maximum value of the suppression factors for $E > \Delta$ using the *ad hoc* scheme described in Appendix B.

Maximum suppression factor	Total time in units of T_0	Number of super-iterations	Times for each super-iteration in units of T_0							
$\max_{E \geqslant \Delta} \; s(\frac{E}{\Delta}\frac{T_{\text{tot}}}{T_0})$	T_{tot}/T_0	n	T_1/T_0	T_2/T_0	T_3/T_0	T_4/T_0	T_5/T_0	T_6/T_0	T_7/T_0	T_8/T_0
4.719×10^{-2}	0.8129	1	0.8129							
8.508×10^{-4}	1.5906	2	0.9361	0.6545						
2.421×10^{-5}	2.4222	3	0.9494	0.6638	0.8090					
7.549×10^{-7}	3.0752	4	0.9785	0.6841	0.8338	0.5788				
7.385×10^{-9}	3.9865	5	0.9764	0.6827	0.8320	0.5776	0.9180			
8.948×10^{-11}	4.7944	6	0.9881	0.6908	0.8419	0.5845	0.9290	0.7601		
5.689×10^{-12}	5.4500	7	0.9925	0.6939	0.8457	0.5871	0.9331	0.7634	0.6343	
1.539×10^{-14}	6.4010	8	0.9895	0.6918	0.8431	0.5853	0.9303	0.7611	0.6324	0.9675

A less dramatic way to view the improvement is to focus on the total time needed to achieve a certain level of suppression, in which case the gain looks to be something like a factor of two in total time. Whichever way one views this, however, the comparison understates the advantages of using an intentional approach. As noted earlier, the random approach yields exponentially large fluctuations making it complicate the achievement with reliable levels of suppression. In contrast, this problem is entirely vitiated by choosing the times intentionally: it provides an upper bound for the suppression factor that is truly reliable.

Moreover, the scheme discussed in Appendix B was constructed entirely *ad hoc* to provide a concrete example of a scheme that substantially outperforms the random approach. It is not implausible that a broad search over parameter space could yield substantially improvements. We leave such a search to future research.

As a practical matter, implementing rodeo projection efficiently would be helped greatly if one produces a table, similar to Table II, but containing a denser set of total times and ideally, better suppression factors. It would then be a simple matter to either pick out the suppression required for the problem at hand and find the time needed to ensure that this is achieved or alternatively fix a maximum time that can be allotted and find the level of suppression that could be achieved. In either case one could then use the times of various super-iterations (or regular iterations) associated with that row of the table to implement the projection efficiently.

A. Incorporating information about the initial state

If one had full information about the spectral function of the initial state, one could choose to optimize things differently. Instead of attempting to minimize $Q = \max_{E \geqslant \Delta} s(\frac{E}{\Delta}\frac{T}{T_0})$ and thereby put an upper bound on S_E, the net suppression factor for the excited states, one could attempt to find a set of (super) iterations with fixed total time that minimized S_E or a set with a fixed value of S_E that minimized the time. This is unlikely to be done since one would need to do this minimization separately for each initial state whose spectral function is known.

On the other hand, if one had access to a table of the sort described above (or even just Table II) and knowledge of the

spectral function, one could either pick a time that one wishes to spend and find the precise value of S_E associated with it (rather than merely having an upper bound that could be orders of magnitude larger than S_E) or perhaps more practically identify the level of suppression in S_E required for the problem at hand and choose the shortest time from the table that achieves it. This time will likely be significantly shorter than the time obtained from the upper bound on S_E.

However, circumstances in which one has access to the full information of spectral function will presumably be quite rare. It seems far more likely that partial information about the spectral function might be available. Even with partial information one may be able to find significantly stronger upper bounds than Eq. (21).

To do this it is useful to construct a function that is the least upper bound of $s(\frac{E}{\Delta}\frac{T}{T_0})$ subject to the constraint that it is monotonically decreasing. Let us denote this function $s^{\text{UB}}(\frac{E}{\Delta}\frac{T}{T_0})$. In Fig. 6 this is done for the $n = 3$ case given in Table II.

Next, suppose that there is a good reason to believe that the most of the spectral strength above the ground state is
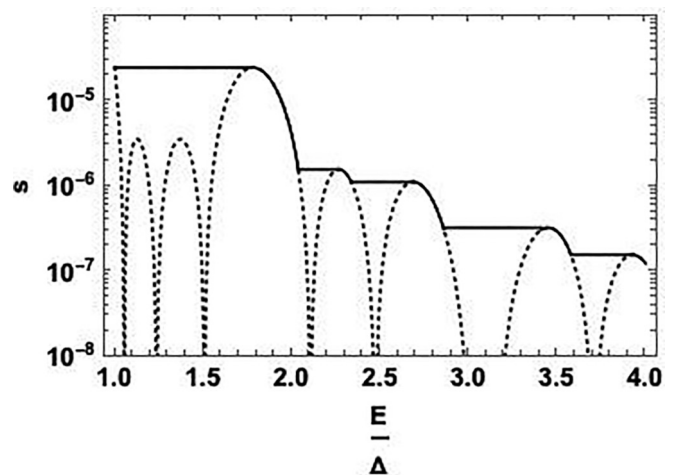

FIG. 6. The solid line is $s^{\text{UB}}(\frac{E}{\Delta}\frac{T}{T_0})$, the least upper bound of $s(\frac{E}{\Delta}\frac{T}{T_0})$ subject to the constraint that it is monotonically decreasing for the $n = 3$ case given in Table II; the dotted line is $s(\frac{E}{\Delta}\frac{T}{T_0})$ itself.

predominantly at energies at least a few times Δ. One can characterize this with an energy E_0 and f, the fraction of the excited state integrated spectral density above E_0:

$$f \equiv \frac{\int_{E_0}^{\infty} dE \ \rho(E)}{\int_{\Delta - \epsilon}^{\infty} dE \ \rho(E)}. \tag{23}$$

The bound on S_E then becomes

$$S_E \leqslant (1 - f) \, s^{\mathrm{UB}} \left(\frac{\Delta}{\Delta} \frac{T_{\mathrm{tot}}}{T_0} \right) + f \, s^{\mathrm{UB}} \left(\frac{E_0}{\Delta} \frac{T_{\mathrm{tot}}}{T_0} \right)$$

$$\leqslant Q. \tag{24}$$

As an example, if one uses three super-iterations with the simple scheme of Appendix B given in Table II, and has $f = 0.99$ and $E_0 = 3$, then the bound becomes $S_E \leqslant 5.591 \times 10^{-7}$, which is more than a factor of 40 smaller than the bound in Table II which assumes no knowledge of the initial state. If $f = 0.9999$ with $E_0 = 8$ the bound becomes $S_E \leqslant 1.194 \times 10^{-8}$, which is a factor of 2000 smaller.

VII. PRACTICAL CONSIDERATIONS

A number of practical issues involving the implementation of the rodeo projection algorithm will arise when suitable quantum computers become available. This section will consider a few of these and discuss ways to deal with them.

A. Cost of super-iterations

In the introduction it was suggested that the time of computation in units of T_0 is a reasonable proxy for the computational cost. While this is largely true, there is an important caveat, apart from allowing the system to evolve in time according to Hamiltonian dynamics, there is an additional overall computational cost associated with running each iteration: one needs to both couple the auxiliary qubit to the system of interest (which involves use of Hadamard gates and a conditional coupling) and to measure the auxiliary qubit. While such costs may be negligible for any given iteration, super-iterations as formally defined in Sec. V require an infinite number of iterations and thus, these fixed costs per iteration would become arbitrarily large.

The obvious solution is to simply use a fairly large but finite number of iterations for each super-iteration. If one uses N iterations in each super-iteration, then

$$s_{\mathrm{sup\,N}}(\zeta_{\mathrm{sup}}) = \prod_{k=1}^{N} \cos^2 \left(\frac{\pi \zeta_{\mathrm{sup}}^j}{2^k} \right) = \frac{j_0^2(\pi \zeta_{\mathrm{sup}})}{j_0^2 \left(\frac{\pi \zeta_{\mathrm{sup}}}{2^N} \right)}. \tag{25}$$

Since $j_0^2 \left(\frac{\pi \zeta_{\mathrm{sup}}}{2^N} \right)$ is very close to unity until ζ_{sup} is of order 2^{N-2}, such a limitation will not noticeably alter any results for suppression until ζ_{sup} gets very large.

One might worry that when ζ_{sup} is an integer multiple of 2^N the super-iteration becomes unity and suppression is lost. This is unlikely to be a practical problem, but it does mean that theoretical upper bound for any given N, the suppression factors given in a table analogous to Table II (or a more sophisticated table) only hold for energies less than some calculable maximum. Thus, if one knows the integrated spectral density above that maximum is negligible, the bound remains valid.

Fortunately such maxima occur at very high energies for relatively modest N. For example, when a single super-iteration is used, the maximum energy (in units of Δ) consistent with the suppression on the table will occur at approximately 2^N. With the values in Table II a more precise maximum energy consistent with the quoted bound is $E_{\mathrm{max}}/\Delta \approx (2^N - 1.430)/0.8129$ so that for $N = 15$, it holds up to energies of $40,308 \times \Delta$. If more than one super-iteration is used, the maximum suppression factor calculated for an arbitrarily large number of iterations per super-iterations will continue to hold to much higher energies than the case of a single iteration. In any case, one can always make N sufficiently large that this is not a practical problem.

B. False negatives

Up until now we have focused on reducing the probability of "false positives," the survival of components other than the ground state after the algorithm has run. We have ignored the problem of "false negatives" since, in the ideal case where the ground-state energy is known precisely and the time evolution and control over timing are perfect, every ground-state component will survive each iteration with unit probability. However, in practice the energy of the ground state is never known exactly and there will be errors due to imperfections in the propagation of the state in time. These can lead to "false negatives" in which the auxiliary qubit is measured in the down state even though the physical system is in the ground state. If one has a reasonable idea of the scale of the uncertainty of the ground state, the effects of imperfect calculation, and the probability that the state is in its ground state, the likelihood of a false negative for any given iteration can be estimated. Assuming that likelihood is small but not totally negligible, the natural thing to do is ignore the fact that the auxiliary qubit was measured down and continue the calculation while keeping track of a nominally unsuccessful run. If the auxiliary qubit is measured to be up through the next super-iterations one should simply assume that a false negative occurred and continue. If a second nominally unsuccessful iteration occurs anytime up through the next super-iteration one should assume that the iteration was truly unsuccessful. It should be very rare for this to care accidentally unless the uncertainty in the ground-state energy is very large.

C. Uncertainties in the energy of the first excited state

The formulation of the problem assumed that the energy of the ground state and the first excited state were known. As discussed above a lack of precise knowledge of the ground-state energy can lead to false negatives, but there is a useful strategy for dealing with these. A lack of precise knowledge of Δ, the energy of the first excited state relative to the ground state, is even easier to mitigate. There are two possibilities to consider: either are the estimate of Δ is larger than the true value, or it is smaller.

Underestimating Δ has a very modest cost. T_0, the time standard, will be larger than needed and hence all times (the proxy for computational cost) will be slightly larger than strictly necessary to achieve the level of suppression intended.

However, the upper bound for suppression will be computed in Table II or its generalization will remain valid since a given value of ζ_{tot} occurs at a lower energy.

Overestimating Δ is more serious: a given value of ζ_{tot} occurs at a lower energies and hence effectively probes the suppression factor at values of the energy below the true value Δ where the suppression factor can be and generically is larger than the computed upper bound.

The obvious way to make sure that Δ is not overestimated is to ensure that the value used is chosen conservatively to be below the estimated value by an amount larger than the uncertainty. This would guarantee a reliable upper bound and a very modest computational cost.

APPENDIX A: A RODEO ITERATION

This Appendix derives various properties of the density matrix and reduced density matrix associated with a single iteration of the rodeo algorithm.

The Hilbert space of the total system is enlarged by the inclusion of an ancilla qubit, which is initialized to be in the state $|\uparrow\rangle$ (in the z basis) and the physical system is in a pure quantum state. Given the inclusion of an additional degree of freedom—which will be entangled with the state of the phys-

ical state, a natural description of the system the system is via density matrices and reduced density matrices.

The physical system prior to the running of the jth iteration of the algorithm is $|\psi\rangle_{\text{phys}}^{j\,\text{initial}}$ given in Eq. (10); the full (pure state) density matrix in the extended Hilbert space is given by the tensor product

$$\hat{\rho}_{\text{total}}^{j\,\text{initial}} = |\uparrow\rangle\langle\uparrow| \otimes \hat{\rho}_{\text{phys}}^{j\,\text{initial}},$$

where the subscripts "phys" and "total" indicate the space describing the physical system being studied and the total system including the ancilla qubit; j simply labels the iteration.

The total system is then subject to unitary evolution:

$$\hat{\rho}_{\text{total}}^{j\,\text{final}} = \hat{U}_j\,\hat{\rho}_{\text{total}}^{j\,\text{initial}}\,\hat{U}_j^{\dagger} \quad \text{where}$$
$$\hat{U}_j = \exp[-i(\hat{p}_x^{\uparrow} \otimes \hat{H}_{\text{phys}} T_j)] \quad \text{with}$$
$$\hat{p}_x^{\uparrow} \equiv \tfrac{1}{2}(1 + \hat{\sigma}_x); \tag{A1}$$

$\hat{p}_x^{\uparrow}$ acts on the ancilla qubit and is a projector onto the up state in the x basis. This is achieved by acting on the ancilla qubit with a Hadamard gate, running Hamiltonian dynamics for the physical system for a time T_j, conditionally on the ancilla qubit being in the "up" state and then again acting on the ancilla qubit with a Hadamard gate.

The crux of the algorithm is the quantum interference between the components in which the Hamiltonian time evolves and the components where it does not.

Following the time evolution the density matrix for the full system is given by

$$
\begin{aligned}
\hat{\rho}_{\text{total}}^{j\,\text{final}} = &\;|\uparrow\rangle\langle\uparrow| \otimes \left[\left|\alpha_g^j\right|^2 |\psi_g\rangle\langle\psi_g| + \sum_{c,c'} \cos\left(\frac{\omega_c T_j}{2}\right)\cos\left(\frac{\omega_c' T_j}{2}\right)(\alpha_c\alpha_c^*|\psi_c\rangle\langle\psi_{c'}| + \alpha_c^*\alpha_{c'}|\psi_{c'}\rangle\langle\psi_c|) \right. \\
&\left. + \sum_c \cos\left(\frac{\omega_c T_j}{2}\right)(\alpha_g\alpha_c^* e^{-i\omega_c T_j/2}|\psi_g\rangle\langle\psi_c| + \alpha_g^*\alpha_c e^{i\omega_c T_j/2}|\psi_c\rangle\langle\psi_g|) \right] \\
&+ |\downarrow\rangle\langle\downarrow| \otimes \left[\sum_c \left|\alpha_c^j\right|^2 \sin^2\left(\frac{\omega_c T_j}{2}\right)|\psi_c\rangle\langle\psi_c| + \sum_{c\neq c'} \sin\left(\frac{\omega_c T_j}{2}\right)\sin\left(\frac{\omega_{c'} T_j}{2}\right)(\alpha_c\alpha_c^*|\psi_c\rangle\langle\psi_{c'}| + \alpha_c^*\alpha_{c'}|\psi_{c'}\rangle\langle\psi_c|) \right] \\
&+ |\downarrow\rangle\langle\uparrow| \otimes \left(\sum_c \frac{1}{2}\left(e^{-i\omega_c T_j} - 1\right)\alpha_c|\psi_c\rangle \right)\left(\alpha_g^*\langle\psi_g| + \sum_{c'} \frac{1}{2}\left(e^{i\omega_{c'} T_j} + 1\right)\alpha_{c'}^*\langle\psi_{c'}| \right) \\
&+ |\uparrow\rangle\langle\downarrow| \otimes \left(\alpha_g^*|\psi_g\rangle + \sum_c \frac{1}{2}\left(e^{-i\omega_c T_j} + 1\right)\alpha_c^*|\psi_{c'}\rangle \right)\left(\sum_c \frac{1}{2}\left(e^{i\omega_{c'} T_j} - 1\right)\alpha_{c'}^*\langle\psi_c| \right) \quad \text{with } \omega_a = \frac{E_a}{\hbar}.
\end{aligned} \tag{A2}
$$

The next step is to measure the state of the auxiliary qubit. This allows one to use Eq.(A2) to write the reduced density matrix for the physical system at the end of the iteration in the following form:

$$
\hat{\rho}_{\text{phys}}^{j\,\text{final}} = \left|\alpha_g^j\right|^2 |\psi_g\rangle\langle\psi_g| + \sum_c \left|\alpha_c^j\right|^2 |\psi_c\rangle\langle\psi_c| + \sum_c \cos\left(\frac{\omega_c T_j}{2}\right)(\alpha_g\alpha_c^* e^{-i\omega_c T_j/2}|\psi_g\rangle\langle\psi_c| + \alpha_g^*\alpha_c e^{i\omega_c T_j/2}|\psi_c\rangle\langle\psi_g|) \tag{A3a}
$$

$$
+ \sum_{c\neq c'} \left[\cos\left(\frac{\omega_c T_j}{2}\right)\cos\left(\frac{\omega_c' T_j}{2}\right) + \sin\left(\frac{\omega_c T_j}{2}\right)\sin\left(\frac{\omega_c' T_j}{2}\right) \right](\alpha_c\alpha_c^*|\psi_c\rangle\langle\psi_{c'}| + \alpha_c^*\alpha_{c'}|\psi_{c'}\rangle\langle\psi_c|)
$$

$$
= P^{j\,s}\hat{\rho}^{j\,s} + P^{j\,u}\hat{\rho}^{j\,u} \quad \text{with} \tag{A3b}
$$

$$
P^{j\,s} = \left|\alpha_g\right|^2 + \sum_c \left|\alpha_c^j\right|^2 \cos^2\left(\frac{\omega_c T_j}{2}\right), \quad P^{u\,s} = \sum_c \left|\alpha_c^j\right|^2 \sin^2\left(\frac{\omega_c T_j}{2}\right), \tag{A3c}
$$

$$\hat{\rho}^{js} = \frac{\left[\alpha_g|\psi_g\rangle + \sum_c \alpha_c^j \frac{1}{2}(1+e^{-i\omega_c T_j})|\psi_c\rangle\right]\left[\alpha_g^*\langle\psi_g| + \sum_c' \alpha_{c'}^{*j}\frac{1}{2}(1+e^{i\omega_{c'}T_j})|\langle\psi_{c'}|\right]}{\left|\alpha_g^j\right|^2 + \sum_c \left|\alpha_c^j\right|^2 \cos^2\left(\frac{\omega_c T_j}{2}\right)}, \tag{A3d}$$

$$\hat{\rho}^{ju} = \frac{\left[\sum_c \alpha_c^j \sin\left(\frac{\omega_c T_j}{2}\right)|\psi_c\rangle\right]\left[\sum_{c'} \alpha_{c'}^{*j}\sin\left(\frac{\omega_{c'}T_j}{2}\right)|\langle\psi_{c'}|\right]}{\sum_c |\alpha_c^j|^2 \sin^2\left(\frac{\omega_c T_j}{2}\right)}; \tag{A3e}$$

in Eq. (A3b), P^{js} is the probability that the auxiliary qubit is in the up state (in the z). and P^{ju} is the probability that it is in the down state; the superscripts s and u denote "successful" and "unsuccessful" iterations.

The nomenclature "successful" and "unsuccessful" are associated with the measurement of the auxiliary. If it is measured to be down (in the z basis), the entanglement with the auxiliary qubit ensures that the physical system is projected onto the state described by $\hat{\rho}^{ju}$ which has no components in the ground state. Since the goal of the algorithm is to enhance the components in the ground state, such a measurement is regarded as an unsuccessful iteration. If the measurement yields an up state—a successful iteration—then the physical system is projected onto the state described by $\hat{\rho}^{ju}$. This is the density matrix for a pure state: the effect of a successful iteration is to transform the state of the physical system

$$|\psi\rangle_{\text{phys}}^{j\,\text{initial}} \xrightarrow[\text{iteration}]{\text{successful}} |\psi\rangle_{\text{phys}}^{j\,\text{final}} \quad \text{with}$$

$$|\psi\rangle_{\text{phys}}^{j\,\text{initial}} = \alpha_g|\psi_g\rangle + \sum_c \alpha_c|\psi_c\rangle \quad \text{and}$$

$$|\psi\rangle_{\text{phys}}^{j\,\text{final}} = \frac{\alpha_g|\psi_g\rangle + \sum_c \frac{1}{2}(1+e^{-i\omega_c T_j})\alpha_c^j|\psi_c\rangle}{\sqrt{\left|\alpha_g^j\right|^2 + \sum_c \left|\alpha_c^j\right|^2 \cos^2\left(\frac{\omega_c T_j}{2}\right)}}. \tag{A4}$$

APPENDIX B: AN *AD HOC* PRESCRIPTION

Table II was obtained using a simple *ad hoc* prescription. The purpose was simply to demonstrate explicitly that there exist schemes that significantly outperform the average results of RRA and they do so without suffering from the exponentially large fluctuations which are inherent in the RRA scheme. It should be noted, however that there is no reason to assume, *a priori*, that this prescription achieves something close to the actual optimum choice.

We denote this, the Whac-a-Mole (WAM) prescription after the arcade game of that name, which bears a certain resemblance to the prescription. In that game, "moles" pop up

in random places and need to be suppressed by whacking them with a soft mallet, after which a mole emerges somewhere else, which also needs to be suppressed via whacking.

The prescription is based on super-iterations. As a first stage, one chooses a super-iteration with a time equal to T_0. This yields a suppression factor with zeros at all integer values of E/Δ. Our interest is in energies greater than Δ. One can identify the largest value of the suppression factor for energies greater than Δ and consider this to be the "mole" to be whacked by choosing the time for the next supper iteration to create a zero a precisely at the energy of the maximum. Such a time will be larger than T_0 so that smallest zero in the suppression function will remain at $E = \Delta$. To do a third iteration one finds the maximum suppression factor for $E > \Delta$ and treats it as the "mole" to be whacked by choosing the time for the next super-iteration to create a zero a precisely at the energy of the maximum. This can be continued for as many cycles as one wishes. Each cycle of this prescription yields progressively narrower peaks at progressively lower scales of the suppression factor and the maxima appear for E/Δ between one and two.

The scheme as described so far is constrained to preserve $E = \Delta$ as the smallest zero of the suppression factor as a function of energy. Such a constraint can be relaxed to further optimize the time by simply using an overall scaling factor of less than unity for the times. Such a scaling does not affect the maximum value of the suppression factors but merely pushes a given value to higher energies. In other words, it acts to push energies that were previously below Δ to above Δ. Since the suppression factor at $E = \Delta$ was initially zero, an overall scaling factor in times that is slightly smaller than unity will only push small suppression factors into the physical region, $E \geqslant \Delta$. To optimize the time, one chooses a scaling factor such that the new suppression factor at $E = \Delta$ is equal to the maximum, since any smaller one will increase the maximum. This scaling provides a very modest reduction in time after the first couple of cycles, but it is a nearly 20% effect if one is using a single super-iteration.

[1] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[2] P. W. Shor, SIAM J. Sci. Statist. Comput. **26**, 1484 (1997).

[3] H. W. Lenstra, Ann. Math. **126**, 649 (1987).

[4] C. Pomerance, Comput. Methods Number Theory **0**, 89 (1982).

[5] C. Pomerance, Notices Am. Math. Soc. **43**, 1473 (1996).

[6] R. P. Brent, BIT Numer. Math. **20**, 176 (1980).

[7] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

[8] E. C. Behrman, G. A. Jongeward, and P. G. Wolynes, J. Chem. Phys. **79**, 6277 (1983).

[9] A. Alexandru, G. Basar, P. F. Bedaque, S. Vartak, and N. C. Warrington, Phys. Rev. Lett. **117**, 081602 (2016).

[10] G. Aarts, J. Phys.: Conf. Ser. **706**, 022004 (2016).

[11] S. P. Jordan, K. S. M. Lee, and J. Preskill, Science **336**, 1130 (2012).

[12] E. Zohar and M. Burrello, Phys. Rev. D **91**, 054506 (2015).

[13] I. M. Georgescu, S. Ashhab, and F. Nori, Rev. Mod. Phys. **86**, 153 (2014).

[14] C. W. Bauer, Z. Davoudi, A. B. Balantekin, T. Bhattacharya, M. Carena, W. A. deJong, P. Draper, A. El-Khadra, N. Gemelke, M. Hanada *et al.*, PRX Quantum **4**, 027001 (2023).

[15] L. Funcke, T. Hartung, K. Jansen, and S. Kühn, Proc. Sci. **430**, 228 (2023).

[16] D. Beck, J. Carlson, Z. Davoudi, J. Formaggio, S. Quaglioni, M. Savage, J. Barata, T. Bhattacharya, M. Bishof, I. Cloet *et al.*, arXiv:2303.00113.

[17] A. Y. Kitaev, arXiv:quant-ph/9511026 (1995).

[18] N. J. Ward, I. Kassal, and A. Aspuru-Guzik, J. Chem. Phys. **130**, 194105 (2009).

[19] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **79**, 2586 (1997).

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

[21] J. C. Aulicino, T. Keen, and B. Peng, Int. J. Quantum Chem. **122**, e26853 (2022).

[22] K. Bharti *et al.*, Rev. Mod. Phys. **94**, 015004 (2022).

[23] X.-M. Zhang, T. Li, and X. Yuan, Phys. Rev. Lett. **129**, 230504 (2022).

[24] A. Carrera Vazquez and S. Woerner, Phys. Rev. Appl. **15**, 034027 (2021).

[25] D. Lee, J. Bonitati, G. Given, C. Hicks, N. Li, B.-N. Lu, A. Rai, A. Sarkar, and J. Watkins, Phys. Lett. B **807**, 135536 (2020).

[26] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, Phys. Rev. A **78**, 042307 (2008).

[27] D. B. Kaplan, N. Klco, and A. Roggero, arXiv:1709.08250.

[28] E. J. Gustafson, Phys. Rev. D **101**, 071504(R) (2020).

[29] K. Choi, D. Lee, J. Bonitati, Z. Qian, and J. Watkins, Phys. Rev. Lett. **127**, 040505 (2021).

[30] Z. Qian, J. Watkins, G. Given, J. Bonitati, K. Choi, and D. Lee, arXiv:2110.07747.

[31] M. Bee-Lindgren, Z. Qian, M. DeCross, N. C. Brown, C. N. Gilbreth, J. Watkins, X. Zhang, and D. Lee, arXiv:2208.13557.

[32] I. Stetcu, A. Baroni, and J. Carlson, arXiv:2211.10545.

[33] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, arXiv:quant-ph/0001106.

[34] K. Wan and I. H. Kim, arXiv:2004.04164.

[35] D. Aharonov and A. Ta-Shma, arXiv:quant-ph/0301023.

[36] E. A. Coello Pérez, J. Bonitati, D. Lee, S. Quaglioni, and K. A. Wendt, Phys. Rev. A **105**, 032403 (2022).

[37] A. N. Ciavarella, S. Caspar, M. Illa, and M. J. Savage, Quantum **7**, 970 (2023).

[38] T. Albash and D. A. Lidar, Rev. Mod. Phys. **90**, 015002 (2018).

[39] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Rep. Prog. Phys. **83**, 054401 (2020).

[40] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).

[41] M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).