

Tunneling in projective quantum Monte Carlo simulations with guiding wave functionsT. Parolini,^{1,2,3} E. M. Inack,^{1,2,4,5} G. Giudici,^{1,2,3} and S. Pilati⁶¹*SISSA - International School for Advanced Studies, I-34136 Trieste, Italy*²*The Abdus Salam International Centre for Theoretical Physics, I-34151 Trieste, Italy*³*Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, I-34136 Trieste, Italy*⁴*INFN, Sezione di Trieste, I-34136 Trieste, Italy*⁵*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada N2L 2Y5*⁶*School of Science and Technology, Physics Division, Università di Camerino, 62032 Camerino (MC), Italy*

(Received 31 August 2019; published 6 December 2019)

Quantum tunneling is a valuable resource exploited by quantum annealers to solve complex optimization problems. Tunneling events also occur during projective quantum Monte Carlo (PQMC) simulations, and in a class of problems characterized by a double-well energy landscape their rate was found to scale linearly with the first energy gap, i.e., even more favorably than in physical quantum annealers, where the rate scales with the gap squared. Here we investigate how a guiding wave function—which is essential to make many-body PQMC simulations computationally feasible—affects the tunneling rate. The chosen test beds are a continuous-space double-well problem, the ferromagnetic quantum Ising chain, and the recently introduced shamrock model. As guiding wave function, we consider an approximate Boltzmann-type ansatz, the numerically exact ground state of the double-well model, and a neural-network wave function based on a Boltzmann machine. Remarkably, for each ansatz we find the same asymptotic linear scaling of the tunneling rate that was previously found in the PQMC simulations performed without a guiding wave function. We also provide a semiclassical theory for the double-well with exact guiding wave function that explains the observed linear scaling. These findings suggest that PQMC simulations guided by an accurate ansatz represent a valuable benchmark for physical quantum annealers and a potentially competitive quantum-inspired optimization technique.

DOI: [10.1103/PhysRevB.100.214303](https://doi.org/10.1103/PhysRevB.100.214303)**I. INTRODUCTION**

Quantum annealers are special purpose adiabatic quantum computers designed to solve complex optimization problems [1]. Compared to alternative classical optimization algorithms, chiefly simulated annealing, they can additionally exploit quantum tunneling to cross energy barriers and reach lower energy solutions [2–5]. Their dominant bottlenecks are the small energy gaps associated to avoided level crossings. Such small gaps typically occur in disordered systems when two well-separated competing states are connected by a tunneling process. This scenario frequently happens in the glassy phases that characterize typical hard optimization problems.

Simulating the real-time dynamics of quantum annealers, e.g., to identify classes of problems where they might outperform classical optimization methods, is possible only for relatively small systems (say, around 30 spins). Quantum Monte Carlo (QMC) simulations have emerged as a useful alternative tool to simulate the quantum annealers' behavior in configurations where the sign problem does not occur [1,6–11]. This is the case, e.g., of the devices currently commercialized by *D*-wave systems (see, e.g., [1,12–15]). In particular, path-integral Monte Carlo (PIMC) [16] and projective QMC (PQMC) [17] algorithms have been adopted, beside other techniques such as the stochastic series expansion algorithm [18,19]. Tunneling events also occur during QMC simulations, similarly to what happens during the quantum annealers' dynamics [20,21]. In various problems characterized

by a double-well energy landscape, the tunneling rate of finite-temperature PIMC simulations was found to scale with the system size, or with the height of the energy barrier, as the square of the first energy gap [22–24]. This is the same scaling predicted by the theory of incoherent quantum tunneling [25], and it is also the scaling of the inverse of the annealing time required by a coherent quantum annealer to avoid diabatic transitions [26].

References [23,24,27] explained these results using a semiclassical theory of instantons in PIMC simulations. In the case of PQMC algorithms, the tunneling rate was found to scale linearly with the gap, providing a quadratic speedup compared to the expected behavior of a quantum annealer [28].¹ Furthermore, the PQMC algorithms displayed this speedup even in the so-called shamrock model [28], where frustrated interactions cause an exponential slowdown of the finite-temperature PIMC dynamics [29]. These findings suggest that PQMC simulations constitute a relevant benchmark for physical quantum annealers and a competitive quantum-inspired classical optimization algorithm [30–32]. In fact, they have recently been

¹A linear scaling was identified also for PIMC simulations performed with open boundary conditions in the inverse-temperature direction [23,24]. However, significant deviations have later been discussed [27]. Furthermore, the computational cost of zero-temperature simulations based on open boundary PIMC algorithms has not been analyzed in detail.

employed to obtain better solutions in optimization problems relevant for medical research, specifically, for pulse-sequence optimization in magnetic resonance fingerprinting [33]. This further highlights the importance of exhaustively characterizing their tunneling dynamics.

The tunneling-time studies mentioned above have considered PQMC algorithms implemented without a guiding wave function (GWF). The GWF, usually a variational ansatz that closely approximates the ground state, guides the PQMC simulation towards the relevant regions of the configuration space. This improves the algorithm's accuracy and efficiency [17]. In fact, without a sufficiently accurate GWF, the computational cost of PQMC simulations increases exponentially with the system size [28,34,35]. In principle, one might expect the GWF to significantly impact the tunneling dynamics, since it alters both the sampling algorithm and the probability distribution sampled at equilibrium. In this paper, we analyze whether the GWF does indeed affect the tunneling time in PQMC simulations, and if it does, to what extent. As test beds, we consider a one-dimensional continuous-space Hamiltonian describing a quantum particle in a double well, the ferromagnetic quantum Ising chain, and the shamrock model. Notice that, in the ferromagnetic phase, also the Ising-type models can be described by an effective double-well potential, with the two polarized states with opposite magnetizations representing the two competing potential minima. We consider different kinds of GWFs, including a Boltzmann-type ansatz that mimics the equilibrium distribution of a classical statistical ensemble and, for the continuous-space model, the numerically exact representation of the ground state. For the quantum Ising chain, we also consider an ansatz that mimics the structure of a generative artificial neural network [36], specifically an unrestricted Boltzmann machine [37] or, in a different jargon, a shadow wave function [38,39]. Remarkably, for all GWFs we consider, we find the same linear scaling (to leading exponential order) of the tunneling rate with the gap as previously found in PQMC simulation performed without GWF. The choice of the GWF only affects the prefactor. We also provide a semiclassical theory based on the Wentzel-Kramers-Brillouin (WKB) approximation, valid for PQMC simulations of the double well with exact GWF, that explains the observed linear scaling.

The rest of the paper is organized as follows: In Sec. II we present the PQMC algorithm, implemented with and without GWF, and the analysis of the tunneling time for the continuous-space double-well model. In the same section, the WKB semiclassical theory of the PQMC tunneling time is reported. The algorithm and the tunneling-time analysis for the quantum Ising chain and for the shamrock model are presented in Sec. III. Our conclusions and the outlook are reported in Sec. IV.

II. TUNNELING TIME IN CONTINUOUS-SPACE DOUBLE-WELL SYSTEMS

In this section, we consider a quantum particle in one spatial dimension, described by the following continuous-space Hamiltonian:

$$\hat{H} = -\frac{1}{2} \frac{d^2}{dx^2} + V(x), \quad (1)$$

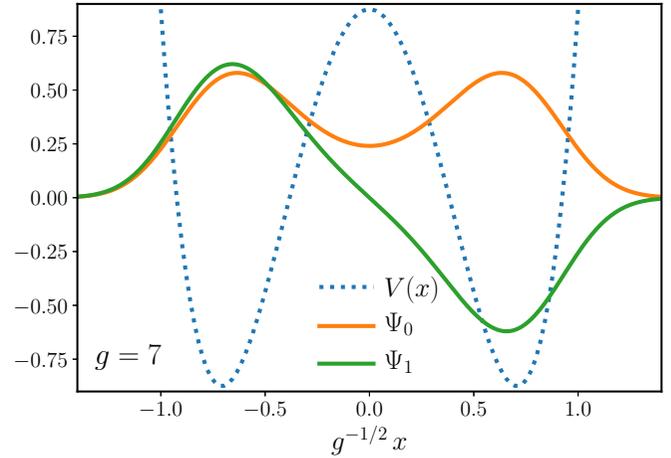


FIG. 1. Profile of the quartic double-well potential Eq. (2) (dotted blue line), shifted for better comparison with the wave functions of the ground state $\Psi_0(x)$ and of the first excited state $\Psi_1(x)$ (orange and green solid lines, respectively).

with the quartic double-well potential

$$V(x) = \frac{x^4}{g} - x^2. \quad (2)$$

The profile of $V(x)$ is visualized in Fig. 1, together with the corresponding ground-state wave function $\Psi_0(x)$ and the first excited state $\Psi_1(x)$. Units are chosen so that $\hbar/m = q = 1$, where \hbar is the reduced Planck constant, m is the particle mass, and $q = \frac{1}{2} \sqrt{V''(x_{L,R})}$ fixes the curvature at the bottom of the well. Here $x_{L,R} = \mp \sqrt{g/2}$ are the minimum points of $V(x)$. The dimensionless parameter g controls the height of the barrier separating the two wells, $V(0) - V(x_L) \propto g$, and the distance between the minimum points, $x_R - x_L \propto \sqrt{g}$. For large g , the barrier height increases relative to the wells' separation and the energy spectrum becomes doubly degenerate, corresponding to the two wells being asymptotically independent. For large but finite g , the two wells are connected by tunneling processes. These processes lift the degeneracies, leading to small tunneling gaps between the energy levels. An approximate expression for the first energy gap $\Delta = E_1 - E_0$ between the ground-state energy E_0 and the first excited level E_1 , valid in the large- g regime, can be determined via the WKB semiclassical theory. One obtains [40]

$$\Delta = 8 \sqrt{\frac{g}{\pi}} \exp\left(-\frac{2g}{3}\right). \quad (3)$$

In an isolated double well, a quantum particle initially prepared in one of the two states,

$$\Psi_{L,R}(x) = \frac{\Psi_0(x) \pm \Psi_1(x)}{\sqrt{2}}, \quad (4)$$

which are localized in the left and right wells, respectively, performs coherent periodic oscillations between the two states at a rate proportional to Δ . As a reference, it is worth mentioning that, instead, a quantum particle coupled to a thermal bath and subjected to a double-well potential would undergo incoherent tunneling at a rate proportional to Δ^2 [25].

A. Diffusion Monte Carlo algorithm

We are interested in comparing the tunneling dynamics generated by PQMC simulations to the quantum tunneling time, and also in identifying their relation with the energy gap Δ . Specifically, we consider simulations performed with the diffusion Monte Carlo (DMC) method, which belongs to the family of PQMC algorithms for continuous-space systems. The DMC algorithm aims at projecting out the ground state by evolving the imaginary-time Schrödinger equation for the time-dependent wave function $\Psi(x, t)$. In one dimension, this equation reads

$$\frac{\partial \Psi(x, t)}{\partial t} = \frac{1}{2} \frac{\partial^2 \Psi(x, t)}{\partial x^2} - [V(x) - E_T] \Psi(x, t). \quad (5)$$

Here, t is the imaginary time and E_T is an energy threshold introduced to stabilize the numerics, as explained later. Following the seminal work of Ref. [41], one can exploit the analogy between Eq. (5) and a modified diffusion equation, i.e., one including a source/sink process corresponding to the second term on the right-hand side, to implement a stochastic simulation. Due to the norm-nonconserving nature of this source/sink term, the simulation has to evolve a large population of equivalent instances of system configurations $\{x_i\}$, in jargon called walkers, subject to diffusion and to a killing and replication process called branching. The sampling algorithm is dictated by a short-time approximation of the Green function of Eq. (1). This is assumed to apply for a short imaginary-time step τ . Long imaginary times $t = N_{\text{DMC}}\tau$ can be reached by iterating many DMC steps N_{DMC} , each corresponding to a time step τ . In the long imaginary-time limit, the walkers sample configurations with a probability distribution proportional to $\Psi(x, t \rightarrow \infty) \propto \Psi_0(x)$. Notice that $\Psi_0(x)$ is assumed to be a real and non-negative function. This is legitimate in the case of stoquastic Hamiltonians, i.e., those not affected by the negative-sign problem [42,43]. We adopt the linear Trotter approximation for the short-time Green function. This leads to the following sampling algorithm. After initializing, say, N_w walkers in some configurations $\{x_i\}$, at every DMC step one applies to each walker i two processes: First, the configuration update $x'_i = x_i + \sqrt{\tau}\delta_i$, where δ_i is sampled from a zero-mean, unit-variance Gaussian distribution; then, one samples an integer number $s_i = \lfloor w_i + r_i \rfloor$, where $w_i = \exp[-\tau(V(x'_i) - E_T)]$ is the walker weight and $r_i \in (0, 1)$ is a uniform random variable, and creates s_i copies of the walker to be included in the population for the next time step. The size of the walker population can be tuned close to a desired target value N_w by appropriately controlling E_T . For this control, we adopt the textbook recipe of Ref. [44]. Therein, the interested readers will also find a more pedagogical description of the DMC algorithm. The systematic bias due to the Trotter approximation can be eliminated via zero time-step extrapolation. A more subtle bias might originate also from the finite walker population N_w [34,35,45]. Indeed, it has been shown that in many-body systems, in order to keep this bias below a chosen small threshold, N_w has to exponentially increase with the system size [28]. In order to reduce or eliminate this systematic bias and to reduce the statistical fluctuations, it is standard practice to adopt an importance sampling approach using a guiding wave function $\Psi_G(x)$ [17]. Typically, $\Psi_G(x)$ is

a parametrized variational ansatz, whose optimal parameters are determined via energy expectation-value minimization. Hence, one evolves the modified imaginary-time Schrödinger equation for the product $\rho(x, t) = \Psi(x, t)\Psi_G(x)$, which takes the form of a Fokker-Planck-type equation with an additional source/sink term. This equation reads

$$\frac{\partial \rho}{\partial t} = \frac{1}{2} \frac{\partial^2 \rho}{\partial x^2} + \frac{\partial}{\partial x} [\tilde{V}'(x)\rho] - [E_L(x) - E_T]\rho, \quad (6)$$

where we wrote ρ for $\rho(x, t)$ to simplify the notation, the local energy is

$$E_L(x) = V(x) - \frac{1}{2\Psi_G(x)} \frac{d^2 \Psi_G(x)}{dx^2}, \quad (7)$$

and we introduced the effective potential

$$\tilde{V}(x) = -\ln \Psi_G(x). \quad (8)$$

Adopting, again, a linear approximation for the Green function, the sampling algorithm is modified as follows: The configuration update includes, beside the Gaussian random term, a deterministic displacement computed as $\tau \frac{d}{dx} \ln \Psi_G(x)$; in the branching process, the potential $V(x)$ is replaced by $E_L(x)$ for the computation of the walker weight w_i . In the long imaginary-time limit, the walkers sample a probability distribution proportional to $\rho(x, t \rightarrow \infty) = \Psi_0(x)\Psi_G(x)$. One should notice that $E_L(x)$ is a constant function if $\Psi_G(x)$ is an exact eigenstate of the Hamiltonian. This completely suppresses the fluctuations of the random walker population, eliminating the finite- N_w bias. In fact, the algorithm's accuracy and efficiency are significantly improved even when $\Psi_G(x)$ is, albeit not exact, a reasonably good approximation of the ground-state wave function.

B. Tunneling time in diffusion Monte Carlo simulations

Our goal is to analyze the relation between the DMC tunneling time ξ , namely the imaginary time required by the walkers to leak from one well to the other, and the physical tunneling time of the real-time dynamics. More precisely, we are interested in the scaling relation between ξ and the inverse energy gap Δ^{-1} . We measure ξ with a protocol analogous to the one adopted in Refs. [23,24] for PIMC simulations. All walkers are initially set at the bottom of the left well $x = x_L$. The DMC simulation is run until a percentage p of the instantaneous walker population overcomes a position threshold in the right well $x_{\text{th}} \geq 0$. The final imaginary time is recorded, and the process is repeated approximately 300 times to accumulate statistics. The fluctuations of the final imaginary times turn out to be approximately normally distributed, and we take the average and its standard deviation as the definition of ξ and of its error bar, respectively. In the simulations reported here, the threshold is set at $x_{\text{th}} = x_R/2$ and the walker percentage at $p = 25\%$. A careful analysis shows that the asymptotic scaling of ξ is independent of this specific choice up to a constant prefactor. Furthermore, the chosen target walker population $N_w \approx 10^4$ is large enough, and the chosen time step small enough (e.g., $\tau = 0.007$ for $1/\Delta > 70$), to eliminate any significant effect on ξ .

First, the DMC algorithm is run without a guiding wave function (GWF). We measure the tunneling time ξ for

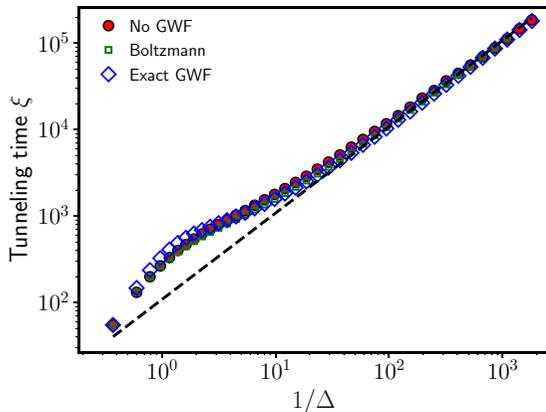


FIG. 2. DMC tunneling time ξ for the quartic double-well potential (2) as a function of the inverse energy gap Δ^{-1} . Three different DMC protocols are shown: The simple DMC algorithm without a GWF (red circles), the DMC algorithm guided by a Boltzmann ansatz (green empty squares), and the one guided by the numerically exact representation of the ground state $\Psi_0(x)$ (blue empty diamonds). The dashed line represents the scaling $\xi \propto \Delta^{-1}$. Here and in all plots, if not visible, the error bars are smaller than the symbol size.

different barrier heights, tuned by varying the parameter g . In Fig. 2, ξ is plotted as a function of the inverse energy gaps Δ^{-1} , which we compute for the different g values we consider using a standard finite-difference method. The discretization is fine enough to ensure that there is no sizable finite-precision effect. In the large- g regime, corresponding to large Δ^{-1} , the tunneling times approach the scaling law $\xi \propto \Delta^{-1}$. This scaling relation has previously been identified in Ref. [28] in QMC simulations of Ising-type models, again performed without a GWF. Next, we run DMC simulations with a GWF. First, as GWF we consider the approximate Boltzmann ansatz $\Psi_G(x) = \exp[-\beta V(x)]$. The fictitious inverse temperature β is fixed by minimizing the variational energy estimate. Second, we consider a numerical representation of the exact ground-state wave function, i.e., we set $\Psi_G(x) = \Psi_0(x)$. The ground-state wave function is obtained via the finite-difference technique. Notice that this ansatz represents the optimal GWF for equilibrium simulations. In Fig. 2, the tunneling times obtained with these two GWFs are compared with the results obtained without GWF. Remarkably, for large g , the same linear relation between ξ and Δ^{-1} is approached. By fitting the three data sets in the large- g regime with the function $\xi(\Delta) = \alpha \Delta^{-b}$, where α and b are the fitting parameters, we obtain the values reported in Table I. In all three cases, the exponent b is consistent with the linear relation corresponding to $b = 1$. The choice of GWF only affects the prefactor α , though for this model the variations are small enough to be masked by statistical uncertainties. These findings indicate that the GWF does not affect the leading scaling relation between tunneling time and inverse energy gap. This is a surprising results, given that introducing the GWF affects both the sampling algorithm and the equilibrium probability distribution of the DMC simulation. A rough explanation can be conjectured by considering the competition between two effects originating from the introduction of the GWF. The first is due to the deterministic drift, which pushes

TABLE I. Fitting parameters α and b , describing the small-gap behavior of the DMC tunneling time ξ in the double-well potential (2) according to the fitting function $\xi(\Delta) = \alpha \Delta^{-b}$. For each protocol, the five rightmost data points shown in Fig. 2 are included in the fit.

DMC	α	b
No GWF	109(11)	0.99(2)
Boltzmann	98(16)	1.01(3)
Exact GWF	112(8)	0.99(1)

walkers away from the potential barrier, inhibiting inter-well crossings. The second is the smoothing out of the weight reduction that occurs when walkers encounter a bump in the potential; this effect reduces the probability of those walkers being eliminated from the population. Our numerical results indicate that these two effects tend to compensate. A more formal explanation of the linear relation between ξ and Δ^{-1} is given in the next subsection.

The double-well potential (2) is characterized by a specific functional relation between the barrier's height and width. In order to verify that our findings do not rely on this particular choice, we introduce an adjustable parameter that allows us to vary the width of the barrier independently of its height. Similarly to Ref. [24], we consider the potential

$$U(x) = \frac{(|x| - x_0)_+^4}{g} - (|x| - x_0)_+^2, \quad (9)$$

where $f(x)_+ \equiv \max\{0, f(x)\}$ and $x_0 \geq 0$. This potential features a *plateau* of width $2x_0$ around the origin, and reduces to $V(x)$ for $x_0 = 0$. In our study, the barrier height is kept constant by fixing g to some value (we use $g = 8$), while x_0 is increased in the interval $x_0 \in [0, 2]$. This has the effect of reducing the gap Δ as well as the tunneling rate $1/\xi$. The GWF chosen for this study is the numerically exact ground-state wave function. Once again, a linear relation is found between ξ and Δ^{-1} , and fitting the dataset with the power law $\xi(\Delta) = \alpha \Delta^{-b}$ in the small-gap regime yields the values $\alpha = 23(1)$ and $b = 0.993(7)$ for the parameters.

C. Semiclassical theory of the DMC tunneling dynamics

We present here a semiclassical theory to explain and generalize our findings. As mentioned above, the DMC algorithm with GWF is described by a Fokker-Planck equation, Eq. (6), containing both the usual drift and diffusion terms, and an additional norm-nonpreserving term corresponding to the branching process. When the GWF coincides with the exact ground-state wave function $\Psi_G(x) = \Psi_0(x)$, this term can be eliminated and branching does not occur. The resulting equation reads

$$\frac{\partial \rho(x, t)}{\partial t} = \frac{1}{2} \frac{\partial^2 \rho(x, t)}{\partial x^2} + \frac{\partial}{\partial x} [\tilde{V}'(x) \rho(x, t)]. \quad (10)$$

This equation describes the stochastic dynamics of a classical particle with distribution $\rho(x, t)$ subject to the effective potential $\tilde{V}(x)$ defined in Eq. (8). The tunneling time corresponding to this stochastic dynamics can then be identified (to exponential accuracy) with the activation time needed for this

classical ensemble to overcome the effective barrier $\Delta\tilde{V} \equiv \tilde{V}(0) - \tilde{V}(x_{\min})$, with initial conditions $\rho(x, 0) = \delta(x - x_L)$ [the normalization can be set to one since the norm of $\rho(x, t)$ is conserved]. Here, x_{\min} indicates the (left) minimum point of $\tilde{V}(x)$, in general different from x_L , while $x = 0$ is its local maximum point, as follows from the form of $\Psi_0(x)$. The computation of the classical activation over a potential barrier is known as Kramers problem [46,47]. The corresponding activation time reads

$$\tau_{\text{act}} = \frac{2\pi}{\sqrt{|\tilde{V}''(x_{\min})\tilde{V}''(0)|}} \exp(2\Delta\tilde{V}). \quad (11)$$

The large- g scaling of τ_{act} is dominated by the exponential function in Eq. (11), so we will focus on that term only. Substituting Eq. (8) into Eq. (11), one obtains

$$\tau_{\text{act}} \sim \left(\frac{\Psi_0(0)}{\Psi_0(x_{\min})} \right)^{-2} \sim \Psi_0(0)^{-2}. \quad (12)$$

Here, we used the fact that $\Psi_0(x_{\min})$ cannot be exponentially small in g if the ground-state wave function is to be normalized to unity [notice that, by definition, $\Psi_0(x)$ achieves its global maximum at x_{\min} , and its amplitude is exponentially small outside of a region of width $O(\sqrt{g})$]. The value of $\Psi_0(0)$ can be estimated using WKB theory, which gives $\Psi_0(0) \sim \exp(-g/3)$ [48]. Substituting this into Eq. (12), and combining with the WKB estimate of the gap, Eq. (3), we finally get

$$\xi \sim \tau_{\text{act}} \sim \frac{1}{\Delta}, \quad (13)$$

where the symbol \sim denotes asymptotic equal scaling for $g \rightarrow \infty$, up to subexponential corrections.

The above argument can actually be generalized to double-well-type potentials other than the quartic double well defined in Eq. (2). In fact, Eq. (12) can be derived from Eq. (10) under quite general assumptions on $V(x)$, as long as a sufficiently accurate approximant of the ground-state wave function is available to use as a GWF, such that the effect of branching may be neglected. Moreover, for a generic double-well potential the ground-state gap in the large-barrier limit can be expressed as [40,49]

$$\Delta \propto \Psi_R(0)\Psi'_R(0), \quad (14)$$

where $\Psi_R(x)$ is the ground-state wave function of a *single* well, as defined in Eq. (4). The above relation holds in the only hypothesis that $\Psi_R(x)$ is asymptotically localized on $x > 0$, such that the condition $\int_{-\infty}^0 [\Psi_R(x)\Psi_R(-x) - \Psi_R(x)^2]dx \ll 1$ is fulfilled for large enough g . We can now resort to WKB theory, which gives the following expression for $\Psi_R(x)$ (to exponential accuracy):

$$\Psi_R(x) = \exp\left(-\int_x^a ds \sqrt{2[V(s) - E_0]}\right), \quad (15)$$

where E_0 is the ground-state energy and a is the (positive) classical turning point, defined by $V(a) = E_0$. This implies that, up to subexponential corrections, $\Delta \propto \Psi_R(0)^2 \propto \Psi_0(0)^2$. Upon substitution into Eq. (12), this leads to the scaling relation (13). Therefore, in the assumption that WKB theory holds in a neighborhood of the origin (e.g., assuming

that the turning points do not approach 0 in the infinite-barrier limit), we see that the DMC tunneling time generically exhibits a Δ^{-1} scaling regardless of the specific form of the potential.

III. TUNNELING TIME IN DISCRETE-BASIS MODELS

In this section we investigate the tunneling time in PQMC simulations of discrete-basis models, specifically, of quantum spin models. The first model we consider is the quantum Ising chain, described by the following Hamiltonian:

$$\hat{H} = \hat{H}_{\text{cl}} + \hat{H}_{\text{kin}}, \quad (16)$$

where $\hat{H}_{\text{cl}} = -J \sum_{i=1}^N \sigma_i^z \sigma_{i+1}^z$ and $\hat{H}_{\text{kin}} = -\Gamma \sum_{i=1}^N \sigma_i^x$. σ_i^x , σ_i^y , and σ_i^z denote Pauli matrices acting on spins at the lattice site i . N is the total number of spins, and we use periodic boundary conditions, i.e., $\sigma_{N+1}^a = \sigma_1^a$, with $a = x, y, z$. The parameter $J > 0$ fixes the strength of the ferromagnetic interactions between nearest-neighbor spins. In the following, we set $J = 1$. Γ is the transverse field intensity. Given $|x_i\rangle$ an eigenstate of σ_i^z having eigenvalue $x_i = 1$ when $|x\rangle = |\uparrow\rangle$ and $x_i = -1$ when $|x\rangle = |\downarrow\rangle$, the quantum state of N spins is indicated by $|x\rangle = |x_1 x_2 \dots x_N\rangle$. The set $\{|x\rangle\}$ of 2^N states forms the computational basis.

At zero temperature, in the ferromagnetic phase $\Gamma < J$, the quantum Ising chain is characterized by an energy landscape with an effective double-well potential, where the magnetization per spin M/N plays the role of reaction coordinate. The minima of the potential connected by the reaction coordinate correspond to the classical states with magnetization $M \simeq \pm N$. These two states are degenerate when $\Gamma = 0$. For finite Γ , provided $\Gamma < J$, quantum fluctuations induce tunneling processes between the two minima, lifting the degeneracy in finite systems. The energy gap between ground state and first excited state is exponentially small in the system size, i.e., $\Delta \propto \exp(-cN)$, where the constant c depends on the transverse field. This closing-gap scenario resembles the Landau-Zener avoided level crossings one typically encounters in adiabatic quantum optimization. There, the small gaps are associated to tunneling processes between competing solutions. In order to avoid diabatic transitions to the first excited state, the total annealing time has to scale as Δ^{-2} . These small gaps represent the bottleneck of adiabatic quantum computing, since for hard optimization problems these gaps often close exponentially fast with the system size [50–54].

A. PQMC simulations of discrete-basis models

Our PQMC simulations for discrete-basis models are based on the continuous-time Green-function Monte Carlo algorithm [55]. This method is exhaustively described in Ref. [56]. Here, we only sketch the main elements. The simulations with importance sampling are implemented by stochastically evolving the modified imaginary-time Schrödinger equation for the product $\rho(x, t) = \Psi(x, t)\Psi_G(x)$:

$$-\frac{\partial}{\partial t} \rho(x, t) = \sum_{x'} [H_{x,x'} - E_T \delta_{x,x'}] \frac{\Psi_G(x)}{\Psi_G(x')} \rho(x', t). \quad (17)$$

Here, $\Psi(x, t) \equiv \langle x | \Psi(t) \rangle$ is the amplitude of the ground-state wave function at imaginary time t , which is assumed to

be real and non-negative. The Hamiltonian matrix elements are $H_{\mathbf{x},\mathbf{x}'} = \langle \mathbf{x} | \hat{H} | \mathbf{x}' \rangle$. E_T is again a reference energy used to stabilize the simulations. When no GWF is used, which corresponds to setting $\Psi_G(\mathbf{x}) = 1$, Eq. (17) becomes the standard imaginary-time Schrödinger equation. Analogously to the continuous-space simulations described in Sec. II A, one evolves a population of walkers undergoing spin-flip updates and branching. An accurate GWF favors updates toward relevant regions of the configuration space and diminishes walker killings and replications. In the infinite imaginary-time limit $t \rightarrow \infty$, attained by iterating many small time steps τ , the walkers sample spin configurations with a probability distribution proportional to $\Psi_0(\mathbf{x})\Psi_G(\mathbf{x})$, where $\Psi_0(\mathbf{x})$ is the ground-state wave function.

In this section we consider two types of GWF. The first is the Boltzmann ansatz:

$$\Psi_G(\mathbf{x}) = \exp[-\beta E_{\text{cl}}(\mathbf{x})]. \quad (18)$$

It resembles the Boltzmann distribution of a classical Ising model with Hamiltonian function $E_{\text{cl}}(\mathbf{x}) = \langle \mathbf{x} | \hat{H}_{\text{cl}} | \mathbf{x} \rangle$. The fictitious temperature β is fixed by minimizing the variational energy $\frac{\langle \Psi_G | \hat{H} | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle}$ using the stochastic gradient descent method [55]. The other GWF is based on a stochastic generative neural network, specifically an unrestricted Boltzmann machine (uRBM) [37]; it is defined as

$$\Psi_G(\mathbf{x}) = \sum_{\mathbf{h}} \phi(\mathbf{x}, \mathbf{h}), \quad (19)$$

where

$$\phi(\mathbf{x}, \mathbf{h}) = \exp \left[\sum_{i=1}^N (K_1 x_i x_{i+1} + K_2 h_i h_{i+1} + K_3 x_i h_i) \right]. \quad (20)$$

The wave function amplitude in each physical, or visible, spin configuration $\mathbf{x} = (x_1, x_2, \dots, x_N)$ is obtained by integrating over all configurations of the N hidden units $\mathbf{h} = (h_1, h_2, \dots, h_N)$, which take the values $h_i = \pm 1$ (with $i = 1, \dots, N$). Periodic boundary conditions are considered both in the visible and in the hidden layers, i.e., $x_{N+1} = x_1$ and $h_{N+1} = h_1$. The three coupling constants K_1 , K_2 , and K_3 fix the interaction strengths between nearest-neighbor visible and hidden spins, and between visible-hidden pairs with the same index, respectively. We determine them via variational optimization using the stochastic reconfiguration method. The analysis reported in Ref. [37] indicated that the optimized uRBMs GWFs are sufficiently accurate to reduce the computational cost of PQMC simulations of the quantum Ising chain down to a polynomial scaling with system size. Differently from the restricted Boltzmann machine originally introduced as a variational ansatz in Ref. [36], the uRBM includes intralayer interactions. In general, this implies that one cannot analytically trace out the hidden spin configurations.² In order to use uRBMs as GWFs, we employ the extended

²As shown in Ref. [57], the uRBM can be mapped to a constrained matrix product state. In one dimension, this representation allows for an analytical treatment of the hidden degrees of freedom. However, we aim at a general framework that could be applied irrespectively of the dimensionality and the interaction range.

PQMC algorithm described in Ref. [37]. It includes a certain number of additional single-spin Metropolis updates of the hidden spins at every PQMC time step. This number has to be made large enough to eliminate spurious correlations among successive walker configuration, which in turn affect the finite- N_w bias. It is quite important to test if and how the possible residual statistical correlations between successive hidden-spin configurations affect the tunneling dynamics.

B. Tunneling times in quantum Ising chains

The tunneling time simulations are performed in the ferromagnetic phase, where the quantum Ising chain is characterized by a double-well potential profile. To measure the PQMC tunneling time ξ , we adopt the protocol of Ref. [28]. All walker configurations are initialized with all spin pointing up, corresponding to the classical state with magnetization $M = \sum_i x_i = N$. This state is close to one of the minima of the effective double-well potential. The PQMC simulation is run until 10% of the walkers has crossed the potential barrier, reaching negative magnetization $M < 0$. The measurement is repeated about 1000 times, taking the average and its standard deviation as definition of ξ and of its error bar, respectively. The PQMC simulations are performed with $N_w = 5000$ – $10\,000$, which is found to be sufficient to eliminate any systematic error on ξ .

Figure 3 displays the tunneling time ξ obtained with the Boltzmann GWF, as a function of the number of spins N , for different transverse field intensities Γ . In the large system-size regime, where the energy gap Δ is small, the exponential growth of ξ closely matches the scaling of the inverse energy gap $\alpha\Delta^{-1}$, where α is an appropriate prefactor. The energy gap values are computed using the exact formula obtained from the free fermion representation of the quantum Ising chain. The tunneling times obtained with the Boltzmann GWF

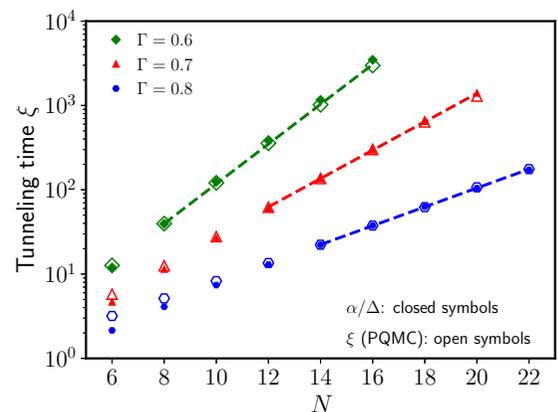


FIG. 3. Tunneling time ξ in PQMC simulations performed with the Boltzmann GWF (open symbols) as a function of the number of spins N in the ferromagnetic Ising chain. Different data sets correspond to different transverse field intensities Γ , with the coupling parameter $J = 1$. The dashed lines represent exponential fitting functions valid in the large- N regime. The closed symbols represent the inverse gap values Δ^{-1} computed with the exact formula obtained from the free fermion representation of the quantum Ising chain and rescaled by an appropriate prefactor $\alpha = O(1)$.

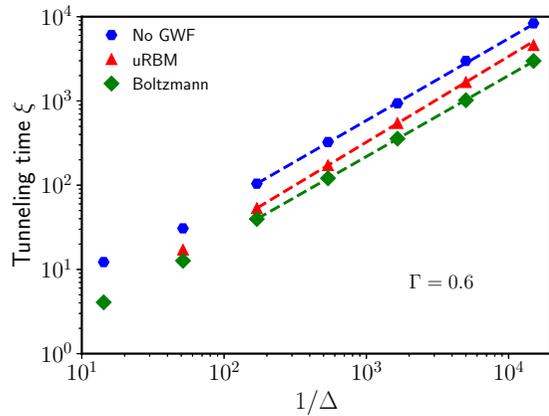


FIG. 4. Tunneling times ξ of the PQMC algorithm implemented without GWF (blue pentagons), with the uRBM GWF (red triangles), and with the Boltzmann GWF (green diamonds), as a function of the inverse energy gap Δ^{-1} , for the quantum Ising chain at $\Gamma = 0.6$. The dashed lines represent the fitting function $\xi(\Delta) = \alpha\Delta^{-b}$, valid in the large Δ^{-1} regime. In all three cases, the fitted exponent is $b \simeq 1$ (see Table II).

and with the uRBM GWF are plotted in Fig. 4 as a function of the corresponding inverse energy gap Δ^{-1} . They are also compared with the results obtained without a GWF (data from Ref. [28]). In all three cases, ξ appears to scale asymptotically linearly with the inverse gap. By fitting the three data sets, in the large Δ^{-1} regime, with the scaling law $\xi(\Delta) = \alpha\Delta^{-b}$, we obtain the values of the fitting parameters α and b reported in Table II. In all cases, the exponent is consistent with the linear scaling $b = 1$.

Shamrock model

To further study the PQMC tunneling dynamics, we address a more challenging quantum spin Hamiltonian, namely the so-called shamrock model. It is described by the following Hamiltonian:

$$\hat{H} = -J\sigma_1^z \sum_{i=2}^N \sigma_i^z + (J - \epsilon) \sum_{i=1}^K \sigma_{2i}^z \sigma_{2i+1}^z - \Gamma \sum_{i=1}^N \sigma_i^x. \quad (21)$$

The N spins are grouped in K rings, which form the leaves of the shamrock. A central spin interacts with a ferromagnetic coupling J with all the other spins in the system. The outer spins in each ring interact antiferromagnetically with interaction energy $J - \epsilon$ where $\epsilon \ll J$ indicates a small interaction energy. Γ is the intensity of the transverse magnetic field. The

TABLE II. Fitting parameters α and b , describing the small-gap behavior of the PQMC tunneling time ξ in the ferromagnetic quantum Ising chain (16), according to the fitting function $\xi(\Delta) = \alpha\Delta^{-b}$. The error bars also take into account the fluctuations due to choosing different fitting windows.

PQMC	α	b
No GWF	0.7(2)	0.97(3)
uRBM	0.32(9)	1.00(2)
Boltzmann	0.28(5)	0.96(3)

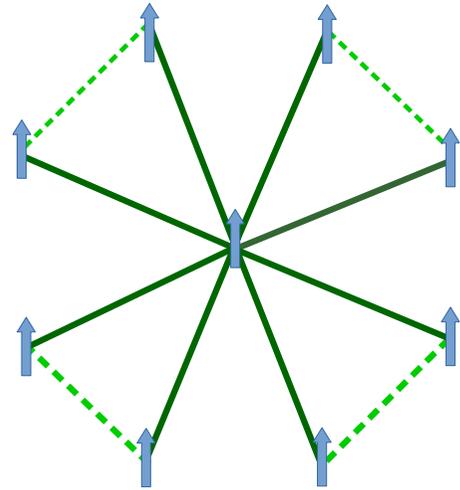


FIG. 5. The shamrock, a model of N frustrated spins in a transverse field. It is made of $K = (N - 1)/2$ leaves, each having three spins. The solid dark-green lines represent ferromagnetic interactions (with interaction strength J) between the central spin and all the other $N - 1$ spins. The dashed light-green lines indicate the antiferromagnetic interactions (with interaction strength $J - \epsilon$) between the outer spins of the same leaf. The overall effect is to create 2^K tunneling paths between the degenerate classical ground states.

connectivity structure of the shamrock model is visualized in Fig. 5. This model was introduced in Ref. [29] as a paradigmatic case where finite-temperature PIMC algorithms cannot efficiently simulate quantum annealing. This was interpreted as an indication that quantum annealing devices have a high potential to provide a quantum speedup in certain classes of optimization problems. It was indeed shown that in this model the PIMC tunneling time scales as $\xi \propto 2^K \Delta^{-2}$, i.e., exponentially worse than the incoherent quantum tunneling

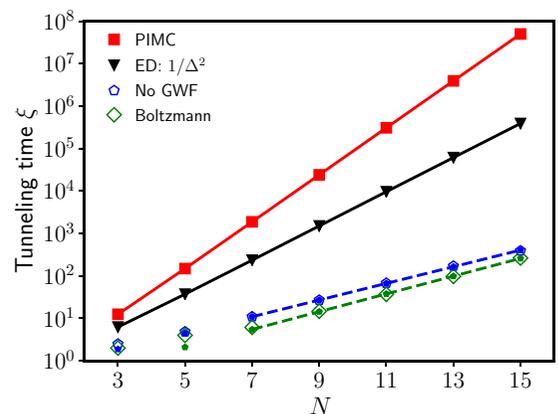


FIG. 6. Tunneling time ξ in the shamrock model as a function of the system size N . The PQMC results obtained with the Boltzmann GWF (green diamonds) and without GWF (blue pentagons) are compared with the scaling of the incoherent quantum tunneling time $1/\Delta^2$ (black triangles), and with the scaling of the finite-temperature PIMC tunneling time $\xi = 2^K/\Delta^2$ [29], where K is the number of leaves in the shamrock. The values of the gap Δ are obtained from exact diagonalization. The model parameters are $\Gamma = 0.5$, $J = 6$, and $\epsilon = 0.2$.

times $\propto \Delta^{-2}$. This slowdown of the PIMC tunneling dynamics originates from the emergence of multiple homotopy-inequivalent paths for tunneling processes between the competing states. We measure the PQMC tunneling times in the shamrock model using the Boltzmann GWF. Notice that in this case the classical energy function in Eq. (18) includes the first two terms of the shamrock Hamiltonian (21). In Fig. 6, these tunneling times are compared with the PQMC results obtained without GWF (data from Ref. [28]), with the scaling of incoherent quantum tunneling, and with the scaling of the PIMC tunneling times. In the large Δ^{-1} regime, the PQMC data with GWF are well described by the fitting function $\xi(\Delta) = \alpha \Delta^{-b}$, where the fitting parameters are $\alpha = 0.32(7)$ and $b = 1.04(3)$. A similar fit, with $b = 0.98(2)$, applies also to the previously reported data, obtained without GWF. These results indicate that even in the shamrock model the PQMC tunneling times asymptotically scale with the inverse gap, independently on the choice of GWF, confirming the quadratic speedup compared to incoherent quantum tunneling.

IV. CONCLUSIONS

We have investigated how guiding wave functions affect the tunneling dynamics of PQMC simulations, considering as test beds a continuous-space double-well problem, the ferromagnetic quantum Ising chain, and the shamrock model with frustrated couplings. As GWFs, both approximate variational *Ansätze* and the numerically computed exact ground-state wave function have been addressed. Remarkably, for all GWFs we find a linear relation between tunneling rate and first energy gap in the asymptotic regime of large tunneling time, corresponding to a high potential barrier in the double well, or to large system sizes in the two Ising-type models. The semiclassical theory we provided explains this linear relation in the case of double-well-type potentials when the exact ground-state wave function is chosen as GWF. It is worth stressing that this linear relation represents a quadratic speedup compared to the expected tunneling rate of a physical quantum annealer. The proof we presented relies on the local validity of the semiclassical approximation for the ground-state wave function. It is an interesting challenge to try to formulate a more general derivation which does not invoke WKB theory, or to instead exhibit a counterexample where a violation of the Δ^{-1} scaling may be observed.

Analyzing if and to what extent QMC algorithms can efficiently simulate quantum tunneling is of critical importance

to understand if quantum annealing devices can outperform classical optimization methods. It is well known that accurate PQMC simulations of the equilibrium properties of large-scale systems are only feasible if a sufficiently accurate GWF is used for importance sampling [17]. Indeed, it has been shown that in the quantum Ising chain the GWF can even change the scaling of the computational cost from being exponential to being polynomial in the system size [37]. The results reported here show that an accurate GWF does not alter the quadratic speedup previously reported for PQMC simulations performed without GWF [28]. Considered together, these findings indicate that PQMC simulations performed with accurate GWFs allow one to efficiently simulate both the equilibrium ground-state properties and also the tunneling dynamics of quantum annealers. Therefore, they can be used as a relevant benchmark in the development of novel quantum annealing devices, and they represent a promising quantum-inspired optimization algorithm [33]. Clearly, more challenging models should be addressed to further benchmark the efficiency of the PQMC tunneling dynamics. Relevant test beds could be Ising spin glasses in higher dimensions. Indeed, one expects that it is harder to obtain accurate variational *Ansätze* for such models. Suitable candidates are restricted [36,58] and unrestricted Boltzmann machines [37,59]. Indeed, both have been shown to be amenable to be used as GWFs in PQMC simulations [37,60]. Deeper neural network *Ansätze*, e.g., the deep convolutional neural networks of Ref. [61] or the recurrent neural networks of Ref. [62], might also be adopted. We leave these studies for future investigations.

ACKNOWLEDGMENTS

We acknowledge useful discussions with A. Gambassi, G. B. Mbeng, G. Santoro, and M. Troyer. S.P. acknowledges financial support from the FAR2018 project of the University of Camerino and from the Italian MIUR under the project PRIN2017 CEnTraL 20172H2SC4. S.P. also acknowledges the CINECA award under the ISCRA initiative, for the availability of high performance computing resources and support. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Economic Development, Job Creation and Trade. G.G. acknowledges support by the ERC under Grant No. 758329 (AGEnTh).

-
- [1] S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, Evidence for quantum annealing with more than one hundred qubits, *Nat. Phys.* **10**, 218 (2014).
 - [2] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, Theory of quantum annealing of an Ising spin glass, *Science* **295**, 2427 (2002).
 - [3] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
 - [4] T. Albash and D. A. Lidar, Adiabatic quantum computation, *Rev. Mod. Phys.* **90**, 015002 (2018).
 - [5] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Perspectives of quantum annealing: Methods and implementations, [arXiv:1903.06559](https://arxiv.org/abs/1903.06559).
 - [6] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, Quantum annealing: A new method for minimizing multidimensional functions, *Chem. Phys. Lett.* **219**, 343 (1994).
 - [7] R. Martoňák, G. E. Santoro, and E. Tosatti, Quantum annealing by the path-integral Monte Carlo method: The two-dimensional random Ising model, *Phys. Rev. B* **66**, 094203 (2002).

- [8] L. Stella and G. E. Santoro, Quantum annealing of an ising spin-glass by green's function monte carlo, *Phys. Rev. E* **75**, 036703 (2007).
- [9] B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, Quantum versus classical annealing of ising spin glasses, *Science* **348**, 215 (2015).
- [10] T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar, Re-examining classical and quantum models for the d-wave one processor, *Eur. Phys. J. Spec. Top.* **224**, 111 (2015).
- [11] G. B. Mbeng, L. Privitera, L. Arceci, and G. E. Santoro, Dynamics of simulated quantum annealing in random ising chains, *Phys. Rev. B* **99**, 064201 (2019).
- [12] M. Johnson, M. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. Berkley, J. Johansson, P. Bunyk *et al.*, Quantum annealing with manufactured spins, *Nature (London)* **473**, 194 (2011).
- [13] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, Experimental signature of programmable quantum annealing, *Nat. Commun.* **4**, 2067 (2013).
- [14] T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, Defining and detecting quantum speedup, *Science* **345**, 420 (2014).
- [15] T. Lanting, A. J. Przybysz, A. Y. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk, N. Dickson, C. Enderud, J. P. Hilton, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, R. Neufeld, T. Oh, I. Perminov *et al.*, Entanglement in A Quantum Annealing Processor, *Phys. Rev. X* **4**, 021041 (2014).
- [16] D. M. Ceperley, Path integrals in the theory of condensed helium, *Rev. Mod. Phys.* **67**, 279 (1995).
- [17] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Quantum monte carlo simulations of solids, *Rev. Mod. Phys.* **73**, 33 (2001).
- [18] C.-W. Liu, A. Polkovnikov, and A. W. Sandvik, Quantum Versus Classical Annealing: Insights from Scaling Theory and Results for Spin Glasses on 3-Regular Graphs, *Phys. Rev. Lett.* **114**, 147203 (2015).
- [19] C.-W. Liu, A. Polkovnikov, and A. W. Sandvik, Quasi-adiabatic quantum Monte Carlo algorithm for quantum evolution in imaginary time, *Phys. Rev. B* **87**, 174302 (2013).
- [20] S. Boixo, V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. H. Amin, A. Y. Smirnov, M. Mohseni, and H. Neven, Computational multiqubit tunneling in programmable quantum annealers, *Nat. Commun.* **7**, 10327 (2016).
- [21] V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. Smelyanskiy, J. Martinis, and H. Neven, What is the Computational Value of Finite-Range Tunneling? *Phys. Rev. X* **6**, 031015 (2016).
- [22] L. T. Brady and W. van Dam, Quantum monte carlo simulations of tunneling in quantum adiabatic optimization, *Phys. Rev. A* **93**, 032304 (2016).
- [23] S. V. Isakov, G. Mazzola, V. N. Smelyanskiy, Z. Jiang, S. Boixo, H. Neven, and M. Troyer, Understanding Quantum Tunneling through Quantum Monte Carlo Simulations, *Phys. Rev. Lett.* **117**, 180402 (2016).
- [24] G. Mazzola, V. N. Smelyanskiy, and M. Troyer, Quantum monte carlo tunneling from quantum chemistry to quantum annealing, *Phys. Rev. B* **96**, 134305 (2017).
- [25] U. Weiss, H. Grabert, P. Hänggi, and P. Riseborough, Incoherent tunneling in a double well, *Phys. Rev. B* **35**, 9535 (1987).
- [26] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, Quantum computation by adiabatic evolution, [arXiv:quant-ph/0001106](https://arxiv.org/abs/quant-ph/0001106).
- [27] Z. Jiang, V. N. Smelyanskiy, S. Boixo, and H. Neven, Path-integral quantum monte carlo simulation with open-boundary conditions, *Phys. Rev. A* **96**, 042330 (2017).
- [28] E. M. Inack, G. Giudici, T. Parolini, G. Santoro, and S. Pilati, Understanding quantum tunneling using diffusion monte carlo simulations, *Phys. Rev. A* **97**, 032307 (2018).
- [29] E. Andriyash and M. H. Amin, Can quantum monte carlo simulate quantum annealing? [arXiv:1703.09277](https://arxiv.org/abs/1703.09277).
- [30] E. M. Inack and S. Pilati, Simulated quantum annealing of double-well and multiwell potentials, *Phys. Rev. E* **92**, 053304 (2015).
- [31] E. Crosson and A. W. Harrow, Simulated quantum annealing can be exponentially faster than classical simulated annealing, in *IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New Brunswick, NJ, 2016), pp. 714–723.
- [32] M. Jarret, S. P. Jordan, and B. Lackey, Adiabatic optimization versus diffusion monte carlo methods, *Phys. Rev. A* **94**, 042318 (2016).
- [33] J. Langston, How the quest for a scalable quantum computer is helping fight cancer (2019), <https://news.microsoft.com/innovation-stories/quantum-computing-mri-cancer-treatment/>.
- [34] N. Nemec, Diffusion monte carlo: Exponential scaling of computational cost for large systems, *Phys. Rev. B* **81**, 035119 (2010).
- [35] M. Boninsegni and S. Moroni, Population size bias in diffusion monte carlo, *Phys. Rev. E* **86**, 056712 (2012).
- [36] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [37] E. M. Inack, G. E. Santoro, L. Dell'Anna, and S. Pilati, Projective quantum monte carlo simulations guided by unrestricted neural network states, *Phys. Rev. B* **98**, 235145 (2018).
- [38] L. Reatto and G. L. Masserini, Shadow wave function for many-boson systems, *Phys. Rev. B* **38**, 4516 (1988).
- [39] S. Vitiello, K. Runge, and M. H. Kalos, Variational Calculations for Solid and Liquid he 4 with A" Shadow" Wave Function, *Phys. Rev. Lett.* **60**, 1970 (1988).
- [40] A. Garg, Tunnel splittings for one-dimensional potential wells revisited, *Am. J. Phys.* **68**, 430 (2000).
- [41] J. B. Anderson, A random-walk simulation of the schrödinger equation: H+ 3, *J. Chem. Phys.* **63**, 1499 (1975).
- [42] S. Bravyi, D. P. Divincenzo, R. Oliveira, and B. M. Terhal, The complexity of stoquastic local hamiltonian problems, *Quantum Inf. Comput.* **8**, 361 (2008).
- [43] S. Bravyi and B. Terhal, Complexity of stoquastic frustration-free hamiltonians, *SIAM J. Comput.* **39**, 1462 (2009).
- [44] J. Thijssen, *Computational Physics* (Cambridge University Press, Cambridge, UK, 2007).
- [45] L. Pollet, N. V. Prokof'ev, and B. V. Svistunov, Stochastic lists: Sampling multivariable functions with population methods, *Phys. Rev. B* **98**, 085102 (2018).
- [46] H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, *Physica* **7**, 284 (1940).
- [47] A. Kamenev, *Field Theory of Non-Equilibrium Systems* (Cambridge University Press, Cambridge, UK, 2011).

- [48] H. J. Müller-Kirsten, *Introduction to Quantum Mechanics: Schrödinger Equation and Path Integral* (World Scientific, Singapore, 2006).
- [49] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed. (Pergamon, New York, 1977).
- [50] E. Farhi, J. Goldstone, S. Gutmann, and D. Nagaj, How to make the quantum adiabatic algorithm fail, *Int. J. Quantum Inf.* **06**, 503 (2008).
- [51] A. P. Young, S. Knysh, and V. N. Smelyanskiy, Size Dependence of the Minimum Excitation Gap in the Quantum Adiabatic Algorithm, *Phys. Rev. Lett.* **101**, 170503 (2008).
- [52] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective, *Phys. Rep.* **523**, 127 (2013).
- [53] C. R. Laumann, R. Moessner, A. Scardicchio, and S. L. Sondhi, Quantum annealing: The fastest route to quantum computation? *Eur. Phys. J.: Spec. Top.* **224**, 75 (2015).
- [54] S. Knysh, Zero-temperature quantum annealing bottlenecks in the spin-glass phase, *Nat. Commun.* **7**, 12370 (2016).
- [55] S. Sorella and L. Capriotti, Green function monte carlo with stochastic reconfiguration: An effective remedy for the sign problem, *Phys. Rev. B* **61**, 2599 (2000).
- [56] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, UK, 2017).
- [57] M. Collura, L. Dell'Anna, T. Felser, and S. Montangero, On the descriptive power of Neural-Networks as constrained Tensor Networks with exponentially large bond dimension, [arXiv:1905.11351](https://arxiv.org/abs/1905.11351).
- [58] R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac, Restricted boltzmann machines in quantum physics, *Nat. Phys.* **15**, 887 (2019).
- [59] M. J. S. Beach, R. G. Melko, T. Grover, and T. H. Hsieh, Making Trotters Sprint: A Variational Imaginary Time Ansatz for Quantum Many-body Systems, *Phys. Rev. B* **100**, 094434 (2019).
- [60] S. Pilati, E. M. Inack, and P. Pieri, Self-learning projective quantum monte carlo simulations guided by restricted boltzmann machines, *Phys. Rev. E* **100**, 043301 (2019).
- [61] K. Choo, T. Neupert, and G. Carleo, Study of the two-dimensional frustrated J1-J2 model with neural network quantum states, *Phys. Rev. B* **100**, 125124 (2019).
- [62] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, *Nat. Mach. Intell.* **1**, 155 (2019).