Scaling analysis and instantons for thermally-assisted tunneling and Quantum Monte Carlo simulations

Zhang Jiang,[1,2] Vadim N. Smelyanskiy,[3] Sergei V. Isakov,[4] Sergio Boixo,[3] Guglielmo Mazzola,[5] Matthias Troyer,[5,6] and Hartmut Neven[3]

[1] *QuAIL, NASA Ames Research Center, Moffett Field, California 94035, USA*
[2] *SGT Inc., 7701 Greenbelt Rd., Suite 400, Greenbelt, Maryland 20770*
[3] *Google, Venice, CA 90291, USA*
[4] *Google, 8002 Zurich, Switzerland*
[5] *Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland*
[6] *Quantum Architectures and Computation Group, Microsoft Research, Redmond, Washington 98052*

We develop an instantonic calculus to derive an analytical expression for the thermally-assisted tunneling decay rate of a metastable state in a fully connected quantum spin model. The tunneling decay problem can be mapped onto the Kramers escape problem of a classical random dynamical field. This dynamical field is simulated efficiently by path integral Quantum Monte Carlo (QMC). We show analytically that the exponential scaling with the number of spins of the thermally-assisted quantum tunneling rate and the escape rate of the QMC process are identical. We relate this effect to the existence of a dominant instantonic tunneling path. The instanton trajectory is described by nonlinear dynamical mean-field theory equations for a single-site magnetization vector, which we solve exactly. Finally, we derive scaling relations for the "spiky" barrier shape when the spin tunneling and QMC rates scale polynomially with the number of spins N while a purely classical over-the-barrier activation rate scales exponentially with N.

I. INTRODUCTION

Computationally hard combinatorial optimization problems can be mapped to classical spin glass models in statistical physics [1]. The energy landscape of the corresponding spin Hamiltonians H_P possesses a large number of spurious local minima. Classical optimization strategies, such as simulated annealing (SA), exploit thermal over–the–barrier transitions for the search trajectory through the energy landscape towards low-energy spin configurations. In quantum optimization algorithms, such as quantum annealing (QA) [2–4] (see also [5] for recent results) tunneling can play a functional role by providing additional pathways to low-energy states [6].

In an archetypical example of QA in spins models, the state evolution is determined by a time-dependent Hamiltonian $H(t) = H_P - \Gamma(t)\sum_j \sigma_j^x$, where σ_j^x is a Pauli operator for the j spin, and $\Gamma(t)$ slowly interpolates between a large value and 0. At sufficiently small values of Γ, all low-energy eigenstates of $H(t)$ are localized in the vicinity of the minima of H_P [7] (pure states in the spin-glass models). In QA the energy landscape is time dependent and the energies of two minima can exchange orders. Due to the tunneling between the minima, this results in an avoided crossing with the energy gap Δ.

In the absence of an environment and for sufficiently slow evolution, QA corresponds to an adiabatic evolution where the system closely follows the instantaneous ground state of $H(t)$ [3]. The state dynamics can be described as a cascade of Landau–Zener transitions at the avoided crossings.

The interaction with an environment can suppress tunneling between two given states of the system. The rate of incoherent tunneling decay is $W \propto \Delta^2/(\hbar^2\gamma)$ when the relaxation rate γ due to environment is much larger than the energy gap, i.e., $\gamma \gg \Delta/\hbar$. However, the environment also gives rise to thermal excitations to higher-energy levels from where the system can tunnel faster [8]. This is called thermally assisted tunneling [9], and has recently been discussed in applications to flux qubit QA [10, 11].

We assume that the system with the microscopic Hamiltonian $H(t)$ has a free-energy minimum associated with a thermodynamically metastable state, and that the incoherent tunneling decay rate W of this state is much smaller than the smallest rate of relaxation γ towards the quasi-equilibrium distribution in the domain associated with this state.

Consider an eigenstate $|\psi_n\rangle$ of $H(t)$ localized in the metastable domain. The energy of $|\psi_n\rangle$ acquires an imaginary part, $E_n - i\hbar W_n/2$, due to incoherent tunneling. Thus, the partition function of the metastable state becomes a complex number $Z_0 = \sum_n e^{-\beta(E_n - i\hbar W_n/2)/\hbar}$, where $\beta/\hbar$ is the inverse temperature. For $\hbar W_n \ll E_n$ we expand the partition function in W_n and express the total tunneling decay rate W of the metastable state in the form

$$W = -\frac{2}{\beta}\frac{\mathrm{Im}(Z_0)}{\mathrm{Re}(Z_0)} = \frac{\sum_n W_n e^{-\beta E_n/\hbar}}{\sum_n e^{-\beta E_n/\hbar}}\ , \qquad (1)$$

where $\mathrm{Re}(Z_0)$ is computed by neglecting the imaginary parts of the energies. The quantity W introduced above is given by the average over the micro-canonical decay rates W_n weighted with the corresponding (quasi) equilibrium populations. Due to the entropic effects, the decay rate W can have a very steep and non-monotonic dependence on temperatures [8]. It can exceed the zero

temperature rate or even the coherent tunneling frequency. Therefore, optimal protocols of QA must also explore the finite-temperature regime.

Tunneling transitions represent the bottlenecks of QA [12] since the tunneling rate W decreases exponentially with the size of the domain D of cotunneling spins (typically, Hamming distance between the minima), $W = B_D e^{-D\,\alpha}$, where B_D is a polynomial prefactor.

In this paper, we investigate analytically thermally assisted tunneling as a *quantum* computational resource for an arbitrary temperatures. We compare the scaling exponent α in the tunneling rate $W_{\rm tunn}$ to the corresponding transition rate of path integral Quantum Monte Carlo (QMC), which is a classical algorithm that is often used to simulate QA in spin glasses [4, 5, 13, 14], and it is the most efficient algorithm to compute exponentially small gaps at first-order phase transitions [5, 15, 16]. We also provide numerical evidence to support the theoretical findings. The numerical results address the case of non-zero bias and finite (but small) temperatures, expanding the numerical findings presented in Ref. [17]

We demonstrate in a closed analytical form for mean-field quantum spin models in transverse field that the transition rate $W_{\rm QMC}$ of QMC and the quantum tunneling rate $W_{\rm tunn}$ have identical exponential scaling with number of co-tunneling spins . This finding applies to the situations where the tunneling is dominated by a most probable path (an instanton). We will employ the path-integral formalism, reminiscent of the one used in the case of continuous [9, 18, 19] and spin [20] systems, such as tunneling of magnetization in nano-magnets [21]. Remarkably, despite the big body of literature in this topic, spin tunneling has only been studied, using path integrals, in systems with fixed total spin. We will introduce a non-perturbative spin path-integral instanton calculus for systems where, due to thermal fluctuations, the total spin is not preserved. We augment our analytical results by detailed numerical studies.

Finally, we compared the quantum tunneling rate, or equivalently the classical QMC rate, with the transition rate corresponding to a classical, purely thermal activation over the barrier. We considered the case where the barrier has a form of a tall and narrow spike. We assumed that the width and height of the spike scale at most linearly with the number of spins N. We found the relations for the parameter range of the two scaling exponents where the quantum tunneling rate (and the classical QMC rate) do not scale exponentially with N while the rate of over-the-barrier thermal escape does.

In Sec. II, we review the results for a thermally assisted quantum spin tunneling rate initially obtained in [8] using the Wentzel-Kramers-Brillouin (WKB) approach. In Sec. III, we developed a Kramers theory of the escape rate for the QMC tunneling simulations. In Sec. IV, we use a path integral approach to establish a detailed connection between the WKB results for the quantum tunneling rate and the associated instanton trajectory and QMC escape rate. In Sec. V, we provide the results of numerical studies. In Sec. VI, we apply the theory developed in this paper to the tunneling problem with small and narrow barriers.

II. THERMALLY ASSISTED TUNNELING IN MULTISPIN SYSTEMS

A. Wentzel-Kramers-Brillouin (WKB) approach

The thermally assisted tunneling rate $W_{\rm tunn}$ (1) in the mean-field models where the total spin quantum number is not conserved was analyzed in Ref. [8] using the discrete WKB approach [22]. We shall summarize below the results in Ref. [8] for the archetypical model of a quantum ferromagnet in an N-spin system [8, 23]

$$\hat{H} = -2\Gamma \hat{S}_x - N g(2\hat{S}_z/N), \quad \hat{S}_\alpha = \frac{1}{2}\sum_{j=1}^{N} \sigma_\alpha^j \,, \qquad (2)$$

where Γ is the strength of the transverse field; σ_α^j is the Pauli matrix of the jth spin, where $\alpha = x, y, z$; $\hat{S}_\alpha$ is the α component of the total spin operator; and g is an arbitrary function of $2\hat{S}_z/N$.

The mean-field interaction energy density $-g(m)$ in (2) has a local and a global minimum. This class of models is known to have large free-energy barriers that lead to exponential (in N) slowing down for QA [8, 23, 24].

The squared total spin operator $\hat{S}^2 = \sum_\alpha S_\alpha^2$ commutes with the Hamiltonian $\hat{H}$ (2). The eigenstates of $\hat{H}$ can be expanded in the basis of the operators $\hat{S}_z$ and $\hat{S}^2$,

$$|\,\Psi\,\rangle = \sum_{S={\rm frac}(N/2)}^{N/2} \sum_{M=-S}^{S} \sum_{\nu=1}^{\Omega(N,S)} C_M^{S,\nu} |M, S, \nu\rangle \,, \qquad (3)$$

where ${\rm frac}(N/2)$ is the fractional part of $N/2$, and $\Omega(N,S) = \binom{N}{N/2-S} - \binom{N}{N/2-S-1}$ is the number of distinct irreducible subspaces with a given total spin quantum number S. The coefficients $C_M^{S,\nu}$ obey the stationary Schrödinger equation

$$-\Gamma \sum_{\alpha=\pm 1} \sqrt{(S+\alpha M+1)(S-\alpha M)}\, C_{M+\alpha}^{S,\nu} - N g(2M/N) C_M^{S,\nu} = E C_M^{S,\nu} \,, \qquad (4)$$

where $M = -S, -S+1, \ldots, S$.

In what follows we will study the limit $N \gg 1$ and assume that $S = \mathcal{O}(N)$. We introduce the normalized total spin quantum number $\ell = 2S/N \in [0,\,1]$. To exponential in N accuracy, the number of distinct irreducible subspaces Ω_ℓ for a given value of ℓ equals

$$\Omega_\ell \sim \exp(N Q_\ell),\; Q_\ell = \sum_{\alpha=\pm 1} \frac{1+\alpha\ell}{2} \ln\left(\frac{2}{1+\alpha\ell}\right) \,, \qquad (5)$$

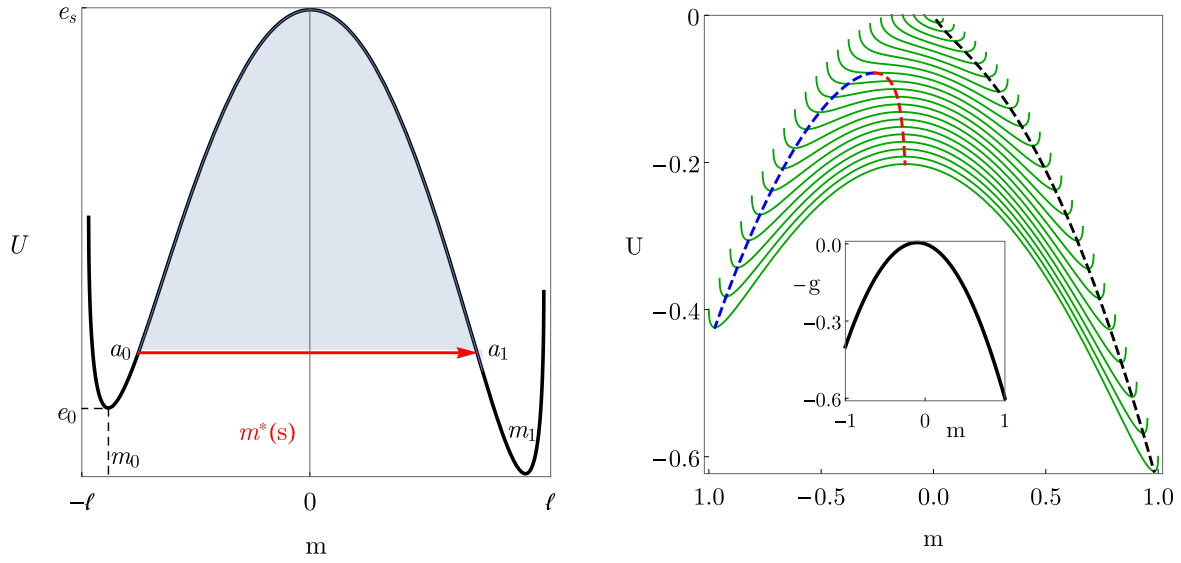


Figure 1. (Left) Plot of the effective potential $U_\ell(m)$ as a function of the magnetization m for the Curie-Weiss model (12) for $\ell = 1$, $\Gamma = 0.4$, and $h = 0.015$. The red line depicts the tunneling path (instanton) at zero temperature. Barrier energies above the instanton energy are shown with blue filling. (Right) Plots of the effective potential for the Curie-Weiss model at different values of ℓ shown with green lines with $h = 0.1$ and $\Gamma = 0.21$. The colored dashed lines correspond to the local minimum (blue), the maximum (red), and the global minimum (black) of $U_\ell(m)$ for different values of ℓ. The total spin parameter ℓ changes in the range from ℓ_c to 1 where $\ell_c = (h^{2/3} + \Gamma^{2/3})^{3/2}$ is the smallest value of ℓ beyond which the effective potential is monostable. The domain of the effective potential for each ℓ is $m \in (-\ell, \ell)$. The inset in the right figure shows the plot of interaction energy density $-g(m)$. We see by comparison of figures in left and right panels that for a given ℓ the turning point at the barrier exit satisfies the condition $a_1 \le \ell$ where the equality is reached only for $\ell = 1$.

where the binary entropic factor Q_ℓ takes the maximum value $Q_{\ell=0} = \ln 2$. To the leading order in S, Eq. (4) can be written in the form $\varepsilon_\ell(m, \hat{p})\, C_m^\ell = e C_m^\ell$ where

$$\varepsilon_\ell(m, \hat{p}) = -\Gamma\sqrt{\ell^2 - m^2}\cos\hat{p} - g(m)\,, \tag{6}$$

momentum $\hat{p} = -(2i/N)\partial/\partial m$, and the rescaled variables are

$$m = \frac{2M}{N}, \quad \ell = \frac{2S}{N}, \quad e = \frac{E}{N}, \quad C_M^{S,\nu} = C_m^\ell. \tag{7}$$

The Hamiltonian $\hat{H}$ (2) has quantized eigenvalues $E_{S,n}$. Their dependence on the quantum number n for each irreducible subspace $\ell = 1, 1 - 2/N, 1 - 4/N, \ldots$ can be obtained using the discrete WKB approach [22, 23].

The eigenfunction C_m^ℓ in the classically forbidden region of m is obtained from the analysis of an auxiliary classical one-dimensional system with energy e, coordinate m, and imaginary "momentum" $\hat{p} \to -ip$ that obeys the Hamilton-Jacobi equation [23]

$$\varepsilon_\ell(m, p) \equiv -\Gamma\sqrt{\ell^2 - m^2}\cosh p - g(m) = e\,. \tag{8}$$

In the classically-forbidden region $\cosh p \ge 1$, leading to the condition

$$-\big[e + g(m)\big] > \Gamma\sqrt{\ell^2 - m^2} > 0\,. \tag{9}$$

From Eq. (8) we obtain the momentum

$$p \equiv p_\ell(m, e) = \operatorname{arcsinh}\frac{\sqrt{\big(e + g(m)\big)^2 - \Gamma^2(\ell^2 - m^2)}}{\Gamma\sqrt{\ell^2 - m^2}}\,. \tag{10}$$

At the boundaries of the forbidden region $m = a_i$ (i=0,1) and $p_\ell(a_i, e) = 0$. These points can be obtained from the solution of the transcendental equation $U_\ell(m) = e$ where for the effective potential

$$U_\ell(m) \equiv \varepsilon_\ell(m, 0) = -\Gamma\sqrt{\ell^2 - m^2} - g(m). \tag{11}$$

Figure 1 shows plots of the effective potential $U_\ell(m)$ for the Curie-Weiss model where interaction energy density $-g(m)$ takes a form

$$-g(m) = -\frac{1}{2}m^2 - hm\,. \tag{12}$$

For values of total momentum ℓ and energy e where metastability exists, the eigenfunction under the barrier takes a standard WKB form [22]

$$C_m^\ell \propto \frac{1}{\sqrt{\partial p_\ell(m, e)/\partial e}} \exp\left(-\int_{a_0(e,\ell)}^{m} dm' p_\ell(m', e)\right)\,.$$

Then, to exponential accuracy in N, the amplitude of the tunneling decay of the metastable state equals to

$$C_{\rm tunn}(e, \ell) \propto \exp\left(-\frac{N}{2}\, S_\ell(e)\right), \tag{13}$$

where $S_\ell(e)$ is the mechanical action under the barrier

$$S_\ell(e) = \int_{a_0(e,\ell)}^{a_1(e,\ell)} p_\ell(m,e)\, dm\,, \tag{14}$$

and $m = a_i$ are boundaries of a classically-forbidden region (cf. Fig. 1).

In the WKB approach one can also compute the action $S_\ell(e)$ by introducing the instanton trajectory in imaginary time $t = -i\tau$ for a system with coordinate m, momentum ip_ℓ, and velocity iv_ℓ. Since we have introduced rescaled variables (7), we also have to rescale the time τ in order to preserve the canonical relations

$$s = 2\tau, \quad v_\ell = \frac{dm_{\rm WKB}}{ds} = -\left(\frac{\partial p_\ell}{\partial e}\right)^{-1}. \tag{15}$$

Using this relation we obtain the instanton trajectory $m = m_{\rm WKB}(s)$ in imaginary time s

$$\begin{aligned}\frac{dm_{\rm WKB}}{ds} &= v_\ell(e, m_{\rm WKB}) \qquad (16)\\ &= \sqrt{\big[e + g(m_{\rm WKB})\big]^2 - \Gamma^2(\ell^2 - m_{\rm WKB}^2)}\,.\end{aligned}$$

Then the action is $S_\ell(e) = \int_0^{s_0} p_\ell v_\ell\, ds$ where s_0 is the duration of motion under the barrier

$$s_0(e,\ell) = -\frac{\partial S_\ell}{\partial e} = \int_{a_0(e,\ell)}^{a_1(e,\ell)} \frac{dm}{v_\ell(e,m)}\,. \tag{17}$$

B. Transition rate for thermally assisted tunneling

Thermally assisted quantum tunneling is relevant to the study of QMC, which is always implemented at finite temperatures. The rate of the thermally-assisted tunneling $W_{\rm tunn}$ can be written in the form (1)

$$W_{\rm tunn} = \sum_{\ell,n} W_{\ell,n} \frac{\Omega_\ell\, e^{-\beta N e_{\ell,n}}}{Z_0} \tag{18}$$

$$Z_0 = \sum_{\ell,n} \Omega_\ell\, e^{-\beta N e_{\ell,n}}\,. \tag{19}$$

Here $W_{\ell,n} \propto |C_{\rm tunn}(e_{\ell,n},\ell)|^2$ is the tunneling decay rate of the states with the set of quantum numbers ℓ and n, and Ω_ℓ gives the number of these states. In accordance with the general prescription (1), each term in (19) is proportional to the probability of thermal activation to the states with the energies $e_{\ell,n}$ that subsequently undergoes a tunneling decay.

In the limit of large N one can replace the summations $\sum_{\ell,n}$ with $\int d\ell \int dn$ and take the integrals using the method of steepest descent. To the leading order one simply needs to maximize the logarithm of the integrand $2\ln[C_{\rm tunn}(e,\ell)] + NQ_\ell - N\beta e$ with respect to ℓ and e [here we used Eq. (5)]. In this way, one replaces the sum of the thermally assisted tunneling decays rates over many channels by the rate for the most probably channel (optimal fluctuation).

We refer to [8] for details of this analysis and simply provide a final result

$$W_{\rm tunn} = B_{\rm tunn} \exp(-N\alpha), \quad \alpha = \beta(\mathfrak{F} - \mathfrak{F}_0)\,, \tag{20}$$

where $B_{\rm tunn}$ is a prefactor and $\mathfrak{F} - \mathfrak{F}_0$ is the difference of the effective "free energies" for the thermally assisted tunneling transition. Here,

$$\beta\,\mathfrak{F} = \min_{e,\,\ell} \big(\beta e + S_\ell(e) - Q(\ell)\big)\,, \tag{21}$$

and $\mathfrak{F}_0 = -\lim_{N\to\infty}(\ln Z_0)/(N\beta)$ is a free energy per spin of the metastable state

$$\beta\,\mathfrak{F}_0 = \min_{\ell} \big(\beta e_0(\ell) - Q(\ell)\big)\,. \tag{22}$$

Here, $e_0(\ell) = U(a_0)$ is the energy of the metastable minimum. The optimal value of ℓ in (22) minimizes the system free energy over the irreducible subspaces associated with the minimum

$$\ell_0 = \operatorname{argmin}\big(\beta e_0(\ell) - Q(\ell)\big)\,. \tag{23}$$

The extremal conditions for Eq. (21) read as

$$\beta = -\frac{\partial S_\ell(e)}{\partial e}\,, \tag{24}$$

$$\frac{\partial Q_\ell}{\partial \ell} = \frac{\partial S_\ell}{\partial \ell}\,. \tag{25}$$

Using (17) and (5) we get

$$\ell = \tanh\left|\frac{\partial S_\ell(e)}{\partial \ell}\right|, \quad s_0(e,\ell) = \beta\,, \tag{26}$$

$$\left|\frac{\partial S_\ell(e)}{\partial \ell}\right| = \ell \int_0^\beta \frac{|e + g(m)|}{\ell^2 - m^2} d\tau\,. \tag{27}$$

As expected, the most probable channel for decay corresponds to the instanton $m_{\rm WKB}(s)$ of period β.

III. KRAMERS ESCAPE RATE FOR QMC TUNNELING SIMULATIONS

A. Simulating tunneling via Quantum Monte Carlo dynamics

The partition function $\mathcal{Z}$ of the model (2) at inverse temperature β can be obtained by using the Suzuki-Trotter formula to map a quantum problem to a classical one with one additional (imaginary-time) dimension $\tau \in (0,\beta)$ [25]. In the limit of infinite number of Trotter slices, $\mathcal{Z}$ is given by an integral over the array of spin paths $\underline{\sigma}(\tau) = \{\sigma_j(\tau)\}_{j=1}^N$ where each path $\sigma_j(\tau) = \pm 1$ is periodical along the imaginary-time axis $\sigma_j(0) = \sigma_j(\beta)$

and parametrized by the locations of points ("kinks") at the axis where the sign of $\sigma_j(\tau)$ changes [26]. The Gibbs probability distribution over paths $\underline{\sigma}(\tau)$ has the form

$$\mathcal{P}_G[\underline{\sigma}(\tau)] = \mathcal{Z}^{-1} \prod_{i=1}^{N} \Gamma^{\kappa[\sigma_j(\tau)]} e^{-N \int_0^\beta g[m(\tau)]d\tau}, \qquad (28)$$

where $m(\tau) = N^{-1} \sum_{j=1}^{N} \sigma_j(\tau)$ and the function $\kappa[\sigma_j(\tau)]$ equals to the number of kinks in a given path $\sigma_j(\tau)$.

QMC samples from the Gibbs distribution (28) using the Metropolis-Hastings algorithm by implementing a series of stochastic updates of the state vector of individual spin paths $\underline{\sigma}(\tau)$. At each QMC step, a transition is proposed from a current system state $\underline{\sigma}(\tau)$ to a new candidate state $\underline{\sigma}'(\tau)$ based on a certain stochastic update rule [often using cluster update methods for paths $\sigma_j(\tau)$ [26]]. The rule is specific for a given implementation of the algorithm, however, the *ratio* of the acceptance and rejection probabilities of the transition proposal $p_{\underline{\sigma}(\tau),\underline{\sigma}'(\tau)}/(1 - p_{\underline{\sigma}(\tau),\underline{\sigma}'(\tau)})$ is the same for any implementation and equal to $P_G[\underline{\sigma}'(\tau)]/P_G[\underline{\sigma}(\tau)]$. We assume that the QMC updates of $\underline{\sigma}(\tau)$ occur at the sequence of random instants of time $(t_1, t_2, \ldots)$ sampled from a Poisson process with a constant rate λ. Then, the stochastic time evolution of the state vector $\underline{\sigma}(\tau, t)$ corresponds to the master equation for the probability distribution function $\mathcal{P}_{\underline{\sigma}(\tau)}(t)$:

$$\frac{\partial P_{\underline{\sigma}(\tau)}}{\partial t} = \sum_{\underline{\sigma}'(\tau)} w_{\underline{\sigma}'(\tau),\underline{\sigma}(\tau)} \mathcal{P}_{\underline{\sigma}'(\tau)} - w_{\underline{\sigma}(\tau),\underline{\sigma}'(\tau)} \mathcal{P}_{\underline{\sigma}(\tau)} \, . \qquad (29)$$

Here, $w_{\underline{\sigma}'(\tau),\underline{\sigma}(\tau)} \propto \lambda p_{\underline{\sigma}'(\tau),\underline{\sigma}(\tau)}$ are transition probabilities that depend explicitly of the stochastic update rule. Based on the above, they must obey the detailed-balance conditions

$$\frac{w_{\underline{\sigma}(\tau),\underline{\sigma}'(\tau)}}{w_{\underline{\sigma}'(\tau),\underline{\sigma}(\tau)}} = \frac{P_G[\underline{\sigma}'(\tau)]}{P_G[\underline{\sigma}(\tau)]} \qquad (30)$$

that guarantee that Gibbs distribution (28) is the stationary (long-time) solution of (29).

Because of the mean-field character of the model (2) it is possible to obtain in a closed form a Gibbs probability measure for the magnetization per spin order parameter $m(\tau)$:

$$P_G[m(\tau)] = \frac{e^{-N\beta\mathcal{F}[m(\tau)]}}{Z_0}, \int Dm(\tau) P_G[m(\tau)] = 1 \quad (31)$$

and the free-energy functional $\mathcal{F}$ has the form [24]

$$\mathcal{F}[m(\tau)] = \frac{1}{\beta} \int_0^\beta [m(\tau) g'(m(\tau)) - g(m(\tau))] d\tau - \frac{1}{\beta} \ln \Lambda[g'(m(\tau))] \, . \qquad (32)$$

Here, $-g(m)$ is the mean-field interaction energy density (2), $g'(m) = dg/dm$, and the functional $\Lambda[\lambda(\tau)]$ equals

$$\Lambda[\lambda(\tau)] = \mathrm{Tr} K^{\beta,0}[\mathbf{B}(\tau)], \qquad (33)$$

$$K^{\tau_2,\tau_1}[\mathbf{B}(\tau)] = \mathrm{T}_+ e^{-\int_{\tau_1}^{\tau_2} d\tau H_0(\tau)} \, , \qquad (34)$$

$$H_0(\tau) = -\mathbf{B}(\tau) \cdot \boldsymbol{\sigma}, \quad \mathbf{B}(\tau) = (\Gamma, 0, \lambda(\tau)) \, , \qquad (35)$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices. The propagator $K^{\tau_2,\tau_1}[\mathbf{B}(\tau)]$ corresponds to a spin-1/2 particle evolving in imaginary time under the action of the magnetic field $\mathbf{B}(\tau)$.

If we consider the order parameter as static and remove the "time" indices, we obtain the free energy function

$$F(m) = m\, g'(m) - g(m) - \frac{1}{\beta} \ln \left(2 \cosh(\beta \sqrt{(g'(m))^2 + \Gamma^2}) \right) . \qquad (36)$$

We denote the minima of the free energy by $m = m_i$, $F_i = F(m_i)$. Here, the index $i = 0$ ($i = 1$) corresponds to the local (global) minimum of $F(m)$.

Quantum tunneling can be simulated with QMC using a general approach that was first considered in a QA context [4, 5, 13, 14]. At $t = 0$ the QMC state vector $\underline{\sigma}(\tau, 0)$ is initiated in the vicinity of the metastable state basin by uniformly sampling from the spin configurations that satisfy the conditions $N^{-1} \sum_{j=1}^{N} \sigma_j(\tau, 0) \simeq m_0$. Every time when the state vector $\underline{\sigma}(\tau, t)$ arrives at the vicinity of the global minimum $N^{-1} \sum_{j=1}^{N} \sigma_j(\tau, t_f) \simeq m_1$, the QMC process is terminated. Repeating this experiment many times, one can determine the average escape time of the QMC process from the metastable states m_0, which is proportional to the inverse of the escape rate W_{QMC} [17]. The numerical studies of W_{QMC} are described in Sec. V.

It is convenient to study the stochastic trajectories $\underline{\sigma}(\tau, t)$ by inspecting their projections $m(\tau, t)$ onto the continuous functional space defined in (32) (see Fig. 2). At low enough temperatures the distribution $P[m(\tau), t]$ corresponding to the QMC dynamics (29) quickly relaxes toward quasi-equilibrium (28) sharply localized in the vicinity of m_0. This relaxation process occurs with the rate $\gamma \gg W_{\mathrm{QMC}}$. The trajectory $m(\tau, t)$ spends a long time ($\sim W_{\mathrm{QMC}}^{-1}$) near the metastable state m_0. Occasionally, a large fluctuation occurs corresponding to the escape event where the path $m(\tau, t)$ moves away from m_0 and arrives eventually at the vicinity of m_1 (see Fig. 2). During the escape event the system initially moves uphill when the free energy $\mathcal{F}[m(\tau, t)]$ is increasing until it reaches the saddle point of the functional $\mathcal{F}$ to be denoted as $m_z(\tau)$ that satisfies the equation

$$\left. \frac{\delta \mathcal{F}[m(\tau)]}{\delta m(\tau)} \right|_{m(\tau)=m_z(\tau)} = 0, \quad m_z(0) = m_z(\beta) \, . \qquad (37)$$

Near the saddle point the escape path $m(\tau, t)$ slows down, because the variation of $\mathcal{F}[m(\tau)]$ is small. Then it goes downhill almost deterministically, so that the free energy

$\mathcal{F}[m(\tau,t)]$ is decreasing until the state m_1 is reached. The quasi-stationary statistical distribution over $m(\tau)$ has the Gibbs form $P_G[m(\tau)]$ everywhere in the domain of the local minimum except in the small vicinity of the saddle point $|\mathcal{F}[m(\tau)] - \mathcal{F}[m_z(\tau)]| \lesssim \beta^{-1}$, where deviations from $P_G[m(\tau)]$ allow for the probability current flow away from the metastable state. This area lies inside the domain marked with dashed line in Fig. 2. The probability for the path $m(\tau,t)$ to reach the vicinity of saddle point that lies inside this domain is $P_G[m(\tau)] \propto \exp\{-F[m(\tau)] - F[m_0]\}$. This is precisely the Boltzmann factor that determines the transition rate exponent in Kramers' theory of escape [27–30]

$$W_{\rm QMC} = B_{\rm QMC}\, e^{-\beta N \Delta\mathcal{F}}, \quad \Delta\mathcal{F} = \mathcal{F}[m_z(\tau)] - F(m_0)\,, \tag{38}$$

where $\Delta\mathcal{F} \gg \beta^{-1}$ and $B_{\rm QMC}$ is a prefactor (polynomial in N) that depends on the specific path update rule used in QMC.

The connection between the saddle points of the free energy functional in a classical one-dimensional field theory and the instantons in the corresponding quantum mechanical tunneling problem was first established using the path-integral formalism in [27, 31] for the case of a particle in a potential. To apply the same argument to our mean-field quantum spin problem we use the partition function $\mathcal{Z}_0$ associated with the Hamiltonian (2) in a form of the path integral normalization factor for $P_G[m(\tau)]$ (31). Then we express the tunneling decay rate in a standard form (1) in terms of the imaginary part of the partition function. For a large number of spins $N \gg 1$ this path-integral can be calculated using the saddle-point method within the instantonic calculus, where the free-energy functional $\mathcal{F}[m(\tau)]$ plays the role of the mechanical action. This gives

$$W_{\rm tunn} = B_{\rm tunn} \exp(-\beta N \Delta\mathcal{F})\,, \tag{39}$$

where $\Delta\mathcal{F}$ is determined by (37) and (38). The dominant contribution to the path integral is given by the instanton $m_z(\tau)$. We will show below that the exponential scaling of $W_{\rm tunn}$ and $W_{\rm QMC}$ with N is the same. The prefactors $B_{\rm QMC}$ and $B_{\rm turn}$ are expected to be different. $B_{\rm tunn}$ can be expressed entirely in terms of the free-energy functional $F[m(\tau)]$ while $B_{\rm QMC}$ depends on the specific path update rule used in QMC.

To complete our analysis, we would like to relate our findings to the WKB analysis of thermally assisted tunneling presented in Sec. II. We will now show in an explicit analytical form that the exponential scaling coefficient $\beta\Delta\mathcal{F}$ of the QMC transition rate with N and the QMC saddle-point solution $m_z(\tau)$ are identical to, respectively, the exponential scaling coefficient α [Eq. (20)] of the WKB tunneling rate with N and the WKB instanton tunneling path $m_{\rm WKB}(\tau)$ [Eq. (16)].

Such analysis is of interest because unlike the extrema of the free-energy functional (36) that obey the "static" condition $m(\tau) = m_{0,1}$, the saddle-point (instanton) trajectories $m_z(\tau)$ are time dependent. They also appear in the context of the QA in mean-field spin models [23, 24, 32, 33] of the type given in (2).

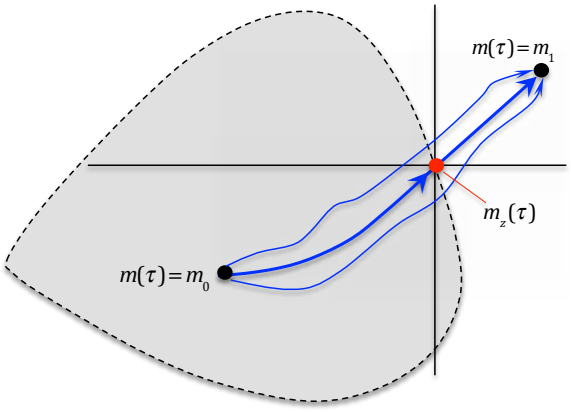


Figure 2. Schematic illustration of the stochastic Kramers escape paths $m(\tau,t)$ depicted with blue lines. The most probable escape path is shown with a solid blue line. It starts near m_0 and arrives at m_1 crossing the vicinity of saddle point $m_z(\tau)$ shown with a red dot. The dashed line depicts the boundary of the region where $\mathcal{F}[m(\tau)] \le \mathcal{F}[m_z(\tau)]$. Inside this region $\mathcal{F}[m(\tau)] - \mathcal{F}[m_z(\tau)] \gg T = \hbar/\beta$.

IV. COMPARISON OF WKB AND QMC TRANSITION RATES

Using the expression (32) for $\mathcal{F}$, we can re-write Eq. (37) in the form of the following two equations:

$$m_z(\tau) = \frac{\delta}{\delta\lambda(\tau)} \ln\Lambda[\lambda(\tau)], \quad \lambda(\tau) = g'[m_z(\tau)]\,, \tag{40}$$

where the functional $\Lambda[\lambda(\tau)]$ is defined in (33)–(35).

We now introduce a vector function in imaginary time $\mathbf{m}(\tau)$=$(m_x(\tau), m_y(\tau), m_z(\tau))$ corresponding to expectation values of the operator $\boldsymbol{\sigma}$ for the spin-$\frac{1}{2}$ particle defined as

$$\begin{aligned}\mathbf{m}(\tau) &= \frac{{\rm Tr}\left(K^{\beta,\tau}\boldsymbol{\sigma}K^{\tau,0}\right)}{{\rm Tr}\,K^{\beta,0}} \\ &= \frac{\delta}{\delta\mathbf{B}(\tau)} \ln {\rm Tr} K^{\beta,0}[\mathbf{B}(\tau)]\,,\end{aligned} \tag{41}$$

where $\mathbf{B}(\tau)$ is defined in Eq. (35). We recognize that Eq. (40) corresponds to the z-component of the vectorial equation (41).

Differentiating this equation with respect to τ and using (35), we obtain

$$\frac{d\mathbf{m}}{d\tau} = \frac{{\rm Tr}\left(K^{\beta,\tau}[H_0(\tau),\boldsymbol{\sigma}]K^{\tau,0}\right)}{{\rm Tr}\,K^{\beta,0}}\,. \tag{42}$$

One can re-write this equation in the following form

$$\frac{d\mathbf{m}}{d\tau} = -2i\frac{\partial\mathcal{H}_0(\mathbf{m})}{\partial\mathbf{m}} \times \mathbf{m}\,, \tag{43}$$

where

$$\mathcal{H}_0(\mathbf{m}) = -\Gamma m_x - g(m_z) \,. \tag{44}$$

and we observe that $\mathbf{B}(\tau) = -\partial\mathcal{H}_0[\mathbf{m}(\tau)]/\partial\mathbf{m}(\tau)$. We have to solve this equation with the periodic boundary condition $\mathbf{m}(0) = \mathbf{m}(\beta)$ [cf. (37)].

Equation (43) allows for two integrals of motion

$$\mathcal{H}_0(\mathbf{m}) = e, \tag{45}$$

$$\mathbf{m}(\tau)\cdot\mathbf{m}(\tau) = \ell^2 \tag{46}$$

where $\mathbf{m}(\tau)\cdot\mathbf{m}(\tau) \equiv m_x^2(\tau) + m_y^2(\tau) + m_z^2(\tau)$. Then, the solution of (43) can be written in the form:

$$m_x = \sqrt{\ell^2 - m_z^2}\cosh p_\ell(m,e) \,, \tag{47}$$

$$m_y = -i\sqrt{\ell^2 - m_z^2}\sinh p_\ell(m,e) \,, \tag{48}$$

$$\frac{dm_z}{d\tau} = 2v_\ell(e,m) \,, \tag{49}$$

where p_ℓ and v_ℓ are given in Eqs. (10) and (16), respectively. The equation for $m_z(\tau)$ is identical to that for the WKB instanton trajectory $m_{\rm WKB}(\tau)$ defined in (16).

Because ℓ is a constant of motion, its value can be determined at $\tau = 0$:

$$\ell = \frac{1}{\operatorname{Tr} K^{\beta,0}}\sqrt{\sum_{j=x,y,z}\big(\operatorname{Tr}(K^{\beta,0}\sigma_j)\big)^2} \,. \tag{50}$$

Since the propagator $K^{\beta,0}$ depends on ℓ, Eq. (50) is a self-consistent condition.

A. The equivalence of WKB and QMC instanton trajectories

We have shown above that the saddle point of QMC satisfies the same differential equations (49) and (16) as the WKB instanton. What remains to be shown is that the optimal value of ℓ for the WKB approach coincides with that given by the self-consistent condition of QMC (50). In the WKB approach, the optimal value of ℓ can be determined by the conditions (26) and (27). In this section, we express the self-consistent condition for QMC in the same form as the corresponding condition in WKB to demonstrate their equivalence.

It is useful to introduce a replica qubit, which allows us to write the self-consistent condition (50) as

$$\begin{aligned}\ell^2(\operatorname{Tr} K^\beta)^2 &= \operatorname{Tr}\Big(K^\beta\otimes K^\beta\sum_{j=x,y,z}\sigma_j\otimes\sigma_j\Big)\\ &= \operatorname{Tr}\Big(K^\beta\otimes K^\beta\big(P_S - 3P_A\big)\Big) \,,\end{aligned} \tag{51}$$

where $K^\beta \equiv K^{\beta,0}$ and P_S (P_A) is the projector onto the symmetric (anti-symmetric) subspace of the two qubits. To analyze the double propagator $K^\beta\otimes K^\beta$, we consider the Hamiltonian

$$H_0^{(2)} = H_0\otimes I + I\otimes H_0, \quad H_0 = -\Gamma\sigma_x - g'(m_z)\sigma_z$$

where H_0 was already introduced in Eq. (35). With the Bell basis $|\,\Phi^+\,\rangle = \frac{1}{\sqrt{2}}(|\,00\,\rangle + |\,11\,\rangle)$, $|\,\Phi^-\,\rangle = \frac{1}{\sqrt{2}}(|\,00\,\rangle - |\,11\,\rangle)$, $|\,\Psi^+\,\rangle = \frac{1}{\sqrt{2}}(|\,01\,\rangle + |\,10\,\rangle)$, $|\,\Psi^-\,\rangle = \frac{1}{\sqrt{2}}(|\,01\,\rangle - |\,10\,\rangle)$, we have

$$-H_0^{(2)}|\,\Phi^+\,\rangle = 2\Gamma|\,\Psi^+\,\rangle + 2g'(m_z)|\,\Phi^-\,\rangle \,, \tag{52}$$

$$-H_0^{(2)}|\,\Phi^-\,\rangle = 2g'(m_z)|\,\Phi^+\,\rangle \,, \tag{53}$$

$$-H_0^{(2)}|\,\Psi^+\,\rangle = 2\Gamma|\,\Phi^+\,\rangle \,, \tag{54}$$

$$-H_0^{(2)}|\,\Psi^-\,\rangle = 0 \,. \tag{55}$$

The anti-symmetric singlet state $|\,\Psi^-\,\rangle\langle\,\Psi^-\,|$ is a dark state, and the triplet states are closed under such evolution. As a consequence, the following identity always holds

$$\operatorname{Tr}\big(K^\beta\otimes K^\beta P_A\big) = \operatorname{Tr} P_A = 1 \,, \tag{56}$$

with which we have

$$\big(\operatorname{Tr} K^\beta\big)^2 - \sum_{j=x,y,z}\big(\operatorname{Tr}(K^\beta\sigma_j)\big)^2 = 4 \,. \tag{57}$$

Using the identity (56), the self-consistent condition (51) can be simplified to

$$\ell^2 = 1 - \frac{4}{\operatorname{Tr}\big(K^\beta\otimes K^\beta P_S\big) + 1} \,. \tag{58}$$

To solve ℓ, we need to know the trace of the double propagator in the symmetric subspace $\operatorname{Tr}\big(K^\beta\otimes K^\beta P_S\big)$.

Consider the time evolution of a state $|\,\Xi(\tau)\,\rangle$ in the symmetric subspace of the two qubits,

$$\begin{aligned}|\,\Xi(\tau)\,\rangle &= K(\tau,0)\otimes K(\tau,0)|\,\Xi(0)\,\rangle\\ &= -\xi_x(\tau)|\,\Phi^-\,\rangle - i\xi_y(\tau)|\,\Phi^+\,\rangle + \xi_z(\tau)|\,\Psi^+\,\rangle \,,\end{aligned} \tag{59}$$

where ξ_x, ξ_y and ξ_z take real values. Using Eqs. (52)–(54), we have the following differential equations for the coefficients:

$$\frac{d\boldsymbol{\xi}}{d\tau} = -2i\frac{\partial\mathcal{H}_0(\mathbf{m})}{\partial\mathbf{m}}\times\boldsymbol{\xi} \,, \tag{60}$$

where $\mathbf{m}$ is determined by the instanton equations (47)–(49). Equation (60) is very similar to Eq. (43), except that it is linearized. Thus, a known solution to Eq. (60) is the instanton solution $\boldsymbol{\xi}(\tau) = \mathbf{m}(\tau)$. Because the instanton solution is periodic, $\mathbf{m}(0)$ is an eigenvector of the propagator $K^\beta\otimes K^\beta$ with eigenvalue 1. To evaluate the trace, we still need to know the other two linearly independent solutions to Eq. (60).

We define the following symmetric bilinear form

$$\mathcal{B}\big(\boldsymbol{\xi}(\tau),\boldsymbol{\eta}(\tau)\big) = \xi_x(\tau)\eta_x(\tau) + \xi_y(\tau)\eta_y(\tau) + \xi_z(\tau)\eta_z(\tau) \,, \tag{61}$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are two solutions to Eq. (60). The bilinear form takes real values, and it is not an inner product (there is no complex conjugate on ξ_y). It is readily

seen that the bilinear form is a constant of motion from Eq. (60),

$$\mathcal{B}\big(\boldsymbol{\xi}(\tau),\boldsymbol{\eta}(\tau)\big)=\mathcal{B}(\boldsymbol{\xi}(0),\boldsymbol{\eta}(0))\,. \tag{62}$$

Let the other two solutions to Eq. (60) satisfy the initial conditions

$$\xi_x^{(\pm)}(0)=-m_z(0)\,,\quad \xi_z^{(\pm)}(0)=m_x(0)\,, \tag{63}$$

$$\xi_y^{(\pm)}(0)=\pm i\sqrt{m_x^2(0)+m_z^2(0)}\,. \tag{64}$$

Notice that

$$\mathcal{B}\big(\boldsymbol{\xi}^{(\pm)}(0),\,\boldsymbol{\xi}^{(\pm)}(0)\big)=\mathcal{B}\big(\boldsymbol{\xi}^{(\pm)}(0),\,\mathbf{m}(0)\big)=0\,. \tag{65}$$

Because all the bilinear forms in Eq. (65) are preserved at time β and $\mathbf{m}(\beta)=\mathbf{m}(0)$, $\boldsymbol{\xi}^{(\pm)}(\beta)$ must proportional to $\boldsymbol{\xi}^{(\pm)}(0)$ or $\boldsymbol{\xi}^{(\mp)}(0)$. The second possibility can be ruled out by noticing that the signs of $i\xi_y^{(\pm)}(\tau)$ cannot change during the evolution, because $i\xi_y^{(\pm)}(\tau)\neq 0$ for all $\tau\in[0,\,\beta]$. As a result, we have

$$\boldsymbol{\xi}^{(\pm)}(\beta)=\kappa_\pm\,\boldsymbol{\xi}^{(\pm)}(0)\,, \tag{66}$$

where proportional factors $\kappa_\pm$ are to be determined. Using Eq. (65), we have

$$-\big(\xi_y^2+\xi_z^2\big)m_x^2=-\xi_y^2m_y^2+2i\xi_z\xi_y m_z m_y+\xi_z^2m_z^2\,, \tag{67}$$

where we neglect the superscripts $(\pm)$ in ξ to simplify the notations. Introducing $\varrho_\pm=-i\xi_y^{(\pm)}/\xi_z^{(\pm)}$, we have

$$(\varrho_\pm^2-1)m_x^2=-\varrho_\pm^2m_y^2+2i\varrho_\pm m_z m_y+m_z^2\,. \tag{68}$$

The solution to the above quadratic equation is

$$\varrho_\pm=\frac{im_zm_y\pm\ell\,m_x}{\ell^2-m_z^2}\,, \tag{69}$$

where we use the condition (46) to simplify things. Putting the definition of $\varrho_\pm$ into Eq. (60), we have

$$\dot\xi_z^{(\pm)}(\tau)=2\Gamma\varrho_\pm(\tau)\xi_z^{(\pm)}(\tau)\,, \tag{70}$$

which can be solved exactly

$$\xi_z^{(\pm)}(\tau)=\xi_z^{(\pm)}(0)\,e^{2\Gamma\int_0^\tau\varrho_\pm(\tau')\,d\tau'}\,. \tag{71}$$

The integral in the above equation can be evaluated using the solution Eq. (69),

$$\Gamma\int_0^\beta\varrho_\pm(\tau)\,d\tau=\pm\ell\int_0^\beta\frac{\big|e+g(m_z)\big|}{\ell^2-m_z^2}\,d\tau\equiv\pm\mathcal{I}\,, \tag{72}$$

where we also use the relations $m_x=\big|e+g(m_z)\big|/\Gamma$ and $im_y=\dot m_z/2\Gamma$. The eigenvalues $\kappa_\pm$ can thus be determined,

$$\kappa_\pm=e^{2\Gamma\int_0^\beta\varrho_\pm(\tau)\,d\tau}=e^{\pm2\mathcal{I}}\,. \tag{73}$$

With the three eigenvalues of the double propagator all solved, we have

$$\mathrm{Tr}(K^\beta\otimes K^\beta P_S)=1+2\cosh\big(2\mathcal{I}\big)\,. \tag{74}$$

Putting the above result into Eq. (58), we have

$$\ell=\sqrt{\frac{\cosh\big(2\mathcal{I}\big)-1}{\cosh\big(2\mathcal{I}\big)+1}}=\tanh\mathcal{I}\,. \tag{75}$$

The integral $\mathcal{I}$ defined in Eq. (72) takes the same form as Eq. (27), although there is a time scale difference of 2 between the two. Comparing Eqs. (26) and (75), we conclude that the optimal value of ℓ in the WKB approach equals to the self-consistent solution of ℓ in the QMC approach.

B. The equivalence of the WKB and QMC scaling of transition rate with the number of spins

Taking into account the expression for $\mathcal{I}=\tan^{-1}\ell$ defined in Eq. (72), after some transformations the WKB action (14) for the extremal trajectory corresponding to the instanton solution (49) can be re-written in the following form:

$$S_\ell=-\beta e-\ell\,\mathcal{I}+\int_0^\beta\left[m_z(\tau)\lambda(\tau)-g(m_z(\tau))\right]d\tau\,. \tag{76}$$

where $\lambda(\tau)=g'(m_z(\tau))$ is given in (40). Using Eq. (74), we have the trace of the QMC propagator,

$$(\mathrm{Tr}\,K^\beta)^2=\mathrm{Tr}(K^\beta\otimes K^\beta P_S)+1=2+2\cosh\big(2\mathcal{I}\big) \tag{77}$$

or equivalently,

$$\mathrm{Tr}\,K^\beta=2\cosh\mathcal{I}\,. \tag{78}$$

With Eq. (75), the logarithm of $\mathrm{Tr}\,K^\beta$ can be expressed as a function of ℓ,

$$\ln\big(\mathrm{Tr}\,K^\beta\big)=\ln\frac{2}{\sqrt{1-\ell^2}}\,. \tag{79}$$

From (76) and (79) the QMC free energy (32) takes the form

$$\begin{aligned}\beta\,\mathcal{F}[m_z(\tau)]&=\int_0^\beta\big(\lambda(\tau)m_z(\tau)-g[m_z(\tau)]\big)\,d\tau-\ln\big(\mathrm{Tr}\,K^\beta\big)\\&=S_\ell+\beta e+\ell\,\tan^{-1}\ell-\ln\frac{2}{\sqrt{1-\ell^2}}\,.\end{aligned} \tag{80}$$

Using the following identity from the definition of the entropic factor Q_ℓ (5),

$$\ln\frac{2}{\sqrt{1-\ell^2}}=Q_\ell+\ell\tan^{-1}\ell\,, \tag{81}$$

we finally obtain

$$\beta\,\mathcal{F}[m_z(\tau)]=S_\ell+\beta e-Q(\ell)\,. \tag{82}$$

After the minimization over ℓ and e is performed, this has exactly the same form as the expression for the effective free-energy in the exponent of the WKB transition rate (21)

$$\min_{e,\ell} \mathcal{F}[m_z(\tau)] = \mathfrak{F}\,.$$

To complete the comparison we note that the static free energy per spin $\mathfrak{F}_0$ that appears in (20) equals to the free energy density $F(m_0)$ from (38). Indeed, those are merely two different expressions for the free energy per spin of the metastable state $m = m_0$. Connecting them requires a cumbersome but straightforward calculation given in Appendix. With that we can establish that

$$\beta\Delta\mathcal{F} = \alpha\,,$$

[cf. (20) and (38)]. Thus, we have shown the equivalence of the exponential factors in WKB and QMC transition rate expressions.

V. NUMERICAL RESULTS

In this section we compare the analytical result for the tunneling escape rate based on Eq. (20) with the QMC escape rate obtained through numerical simulations performed with continuous-time path integral quantum Monte Carlo (QMC) [26]. We study the quantum Curie-Weiss model with the Hamiltonian (2), where $g(m)$ is given in (12). The corresponding effective potential $U_\ell(m)$ for $\ell = 2S/N = 1$ is depicted in Fig. 1.

We follow the same method as in Ref. [17] to obtain the exponential scaling with N in the numerical decay rate in QMC simulations, $W_{\rm QMC} \propto \exp(-\alpha N)$. However, we extend the study [17] to the cases of nonzero biases and finite temperatures. We initialize all spin worldlines in the neighborhood of the higher (left) local minima of $U_\ell(m)$ by setting $\sigma_j(\tau) = -1$. We measure the number of QMC sweeps (defined as one attempted update per spin worldline) required to decay from the metastable state. This is done by counting the number of sweeps until at least 25% of the replicas reverse their magnetization to $\sigma_j(\tau) = 1$ [34]. We obtain the average number of sweeps required for a given number of spins N by repeating this measurement a large number of times. This was done for $N \in \{8, 10, 12, 14, 16\}$. The dependence of the number of sweeps on N is well fitted by the expression $\exp(\alpha N)/N$ (see Ref. [17]). We finally obtain the exponent α from the exponential fit with $12 \le N \le 16$. In Figs. 3(a) and 3(b) the exponent α obtained from numerical QMC simulations is compared to that obtained from the analytical approach developed in the previous sections.

For the analytical value, we solve numerically the rate for thermally assisted tunneling from the expressions (32) and (39) (their equivalence to the WKB result (20) was shown in the previous section). We use the instanton solution Eq. (16) for a particular $\ell = \ell_1$ to construct the 2-dimensional matrix $H_0(\tau)$ in Eq. (35), where $\lambda(\tau)$ is a function of the instanton solution. The 2-dimensional matrix $H_0(\tau)$ is then used to calculate the propagator Eq. (34). Knowing the propagator allows one to calculate ℓ_2 using Eq. (50). If $\ell_1 \neq \ell_2$, the original value ℓ_1 is modified until it equals to the corresponding ℓ_2; this value $\ell_1 = \ell_2$ is the self-consistent solution ℓ^*. The free energy $\mathcal{F}[m_z(\tau)]$ of the instanton trajectory with ℓ^* (the saddle point) can thus be calculated using Eq. (32). The static free energy for the metastable state $F(m_0)$ is calculated by minimizing $F(m)$ in Eq. (36) over m. It can be shown analytically that the extreme points of Eq. (36) equal to those of the effective potential (after optimizing over ℓ). The escape rate $\beta(\mathcal{F}[m_z(\tau)] - F(m_0))$ given in (32), (39) is proportional to the difference of the instanton free energy and the static free energy corresponding to the local minimum.

VI. PROBLEMS WITH SMALL AND NARROW BARRIER

It is instructive to compare the quantum tunneling rate, or equivalently the classical QMC rate, with the transition rate corresponding to a classical, purely thermal activation over the barrier. The latter is the principal mechanism for the transitions between states in simulated annealing (SA) optimization algorithms. In the present framework, that cost function is $-g(m)$, which has multiple minima separated by a barrier. As an illustrating example, we consider $-g(m)$ shown in Fig. 4,

$$-g(m) = -g_0(m) + \Delta g\, f\left(\frac{m - m_b}{\Delta m}\right), \qquad (83)$$

which is monotonically decreasing except for the small region around $m = m_b$, where there is a narrow spike of the height Δg and the typical width Δm. Problems of this type were studied previously in [35–37].

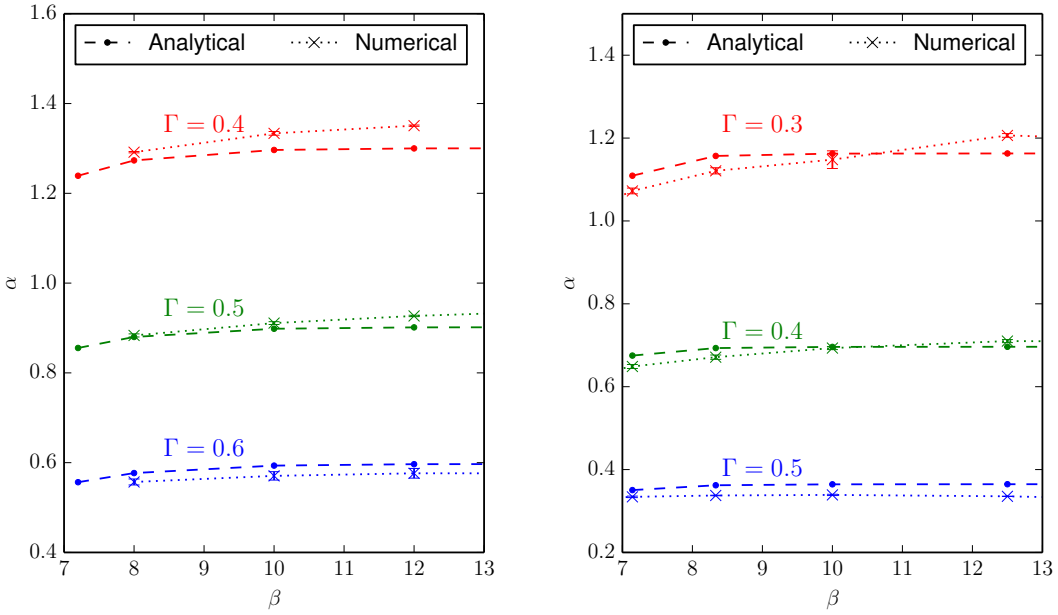
The function f describes the shape of the spike (energy barrier), $f'(0) = 0, f''(0) < 0$. We set the following relations

$$\Delta g = cN^{-\chi}, \quad \Delta m = dN^{-\delta}, \quad \chi < \delta < 1 \qquad (84)$$

where $c, d = \mathcal{O}(1)$ are constants. The main results will not depend on the specific form of the functions g_0 and f other than they are continuous and $|d^k g(m)/dm^k| = \mathcal{O}(1)$ for $k = 0, 1, 2$. We also assume that $f(q) \to 0$ exponentially quickly for $|q| \gg 1$.

The global minimum of the cost function $-g(m)$ corresponds to $m = 1$ and there is a local minimum just to the left of the spike at $m_b - m \ll 1$ as can be seen in Fig. 4. If we initially prepare the system in a state with $m < m_b$, then thermal excitations will cause an over-the-barrier transitions with a rate that scales as $e^{-N\Delta g}$. Therefore, simulated annealing solves the corresponding optimization problem with high probability in time that scales as $\exp(cN^{1-\chi})$.

During QA with the Hamiltonian (2), the transverse field $\Gamma = \Gamma(t)$ is varied in time. It starts at sufficiently



(a) Bias = 0 (b) Bias = 0.1

Figure 3. Dependence on the inverse temperature β of the the scaling exponent α (y-axis) in the analytical expression for the tunneling decay rate (20) and the scaling exponent for the QMC escape rate obtained via continuous Path Integral Quantum Monte Carlo simulations. Both results are obtained for the the Curie-Weiss model (12) for different values of transverse field Γ and bias h. The x-axis corresponds to different values for the inverse temperature β. (a) Different colors correspond to different values of the transverse field $\Gamma \in \{0.4, 0.5, 0.6\}$ for zero bias $h = 0$. (b) Different colors correspond to different values of $\Gamma \in \{0.3, 0.4, 0.5\}$ with $h = 0.1$. In both figures the $\cdot$ symbols with dashed lines correspond to the analytical values while the $\times$ symbols with dotted lines correspond to numerical values. Error bars correspond to the numerical fitting of the exponent α in the numerical data.

large initial value of $\Gamma \gg 1$ when all spins are polarized in the x direction. Then, $\Gamma(t)$ is slowly reduced to zero at the end of the algorithm. We consider the case of low temperatures where thermal fluctuations can be neglected (see below). If the evolution is adiabatic, then the system stays at the instantaneous ground state of $H(t)$ at all times during QA. This state corresponds to the irreducible subspace with the maximum total spin $\ell = 1$. At the end of QA the system arrives at the global minimum of $-g(m)$ corresponding to all spins pointing in the positive z direction.

The effective potential $U(m, \Gamma) = -\Gamma\sqrt{1-m^2} - g(m)$ [cf. (11)] is time dependent as $\Gamma(t)$ evolves during QA. The extreme points of the potential are obtained from the equation

$$\frac{\partial U}{\partial m} = \frac{m\Gamma}{\sqrt{1-m^2}} - g_0'(m) + \frac{\Delta g}{\Delta m} f'\left(\frac{m-m_b}{\Delta m}\right) = 0\,. \quad (85)$$

Following the discussion in Sec. II we denote as $m_{0,1}(\Gamma)$ the two instantaneous minima of the effective potential. The maximum of the potential $m_2(\Gamma)$ is very close to m_b for all Γ:

$$m_2(\Gamma) - m_b \simeq \frac{d^2 N^{\chi-2\delta}}{c|f''(0)|}\left(\frac{m_b\Gamma}{1-m_b^2} - g_0'(m_b)\right)\,. \quad (86)$$

At the beginning of QA when the term $\propto \Gamma$ in (85) is dominating, the system resides in the vicinity of the global minimum at $m_0 = 0$. At the end of QA the system is expected to be at the global minimum $m_1 = 1$. Because the barrier is vary narrow, we can find the value $\Gamma = \Gamma_c$ at which the minima exchange orders,

$$\Gamma_c \simeq g_0'(m_b)\sqrt{\frac{1}{m_b^2} - 1}\,. \quad (87)$$

In the vicinity of this point we have $m_{0,1}(\Gamma_c) \simeq m_b$. The value of the effective potential at the minima is $U[m_{0,1}(\Gamma_c)] \simeq -g_0(m_b)$.

To calculate the tunneling rate the action under the barrier (14) needs to be calculated at the instanton trajectory with the energy corresponding to the minima of U at $\Gamma = \Gamma_c$. The maximum value of the momentum $p(m)$ (10) is reached at the "middle" point $m = m_b$ under the barrier and to the leading order in Δg equals

$$p(m_b) = \gamma \Delta g^{1/2}, \;\; \gamma = \left(g_0'(m_b)\sinh(\ln m_b^{-1})\right)^{1/2} \quad (88)$$

Because $p(m)$ equals zero at the turning points the instantonic action can be estimated as $S = \mu p(m_b)\Delta m$

with $\mu = \mathcal{O}(1)$. The factor $\mu < 1$ depends on the shape of the barrier function f (83). For the limiting case of a rectangular barrier $\mu \to 1$. The tunneling transition rate has the form

$$W_{\rm tunn} = B_{\rm tunn}\, e^{-N\gamma\mu\Delta m\Delta g^{1/2}}\,, \tag{89}$$

where $B_{\rm tunn}$ is a prefactor that scales polynomially with N. Using (84) we get

$$W_{\rm tunn} = B_{\rm tunn}\, e^{-\kappa N^{1-\delta-\chi/2}}, \quad \kappa = c^{1/2} d\gamma\mu\,. \tag{90}$$

For the WKB analysis to be valid the tunneling action (14) must obey $S \gg 1$. This implies that one of the conditions must be true: either $\delta + \frac{\chi}{2} < 1$ or $\delta + \frac{\chi}{2} = 1$, $\kappa \gg 1$.

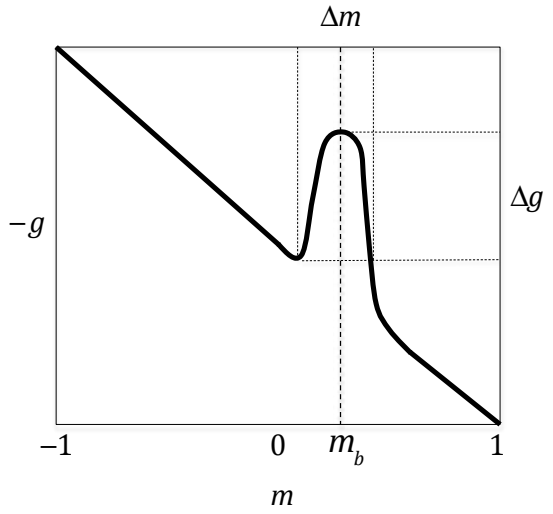


Figure 4. Cartoon of the energy function $-g(m)$ for a problem with small and narrow barrier. The barrier dimensions obey $\Delta m \ll \Delta g \ll 1$ [cf. (84].

Using the Eqs. (26) with $\ell = 1$, for thermal fluctuations to be neglected in QA the temperature must satisfy the condition

$$T \ll T_c = \frac{\hbar}{s_0(e_b)}, \quad e_b = U(\Gamma_c, m_b) \tag{91}$$

where $2s_0(e_b)$ [Eq. (17)] is the period of the instanton motion under the barrier with the energy close to the barrier top $U(m_b, \Gamma_c)$ and $\Gamma = \Gamma_c$. It can be immediately seen from the equations for the instanton velocity (16) and (84) that $s_0(e_b) \sim |g''(m_b)|^{-1/2} \sim N^{-\delta+\chi/2}$ is decreasing and therefore the temperature T_c is increasing with N. Therefore, the rate of the tunneling transition approaches the zero-temperature limit for any fixed T as N increases. Because the exponents in the quantum tunneling rate $W_{\rm turn}$ and the QMC transition rate $W_{\rm QMC}$ are identical, the latter is also well approximated by the zero-temperature limit when $T \ll T_0$.

Let us assume that the following condition is satisfied

$$\delta + \frac{\chi}{2} = 1, \quad \kappa \gg 1, \quad (\chi < \delta < 1)\,. \tag{92}$$

In this case, both the quantum tunneling and the QMC transition rates can be analyzed using the methods developed in the former sections. They scale only polynomially with N, because the exponent in Eq. (90) does not depend on N. However, the rate of the purely thermal transition scales exponentially with N. This implies that, while quantum tunneling is exponentially faster than classical SA algorithms, it still does not offer scaling advantages over classical algorithms for the situations where the path integral is dominated by a single path (instanton). We note that the above conclusion will hold also for a broader range than that given in (92):

$$1 > \chi \geq 2(1-\delta) \geq 0\,. \tag{93}$$

In this case, the QMC transition rate and quantum tunneling rate can only increase compared to the case (92) while the thermal transition rate is still exponential. An example of this situation is $\chi = 0$ and $\delta = 1$ corresponding to the barrier in the interaction energy $-Ng(m)$ that has for form of Kronecker delta with the height that scales as N. However, neither WKB nor Kramers escape theory (for QMC) do work in this case and the actual transition rate expressions should be analyzed by different methods.

VII. DISCUSSION AND SUMMARY

In this paper, we considered QMC simulations of *thermally assisted* quantum tunneling for an N-spin mean-field model. We demonstrated, in a closed analytical form, the equivalence between the exponential scaling of the QMC transition rate and the actual physical tunneling rate with the number of co-tunneling spins. This equivalence was established numerically in a previous study [17] in the effective "zero-temperature" limit where the effect of thermal excitations on the tunneling rate can be neglected, and under the condition of zero bias. In this paper, we provided a detailed theoretical description of the results of [17] and extended them to the case of thermally assisted tunneling at finite temperatures and arbitrary biases. We also provided numerical QMC study to complement theoretical results.

The findings of the identical scaling of QMC and quantum tunneling seem counterintuitive at a first glance; incoherent tunneling decay is a non-equilibrium process while QMC simulations describe fluctuations around equilibria. In mean-field spin models, the density of states increases exponentially with energy, therefore providing for a large number of tunneling channels at finite temperature (1). Both the quantum system and QMC simulation process thermalize in the metastable domain much faster than a Kramers transition or a tunneling decay occurs. However, it is not clear ahead of time that QMC explores the decay channels in the same way as the quantum system.

The rate of quantum spin tunneling can be written as a product of a polynomial (in N) prefactor and an ex-

ponential function. The exponential function dominates at the large-N limit, and it can be determined by the change in the effective free-energy functional $\mathfrak{F}[m(\tau)]$ between the values calculated at the instanton and the local minimum [see Eq. (20)]. Tunneling can thus be described by a most probable path (instanton) corresponding to a single channel that minimizes the effective energy.

The stochastic process in the QMC simulation samples the quasi-equilibrium distribution determined by the classical mean-field free-energy functional $\mathfrak{F}[m(\tau)]$ [see Eq. (32)]. The Kramers escape event in the stochastic process describes the transition from a local minimum to the global minimum, which is dominated by a single "transition state" (a saddle point of $\mathfrak{F}[m(\tau)]$) that the system needs to reach in order to make an escape from the metastable state. This transition state corresponds to a quantum instanton, and the change in free energy needed to reach this state is the same as that in the quantum case. This explains the equivalence in the exponential scaling of the QMC transition rate and thermally assisted quantum tunneling rate.

We find the solution to the instanton and the change in free energy in a closed analytical form for a general mean-field quantum spin model with the Hamiltonian $\hat{H} = -2\Gamma\hat{S}_x - Ng(2\hat{S}_z/N)$. This is achieved by establishing a detailed connection between the $\mathfrak{F}[m(\tau)]$ [Eq. (32)] and the analysis performed with WKB method that deals explicitly with the system eigenstates and takes into account the degeneracy of the collective spin states with different total spin.

We note that despite a substantial body of work on the models of this type [23, 24, 32, 33], the closed analytical form of the instanton based on the free-energy functional $\mathcal{F}[m(\tau)]$ has not been obtained previously.

We think that the spin-instanton method developed in this paper can be generalized to the case of fully connected spin-glass models in transverse field with first-order phase transition such as the p-spin models with $p > 2$. Specifically, the mathematical approach developed in Sec. IV A can be used to calculate the instanton solutions in the replicated free energy within the one-step replica symmetry ansatz [38].

We analyzed the problems with a narrow and tall barrier in the cost function. Assuming that the barrier height $N\Delta g = \mathcal{O}(N^{1-\chi})$ and width $N\Delta m = \mathcal{O}(N^{1-\delta})$ we established that under the condition $1 > \chi > 2(1-\delta) > 0$ the quantum tunneling and QMC rates scale polynomially with N while the rate of purely classical thermal activation over the barrier scales exponentially with N. An interesting case $1-\delta \ll 1$ corresponds to a very narrow barrier. In this case, the scaling exponent $1-\chi$ for the barrier height $N\Delta g$ can be very close to 1 for QMC and quantum tunneling rates to be polynomial functions of N.

Recently, Brady and van Dam [39] found numerical evidence that QMC algorithms will succeed in the same regimes where quantum adiabatic optimization succeeds. More recently, Crosson and Harrow [37] considered a bit-symmetric cost function with a thin, high-energy barrier. They proved that the Markov chain underlying QMC finds the global minimum in polynomial time (in N) if the height of the barrier scales less than order $N^{1/2}$. They also conjectured that this is true even for higher barriers that scale as N.

We note that the result obtained in Ref. [37] refers to a specific form of the interaction energy density $-g(m)$ [Eq. (2)] with a delta-function barrier and represents a particular choice of the broad class of barriers discussed in our paper where QMC and quantum tunneling scales polynomially.

In passing, we would like to briefly mention several areas open for further investigation where obstructions for the efficient simulation of quantum tunneling with QMC might exist. One of such obstructions is that QMC is not always ergodic. QMC performs dynamics of paths, which is not a representation of the dynamics of the corresponding system. A well-known problem resulting from this is that QMC might have conserved quantities not present in the physical system, such as the number of world lines (particles, magnetization), braiding, or winding numbers [40, 41].

QMC may also be less efficient compared to QA in the optimization problems that require multidimensional tunneling to reach the solution. Often in these problems the semiclassical action under the barrier $S(\mathbf{x})$ is not purely imaginary and displays complex features due to the presence of caustics, non-integrability, and non-analyticity. In this case, *no* tunneling path can be defined and a Huygens-type wave propagation should be carried out that involves both $\mathrm{Re}\, S(x)$ and $\mathrm{Im}\, S(x)$ [42, 43]. Due to the highly oscillating nature of the wave function $\Psi(x)$ in the classically forbidden region, it is not clear if the associated probability $|\Psi(x)|^2$ can be faithfully recovered with QMC.

It is an important open question as to how QMC will perform in comparison with QA in the problems that exhibit many-body location and delocalization (MBLD) transitions at finite values of transverse fields [44]. In the problems with disorder and frustration delocalized states can exist in the range of energies with exponential many local minima separated by large Hamming distances from each other. A multitude of tunneling paths connects these minima together and positive interference gives rise to extended states in the space of spin configurations above the mobility edge [45]. It is interesting to explore if the properties of the delocalized phase are important for the QA dynamics towards regions of lower energies where approximate solutions can be obtained in the vicinity of the MBLD transition. In contrast, QMC tunneling, being a classical phenomena, only connects a pair of minima at a time without reproducing the positive interference among exponentially many paths.

An interesting case where QA can have a scaling advantage over classical algorithms refers to the tunneling in non-stoquastic spin Hamiltonians where the negative sign problem prevents a matching QMC algorithm.

Finally, it was found in Ref. [17] that a version of QMC with open boundary conditions can provide a quadratic speedup compared to incoherent tunneling rate for the Hamiltonian (2) (i.e., the scaling of QMC escape rate with N matches that of quantum tunneling *amplitude* instead of the rate). This result can be obtained analytically by a direct extension of the present analysis, which is referred to future studies.

VIII. ACKNOWLEDGEMENTS

The authors would like to acknowledge support from the NASA Advanced Exploration Systems program and NASA Ames Research Center. This work was supported in part by the AFRL Information Directorate under grant F4HBKC4162G001, the Office of the Director of National Intelligence (ODNI), and the Intelligence Advanced Research Projects Activity (IARPA), via IAA 145483. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of ODNI, IARPA, AFRL, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright annotation thereon.

[1] Y. Fu and P. W. Anderson, "Application of statistical mechanics to NP-complete problems in combinatorial optimisation," J. Phys. A: Math. Gen. **19**, 1605 (1986).

[2] Tadashi Kadowaki and Hidetoshi Nishimori, "Quantum annealing in the transverse ising model," Phys. Rev. E **58**, 5355–5363 (1998); J. Brooke, D. Bitko, T. F. Rosenbaum, and G. Aeppli, "Quantum Annealing of a Disordered Magnet," Science **284**, 779–781 (1999).

[3] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda, "A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem," Science **292**, 472–475 (2001).

[4] Giuseppe E. Santoro, Roman Martonak, Erio Tosatti, and Roberto Car, "Theory of quantum annealing of an ising spin glass," Science **295**, 2427–2430 (2002).

[5] Sergio Boixo, Troels F. Ronnow, Sergei V. Isakov, Zhihui Wang, David Wecker, Daniel A. Lidar, John M. Martinis, and Matthias Troyer, "Evidence for quantum annealing with more than one hundred qubits," Nature Physics **10**, 218–224 (2014); T F Rønnow, Zhihui Wang, Joshua Job, Sergio Boixo, S V Isakov, D Wecker, John M Martinis, Daniel A Lidar, and Matthias Troyer, "Defining and detecting quantum speedup," Science **345**, 420–424 (2014); James King, Sheir Yarkoni, Mayssam M Nevisi, Jeremy P Hilton, and Catherine C McGeoch, "Benchmarking a quantum annealing processor with the time-to-target metric," arXiv:1508.05087.

[6] Sergio Boixo, Vadim N. Smelyanskiy, Alireza Shabani, Sergei V. Isakov, Mark Dykman, Vasil S. Denchev, Mohammad Amin, Anatoly Smirnov, Masoud Mohseni, and Hartmut Neven, "Computational Role of Collective Tunneling in a Quantum Annealer," arXiv:1411.4036 .

[7] Boris Altshuler, Hari Krovi, and Jérémie Roland, "Anderson localization makes adiabatic quantum optimization fail," Proceedings of the National Academy of Sciences **107**, 12446–12450 (2010).

[8] Kostyantyn Kechedzhi and Vadim N. Smelyanskiy, "Open-System Quantum Annealing in Mean-Field Models with Exponential Degeneracy," Physical Review X **6**, 021028 (2016).

[9] A. I. Larkin and Yu. N. Ovchinnikov, "Quantum tunneling with dissipation," JETP **37**, 382 (1983), [Pis'ma Zh. Eksp. Teor. Fiz. **37**, 322 (1983)]; D. A. Garanin and E. M. Chudnovsky, "Thermally activated resonant magnetization tunneling in molecular magnets: Mn_{12} Ac and others," Physical Review B **56**, 11102 (1997); Ian Affleck, "Quantum-Statistical Metastability," Phys. Rev. Lett. **46**, 388–391 (1981).

[10] Mohammad H. S. Amin and Dmitri V. Averin, "Macroscopic Resonant Tunneling in the Presence of Low Frequency Noise," Phys. Rev. Lett. **100**, 197001 (2008).

[11] N. G. Dickson, M. W. Johnson, M. H. Amin, *et al.*, "Thermally assisted quantum annealing of a 16-qubit problem," Nat Commun **4**, 1903 (2013).

[12] Sergey Knysh, "Zero-temperature quantum annealing bottlenecks in the spin-glass phase," Nature Communications **7**, 12370 (2016).

[13] Roman Martoňák, Giuseppe E. Santoro, and Erio Tosatti, "Quantum annealing by the path-integral monte carlo method: The two-dimensional random ising model," Phys. Rev. B **66**, 094203 (2002); Roman Martoňák, G. E. Santoro, and E. Tosatti, "Quantum annealing of the traveling-salesman problem," Physical Review E **70**, 057701 (2004); Demian A. Battaglia, Giuseppe E. Santoro, and Erio Tosatti, "Optimization by quantum annealing: Lessons from hard satisfiability problems," Phys. Rev. E **71**, 066707 (2005); Giuseppe E. Santoro and Erio Tosatti, "Optimization using quantum mechanics: quantum annealing through adiabatic evolution," Journal of Physics A: Mathematical and General **39**, R393–R431 (2006).

[14] Bettina Heim, Troels F. Rønnow, Sergei V. Isakov, and Matthias Troyer, "Quantum versus classical annealing of ising spin glasses," Science **348**, 215–217 (2015).

[15] AP Young, S Knysh, and VN Smelyanskiy, "First-order phase transition in the quantum adiabatic algorithm," Physical review letters **104**, 020502 (2010).

[16] Itay Hen and AP Young, "Exponential complexity of the quantum adiabatic algorithm for certain satisfiability problems," Physical Review E **84**, 061152 (2011).

[17] Sergei V. Isakov, Guglielmo Mazzola, Vadim N. Smelyanskiy, Zhang Jiang, Sergio Boixo, Hartmut Neven, and Matthias Troyer, "Understanding Quantum Tunneling through Quantum Monte Carlo Simulations," Physical Review Letters **117**, 180402 (2016).

[18] Sidney Coleman, "Fate of the false vacuum: Semiclassical theory," Phys. Rev. D **15**, 2929 (1977).

[19] A. O. Caldeira and A. J. Leggett, "Quantum tunnelling in a dissipative system," Annals of Physics **149**, 374–456 (1983).

[20] S. A. Owerre and M. B. Paranjape, "Macroscopic quantum tunneling and quantum–classical phase transitions of the escape rate in large spin systems," Physics Reports **546**, 1–60 (2015).

[21] Eugene M Chudnovsky and Javier Tejada, *Macroscopic quantum tunneling of the magnetic moment*, Vol. 4 (Cambridge University Press, 2005).

[22] Anupam Garg, "Application of the discrete Wentzel-Kramers-Brillouin method," J. Math. Phys **39**, 5166 (1998).

[23] Victor Bapst and Guilhem Semerjian, "On quantum mean-field models and their quantum annealing," Journal of Statistical Mechanics: Theory and Experiment **2012**, P06007 (2012).

[24] T Jörg, F Krzakala, J Kurchan, AC Maggs, and J Pujos, "Energy gaps in quantum first-order mean-field-like transitions: The problems that quantum annealing cannot solve," EPL (Europhysics Letters) **89**, 40004 (2010).

[25] Masuo Suzuki, "Relationship between d-dimensional quantal spin systems and $(d+1)$-dimensional ising systems—equivalence, critical exponents and systematic approximants of the partition function and spin correlations," Progress of Theoretical Physics **56**, 1454–1469 (1976).

[26] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, "Exact, complete, and universal continuous-time worldline Monte Carlo approach to the statistics of discrete quantum systems," Journal of Experimental and Theoretical Physics **87**, 310–321 (1998), [Zh. Eksp. Teor. Fiz. **114**, 570 (1998)]; Heiko Rieger and Naoki Kawashima, "Application of a continuous time cluster algorithm to the two-dimensional random quantum Ising ferromagnet," The European Physical Journal B-Condensed Matter and Complex Systems **9**, 233–236 (1999).

[27] J. S. Langer, "Statistical theory of the decay of metastable states," Annals of Physics **54**, 258–275 (1969).

[28] M. I. Dykman and M. A. Krivoglaz, "Theory of fluctuational transitions between the stable states of a non-linear oscillator," Sov. Phys. JETP **77**, 60–73 (1979), [Zh. Eksp. Teor. Fiz. **77**, 60 (1979)].

[29] Alex Kamenev, *Field theory of non-equilibrium systems* (Cambridge University Press, 2011).

[30] Peter Hänggi, Peter Talkner, and Michal Borkovec, "Reaction-rate theory: fifty years after Kramers," Reviews of Modern Physics **62**, 251 (1990).

[31] J. S. Langer, "Theory of the condensation point," Annals of Physics **41**, 108–157 (1967).

[32] A Boulatov and V. N. Smelyanskiy, "Quantum adiabatic algorithms and large spin tunnelling," Physical Review A **68**, 062321 (2003).

[33] Beatriz Seoane and Hidetoshi Nishimori, "Many-body transverse interactions in the quantum annealing of the p-spin ferromagnet," Journal of Physics A: Mathematical and Theoretical **45**, 435301 (2012).

[34] The choice of 25% is somewhat arbitrary. We find that waiting 25% or 50% of the replicas reverse magnetization results in only small differences in the fitting exponent α. This difference is slightly larger for cases when the number of sweeps for a transition in QMC is small.

[35] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, "Quantum adiabatic evolution algorithms versus simulated annealing," arXiv preprint quant-ph/0201031 (2002).

[36] Ben W. Reichardt, "The quantum adiabatic optimization algorithm and local minima," in *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing*, STOC '04 (ACM, New York, NY, USA, 2004) pp. 502–510.

[37] E. Crosson and A. W. Harrow, "Simulated Quantum Annealing Can Be Exponentially Faster Than Classical Simulated Annealing," in *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)* (2016) pp. 714–723.

[38] Thomas Jörg, Florent Krzakala, Jorge Kurchan, and AC Maggs, "Simple glass models and their quantum annealing," Physical review letters **101**, 147204 (2008).

[39] Lucas T. Brady and Wim van Dam, "Quantum Monte Carlo simulations of tunneling in quantum adiabatic optimization," Physical Review A **93**, 032304 (2016).

[40] H. G. Evertz, "The Loop Algorithm," Advances in Physics **52**, 1–66 (2003), arXiv: cond-mat/9707221.

[41] Matthew B. Hastings, "Obstructions to classically simulating the quantum adiabatic algorithm." Quantum Information & Computation **13**, 1038–1076 (2013).

[42] Z. H. Huang, T. E. Feuchtwang, P. H. Cutler, and E. Kazes, "Wentzel-Kramers-Brillouin method in multidimensional tunneling," Phys. Rev. A **41**, 32 (1990).

[43] Shoji Takada and Hiroki Nakamura, "Wentzel–Kramers–Brillouin theory of multidimensional tunneling: General theory for energy splitting," J. Chem. Phys. **100**, 98–113 (1994).

[44] Christopher R. Laumann, Roderich Moessner, Antonello Scardicchio, and S. L. Sondhi, "Quantum annealing: the fastest route to quantum computation?" Eur. Phys. J. Special Topics **224**, 75–88 (2015).

[45] Christopher R. Laumann, A. Pal, and A. Scardicchio, "Many-body mobility edge in a mean-field quantum spin glass," Phys. Rev. Lett. **113**, 200405 (2014).

Appendix

In this appendix, we show that the static free energy per spin $\mathfrak{F}_0$ that appears in (20) equals to the free energy density $F(m_0)$ from (38). For the static solution $m(\tau) = m_0$ corresponding to the local minimum of the QMC free energy (36), the propagator Eq. (34) becomes

$$K^{\beta,0} = e^{\beta[\Gamma\sigma_x + g'(m_0)\sigma_z]} \,. \tag{94}$$

Consequently, the self-consistent condition (50) becomes

$$\ell = \tanh\left(\beta\sqrt{\Gamma^2 + \left(g'(m_0)\right)^2}\right) . \tag{95}$$

The condition for extreme points of the static free energy Eq. (36) is

$$m\sqrt{\Gamma^2 + \left(g'(m)\right)^2} - g'(m)\ell = 0 \,, \tag{96}$$

where we use Eq. (95) to simplify the above expression. Solving $g'(m)$ from the above equation, we have

$$g'(m) = \Gamma m \, (\ell^2 - m^2)^{-1/2} \, . \tag{97}$$

Thus, the pair of equations (95) and (97) determine the local minimum $m = m_0$ and other extreme points of the "static" free energy (36).

In the WKB approach, the effective potential $U_\ell(m) = \varepsilon_\ell(m, 0)$ (8) takes the form

$$U_\ell(m) = -\Gamma \sqrt{\ell^2 - m^2} - g(m) \, . \tag{98}$$

The extreme points of the effective potential satisfy

$$g'(m) = \Gamma m \, (\ell^2 - m^2)^{-1/2} \, , \tag{99}$$

which is the same as Eq. (97). We will use the following equivalent form of Eq. (99),

$$\Gamma^2 + \big(g'(m)\big)^2 = \Gamma^2 l^2 \, (\ell^2 - m^2)^{-1} \, . \tag{100}$$

According to Eq. (23) the optimal values of ℓ satisfy

$$\frac{dQ_\ell}{d\ell} = \beta \, \frac{\partial U_\ell(m)}{\partial \ell} = -\beta \, \Gamma \ell \, (\ell^2 - m^2)^{-1/2} \, , \tag{101}$$

where Q_ℓ is the binary entropy. Putting Eq. (100) into Eq. (101), we have

$$\frac{dQ_\ell}{d\ell} = -\beta \sqrt{\Gamma^2 + \big(g'(m)\big)^2} \, , \tag{102}$$

which is equivalent to Eq. (95) by using the definition of the entropic factor Q_ℓ in Eq. (5). Thus, we have shown that the extreme points of the stationary solutions for QMC and WKB are the same.

Using the conditions (95) and (97), the QMC free energy (36) at the extremum takes the form

$$F_0 = \Gamma m^2 \, (\ell^2 - m^2)^{-1/2} - g(m) - \frac{1}{\beta} \ln \frac{2}{\sqrt{1 - \ell^2}} \, . \tag{103}$$

From the definition of the entropy Eq. (5), we have

$$\ln \frac{2}{\sqrt{1 - \ell^2}} = Q_\ell + \ell \tan^{-1} \ell \, . \tag{104}$$

From Eqs. (95), (96), and (97), we also have

$$\tan^{-1} \ell = \beta \Gamma \ell \, (\ell^2 - m^2)^{-1/2} \, . \tag{105}$$

Putting Eq. (105) into Eq. (104) and then putting the result into Eq. (103), we have

$$F_0 = F(m_0) = -\Gamma \sqrt{\ell^2 - m^2} - g(m) - \frac{1}{\beta} \, Q(\ell) \, , \tag{106}$$

which is exactly the static WKB free energy $\beta \mathfrak{F}_0$ given that we are using the optimal values of $m = m_0$ and ℓ as described above.