Thermal coupled cluster theory for SU(2) systems

Gaurav Harsha and Yi Xu

Department of Physics and Astronomy, Rice University, Houston, Texas 77005, USA

Thomas M. Henderson and Gustavo E. Scuseria

Department of Physics and Astronomy and Department of Chemistry, Rice University, Houston, Texas 77005, USA

Coupled cluster (CC) has established itself as a powerful theory to study correlated quantum many-body systems. Finite-temperature generalizations of CC theory have attracted considerable interest and have been shown to work as nicely as the ground-state theory. However, most of these recent developments address only fermionic or bosonic systems. The distinct structure of the *su*(2) algebra requires the development of a similar thermal CC theory for spin degrees of freedom. In this paper, we provide a formulation of our thermofield-inspired thermal CC for SU(2) systems. We apply the thermal CC to the Lipkin-Meshkov-Glick system as well as the one-dimensional transverse field Ising model as benchmark applications to highlight the accuracy of thermal CC in the study of finite-temperature phase diagrams in SU(2) systems.

I. INTRODUCTION

The *su*(2) algebra forms the basis for understanding a wide range of phenomena in condensed matter as well as chemical systems. For instance, spin systems, such as the transverse field Ising [1] model and variants of the Heisenberg model, provide one of the most effective ways to study magnetic properties in materials. Hamiltonians for some fermionic systems can also be represented using the *su*(2) generators, e.g., the Lipkin-Meshkov-Glick [2–4] and the reduced BCS models [5–7]. The computational cost required to get the exact solution to the Schrödinger equation (even for the ground state) of interacting quantum many-body systems, which includes the SU(2) models under consideration, grows exponentially with the size of the system. Significant research has been devoted to finding methods which provide sufficiently accurate solutions to the many-body Schrödinger equation at an affordable cost. Prominent examples of such methods for SU(2) systems include various flavors of quantum Monte Carlo (QMC) [8,9], density matrix renormalization group (DMRG) and tensor-product states [10,11], pseudofermion functional renormalization group (PFRG) [12], cluster-based methods [13–17], and coupled cluster (CC) theory [18,19]. Among the deterministic alternatives, DMRG provides highly accurate results for one-dimensional lattices with short-range interactions but is known to struggle with systems in higher dimensions and in the presence of long-range interactions. Contrarily, truncated CC theory in the broken-symmetry basis (e.g., in the on-site basis for spin Hamiltonians, and where symmetries may not be preserved), though not as good as DMRG in one dimension, is reasonably accurate and its performance is not affected by the dimensionality of the lattice. This has led to an extensive application of the CC theory to SU(2) systems [20–29].

The number of accurate methods available for the study of excited states and thermal properties of spin models is markedly smaller than that for the ground state. In particular, we have a handful of deterministic methods, such as ancilla DMRG [30–32] and thermal generalizations of the PFRG [33], and a slightly bigger set of stochastic methods such as finite-temperature generalizations of QMC [34–38] and minimally entangled typical thermal states (METTSs) [39,40].

The development of finite-temperature wave-function methods for electronic structure theory has attracted considerable interest in recent years. Several thermal generalizations of ground-state wave-function methods, especially the CC theory, have been introduced [41–51]. This includes our thermofield-inspired coupled cluster [50] which constructs a CC approximation to the thermofield double state, a single wave function in an enlarged Hilbert space that provides an exact representation of the thermal density matrix. Most of these thermal wave function theories, including the thermal CC mentioned above, are tailored to work with bosons or fermions in the grand-canonical ensemble. The difference in the structure of *su*(2) and fermion or boson algebras requires further work to extend these thermal wave-function theories to spin systems. In this paper, we explore a thermal wave-function formalism for SU(2) systems and construct a thermal CC theory within this framework. To assess its performance, we benchmark our method on the Lipkin-Meshkov-Glick model (referred as the Lipkin model hereafter) and the one-dimensional transverse field Ising model (TFIM).

II. THERMOFIELD DYNAMICS

Thermofield dynamics [52–55] provides a prescription for purification of the finite-temperature ensemble density matrix and constructs a single wave function, often known as the

thermal or the *thermofield double* state, that can exactly capture the thermal behavior of quantum systems. It does so by working in an enlarged space which comprises the original Hilbert space and a conjugate copy. The ensemble thermal average in the physical space becomes an expectation value over the purified thermal state in the doubled space,

$$\langle \mathcal{O} \rangle = \text{Tr}(e^{-\beta H} \mathcal{O}) = \frac{\langle \Psi(\beta)|\mathcal{O}|\Psi(\beta)\rangle}{\langle \Psi(\beta)|\Psi(\beta)\rangle}, \tag{1}$$

where β is the inverse temperature and H is the Hamiltonian. For fermions in the grand-canonical ensemble, we define the thermal state $|\Psi(\beta)\rangle$ as

$$|\Psi(\beta)\rangle = e^{-\beta H/2}|\mathbb{I}\rangle, \tag{2a}$$

$$|\Psi(0)\rangle = |\mathbb{I}\rangle = \prod_p (1 + c_p^\dagger \tilde{c}_p^\dagger)|-; -\rangle, \tag{2b}$$

where $c_p^\dagger$ ($\tilde{c}_p^\dagger$) creates a particle in the pth spin orbital in the physical (auxiliary) space, and the product in Eq. (2b) runs over all spin-orbital indices, while the state $|-; -\rangle$ denotes the vacuum for both the physical and conjugate spaces. The identity state $|\mathbb{I}\rangle$ is the exact infinite-temperature thermal state, and derives its name from the infinite-temperature density matrix, which is the identity matrix. The norm of the thermal state gives the partition function. For SU(2) spins, we adopt a similar definition for the identity state that was proposed in Refs. [56–58],

$$|\Psi(0)\rangle = |\mathbb{I}\rangle = \prod_p (1 + J_p^+ \tilde{J}_p^+)|0\rangle, \tag{3a}$$

$$|0\rangle = |\downarrow\downarrow\downarrow\cdots; \downarrow\downarrow\downarrow\cdots\rangle, \tag{3b}$$

where J_p^+ ($\tilde{J}_p^+$) is the conventional spin-1/2 ladder operator for the pth physical (auxiliary) spin and the state $|0\rangle$ describes all the physical and auxiliary spins, which are written respectively to the left and right of the semicolon in Eq. (3b), pointing downwards.

By definition, the thermal state obeys the imaginary-time evolution equation,

$$\frac{\partial}{\partial \beta}|\Psi(\beta)\rangle = -\frac{1}{2}H|\Psi(\beta)\rangle, \tag{4}$$

which can be integrated from $\beta = 0$, where $|\mathbb{I}\rangle$ is the exact thermal state, to the desired value of β. Exact integration, however, scales exponentially with the system size and approximations to this imaginary-time evolution are required. We have explored similar theory to study both the canonical and grand-canonical ensemble thermal properties of correlated fermionic systems using thermal generalizations of configuration interaction and coupled cluster theory, and refer the reader to Refs. [49–51] and references therein for further details on thermofield theory.

A. Mean-field theory

The simplest approximation to construct the thermal state is the mean-field approach, where an effective one-body Hamiltonian of the form $H_0 = \sum_p \epsilon_p J_p^z$ is used to evolve the

thermal state and results in

$$|\Phi(\beta)\rangle = e^{-\beta H_0/2}|\mathbb{I}\rangle,$$
$$= \prod_p (e^{\beta \epsilon_p/4} + e^{-\beta \epsilon_p/4} J_p^+ \tilde{J}_p^+)|0\rangle, \tag{5}$$

which we can normalize into a spin-BCS form,

$$|0(\beta)\rangle = \prod_p \left(u_p + v_p J_p^+ \tilde{J}_p^+\right)|0\rangle, \tag{6}$$

where $u_p = 1/\sqrt{1 + e^{-\beta \epsilon_p}}$ and $v_p = \sqrt{1 - u_p^2}$. The BCS parameters, u_p and v_p, can also be found by minimizing the mean-field free energy of the system. In the following discussion and results, we will not consider such a reference optimization.

B. Correlated theory

The mean-field thermal state serves as a reference point to build correlated approximations to the thermal state. Typically, a wave operator $\Omega(\beta)$ is used to build a configuration interaction (CI)- or CC-like expansion of the wave function,

$$|\Psi(\beta)\rangle \simeq \Omega(\beta) |\Phi(\beta)\rangle. \tag{7}$$

This form of the correlated thermal state is reminiscent of the interaction picture approach. In most ground-state correlated wave-function theories, it is convenient to express the wave operator Ω as excitations on the mean-field reference. For spin systems, this is achieved by transforming the problem to a new $su(2)$ basis in which the mean field is a vector product of down quasispins at each site. Then, Ω can simply be built out of the transformed J^+ ladder operators. Similarly, for thermal wave-function theories, we use a canonical transformation that rotates the operator basis in such a way that, in the new basis, the thermal mean-field state $|0(\beta)\rangle$ has the same form as $|0\rangle$ in Eq. (3b), i.e., with all the physical and auxiliary quasispins pointing downwards. At each lattice site, the 15 generators,

$$J^\mu, \tilde{J}^\nu, J^\mu \otimes \tilde{J}^\nu, \quad \forall \mu, \nu \in \{\pm, z\},$$

collectively span the $su(4)$ algebra. The thermal canonical transformation that we seek is a basis rotation in this $su(4)$ algebra. It was first proposed by Suzuki [56] and is defined as

$$S_p^\pm(\beta) = u_p J_p^\pm + 2v_p J_p^z \tilde{J}_p^\mp, \tag{8a}$$

$$\tilde{S}_p^\pm(\beta) = u_p \tilde{J}_p^\pm + 2v_p \tilde{J}_p^z J_p^\mp, \tag{8b}$$

$$S_p^z(\beta) = u_p^2 J_p^z - v_p^2 \tilde{J}_p^z - u_p v_p (J_p^+ \tilde{J}_p^+ + J_p^- \tilde{J}_p^-), \tag{8c}$$

$$\tilde{S}_p^z(\beta) = u_p^2 \tilde{J}_p^z - v_p^2 J_p^z - u_p v_p (J_p^+ \tilde{J}_p^+ + J_p^- \tilde{J}_p^-), \tag{8d}$$

where the coefficients u_p and v_p are the same as the ones discussed above. The new S^- and $\tilde{S}^-$ operators annihilate the mean-field reference, i.e.,

$$S_p^-(\beta)|0(\beta)\rangle = 0 = \tilde{S}_p^-(\beta)|0(\beta)\rangle. \tag{9}$$

The transformation in Eqs. (8) can, in fact, be derived by realizing that

$$e^{-\beta H_0/2} S_p^-(0) e^{\beta H_0/2} |0(\beta)\rangle = 0, \quad (10a)$$

$$\Rightarrow \quad S_p^-(\beta) = e^{-\beta H_0/2} S_p^-(0) e^{\beta H_0/2}, \quad (10b)$$

where $S_p^-(0) = (J_p^- + 2J_p^z \tilde{J}_p^+)/\sqrt{2}$. The inverse transformation of Eqs. (8) can be obtained by swapping J with S, and taking $v_p \to -v_p$. For the sake of brevity, in the remainder of this paper, we will use the label S for thermal and J for zero-temperature $su(2)$ operators, and drop the explicit β dependence.

One can also directly envision the basis rotation in Eqs. (8) as a nonlinear canonical transformation of the $su(2)$ algebra. In our implementation, we prefer to work with the $su(4)$ representation. We use the symbolic algebraic manipulator *drudge* [59] to encode the $su(4)$ commutation relations and perform the necessary operator algebra to obtain the expressions for the equations discussed below.

III. COUPLED CLUSTER THEORY

In coupled cluster theory, we parametrize the wave function using the exponential of an excitation operator acting on a mean-field reference state. For the thermal state, we get

$$|\Psi(\beta)\rangle = e^{T(\beta)} |\Phi(\beta)\rangle, \quad (11)$$

where $T(\beta)$ creates excitations on the thermal mean-field reference $|\Phi(\beta)\rangle$. The exponential form generally ensures that the computed properties are size extensive and size consistent provided the reference state has these properties to begin with. For all practical applications, the cluster operator T is truncated to a finite order in excitation rank; e.g., CC truncated to single and double excitation is called CCSD. Due to the nonlinear nature of the thermal transformation in Eqs. (8), the Hamiltonian becomes quartic in $su(2)$ generators, although it remains quadratic in the $su(4)$ generators. Therefore, we choose a cluster operator which is quadratic in terms of the $su(4)$ generators in order to capture the exact finite-temperature behavior in the simplest two-site systems. Although symmetries of the system under consideration can be used to simplify its structure, the most general form of the cluster operator $T(\beta)$, with single and double excitations, is

$$T(\beta) = t_0 + T_1 + T_2, \quad (12a)$$

$$T_1 = \sum_p t_p S_p^+ + \sum_p \tilde{t}_p \tilde{S}_p^+ + \sum_p \alpha_p Y_p^{++}, \quad (12b)$$

$$T_2 = \frac{1}{2} \sum_{pq} (t_{pq} S_p^+ S_q^+ + \tilde{t}_{pq} \tilde{S}_p^+ \tilde{S}_q^+ + \alpha_{pq} Y_p^{++} Y_q^{++})$$
$$+ \sum_{pq} m_{pq} S_p^+ \tilde{S}_q^+, \quad (12c)$$

where we define $Y_p^{\mu\nu} = S_p^\mu \otimes \tilde{S}_p^\nu$. The scalar parameter t_0 keeps track of the norm of the thermal CC state (i.e., the partition function). For brevity, we will write $T = \sum_\mu t_\mu \tau_\mu$, where t_μ and τ_μ are compact notations for amplitudes and operators, respectively. Substituting the CC ansatz [Eq. (11)]

into the imaginary-time evolution equation [Eq. (4)], we get

$$\left(e^{-T} \frac{\partial}{\partial \beta} e^T \right) |\Phi(\beta)\rangle = -\frac{1}{2}(\bar{H} - H_0) |\Phi(\beta)\rangle, \quad (13)$$

where the similarity transformed Hamiltonian, $\bar{H} = e^{-T} H e^T$, can be expanded using the Baker-Campbell-Hausdorff (BCH) expansion. While the BCH expansion truncates at fourth order for a general fermionic Hamiltonian, due to the nontrivial nature of the transformation in Eqs. (8), it truncates at eighth order for SU(2) Hamiltonians.

The cluster operator T is constructed from the ladder operators in a β-dependent basis. In general,

$$\left[\frac{\partial \tau_\mu}{\partial \beta}, T \right] \neq 0, \quad (14)$$

for $\tau_\mu \in \{S^\pm, \tilde{S}^\pm, S^z, \tilde{S}^z, \ldots\}$. Therefore, the similarity transformation of the β derivative in the left-hand side of Eq. (13) should be performed using the Wilcox identity [60] (see Ref. [49] for details) and gives

$$e^{-T} \frac{\partial}{\partial \beta} e^T = \sum_\mu \frac{\partial t_\mu}{\partial \beta} \tau_\mu + D, \quad (15)$$

where D represents the contributions from the derivative of the operator part of T, and is given by

$$D = \sum_\mu t_\mu \bar{\tau}_\mu, \quad (16a)$$

$$\bar{\tau}_\mu = \frac{\partial \hat{\tau}_\mu}{\partial \beta} + \frac{1}{2!}\left[\frac{\partial \hat{\tau}_\mu}{\partial \beta}, T \right] + \frac{1}{3!}\left[\left[\frac{\partial \hat{\tau}_\mu}{\partial \beta}, T \right], T \right] + \cdots. \quad (16b)$$

After these manipulations, we arrive at the imaginary-time evolution equation for the amplitudes,

$$\sum_\mu \frac{\partial t_\mu}{\partial \beta} \tau_\mu |\Phi(\beta)\rangle = \left[-\frac{1}{2}(\bar{H} - H_0) - D \right] |\Phi(\beta)\rangle, \quad (17)$$

which can be projected against various subspaces to yield the evolution equations for the CC parameters $\{t_\mu\}$,

$$\sum_\mu \langle \tau_\nu^\dagger \tau_\mu \rangle \frac{\partial t_\mu}{\partial \beta} = \left\langle \tau_\nu^\dagger \left[-\frac{1}{2}(\bar{H} - H_0) - D \right] \right\rangle, \quad (18)$$

where the expectation value is calculated over the normalized mean-field thermal state, i.e.,

$$\langle X \rangle = \langle 0(\beta)|X|0(\beta)\rangle = \frac{\langle \Phi(\beta)|X|\Phi(\beta)\rangle}{\langle \Phi(\beta)|\Phi(\beta)\rangle}. \quad (19)$$

In practice, the excitation operators are orthogonal, so $\langle \tau_\mu^\dagger \tau_\nu \rangle = \delta_{\mu\nu} \langle \tau_\nu^\dagger \tau_\nu \rangle$. The system of first-order differential equations in Eq. (18) can be integrated starting from $\beta = 0$, where the exact initial value for the cluster amplitudes is known [$t_\mu(\beta = 0) = 0$], to the desired inverse temperature. For all the results discussed in this paper, we use dopri5 [61,62], a fourth-order Runge-Kutta algorithm with adaptive grid size (available in SCIPY [63]), with a tolerance value of 10^{-5} to perform the integration.

IV. THERMAL PROPERTIES

Analogous to the ground-state methods, in approximate finite-temperature wave-function theories, we generally have

two different ways to evaluate thermal properties: Expectation values and free-energy derivatives. In this section, we will provide a brief overview of these techniques. Further details are discussed in Appendix D.

A. CC expectation value

For coupled cluster theory, in the first approach, we compute properties as asymmetric expectation values. For an observable $\mathcal{O}$, we have

$$\langle \mathcal{O} \rangle = \langle 0(\beta)|(1+Z)e^{-T}\mathcal{O}e^{T}|0(\beta)\rangle, \tag{20}$$

where we have a CC approximation for the ket, and the bra, approximated using CI, is defined as

$$\langle \Psi_{CI}(\beta)| = \langle \Phi(\beta)|(1+Z)e^{z_0}e^{-T}. \tag{21}$$

The scalar z_0 tracks the norm of the bra thermal state and the CI operator Z is given by

$$Z = Z_1 + Z_2, \tag{22a}$$

$$Z_1 = \sum_p z_p S_p^- + \sum_p \tilde{z}_p \tilde{S}_p^- + \sum_p \gamma_p Y_p^{--}, \tag{22b}$$

$$Z_2 = \frac{1}{2}\sum_{pq}(z_{pq}S_p^- S_q^- + \tilde{z}_{pq}\tilde{S}_p^- \tilde{S}_q^- + \gamma_{pq}Y_p^{--}Y_q^{--})$$
$$+ \sum_{pq} l_{pq}S_p^- \tilde{S}_q^-. \tag{22c}$$

We can reparametrize the bra thermal state as

$$\langle \Psi_{CI}(\beta)| = \langle \Phi(\beta)|(1+W)e^{w_0}, \tag{23}$$

such that the W operator has the same structure as Z, and the two are related through a disentangled similarity transformation,

$$\langle \Phi(\beta)|(1+Z)e^{z_0}e^{-T} = \langle \Phi(\beta)|(1+W)e^{w_0-t_0}. \tag{24}$$

The thermal bra obeys an imaginary-time evolution equation, similar to the ket,

$$\frac{\partial \langle \Psi_{CI}(\beta)|}{\partial \beta} = -\frac{1}{2}\langle \Psi_{CI}(\beta)|H, \tag{25}$$

from which we can derive the evolution equations for the coefficients in the W (or the Z) operator, in the same way as for the CC theory discussed in Sec. III.

B. Free energy derivatives

In the second approach, which we also call the λ-derivative approach, we use Lagrange multipliers to compute finite-temperature properties. First, we define the λ-dependent Hamiltonian as $H(\lambda) = H + \lambda \mathcal{O}$ and use it perform the imaginary-time evolution of the thermal state. Then, the property $\mathcal{O}$ can be defined as the λ derivative of the free energy F, as

$$\langle \mathcal{O} \rangle (\lambda) = \frac{\partial F}{\partial \lambda}. \tag{26}$$

Usually, we are interested in evaluating $\langle \mathcal{O} \rangle (\lambda = 0)$. For thermal CC, the partition function and the free energy F are defined as

$$\mathcal{Z} = \langle \Psi_{CI}|\Psi_{CC}\rangle = e^{t_0+z_0}\langle \Phi(\beta)|\Phi(\beta)\rangle, \tag{27a}$$

$$F = -\frac{1}{\beta}\log \mathcal{Z}. \tag{27b}$$

We can also define the free energy in terms of the integral of internal energy,

$$F(\beta) = \frac{1}{\beta}\int_0^\beta d\tau E(\tau). \tag{28}$$

While constructing the thermal state in the λ-derivative approach, we should ideally use a mean-field reference that also depends on λ, i.e., define the partition function as

$$\frac{\mathcal{Z}}{\mathcal{Z}_0} = e^{w_0+t_0}\langle 0(\beta, \lambda)|(1+W)e^T|0(\beta, \lambda)\rangle. \tag{29}$$

Comparing with *ab initio* CC theory, a λ-dependent reference is similar to orbital-optimized linear response CC theory. Properties calculated using CC expectation values and the λ-derivative approach, with or without an optimized mean field, are generally different (see Appendix D for a proof). However, as the CC approximation becomes accurate, properties from CC expectation and λ-derivative formalisms will become equivalent. We present results for both these techniques in the following section.

C. Definition of thermal state

For all the theory and results presented in this paper, we define the bra and the ket thermal states as

$$\langle \Psi(\beta)| = \langle \mathbb{I}|e^{-\beta H/2}, \quad |\Psi(\beta)\rangle = e^{-\beta H/2}|\mathbb{I}\rangle. \tag{30}$$

However, as has been discussed in Ref. [49], we can also define these thermal states as

$$\langle \Psi_\sigma(\beta)| = \langle \mathbb{I}|e^{-(1-\sigma)\beta H}, \quad |\Psi_\sigma(\beta)\rangle = e^{-\sigma\beta H}|\mathbb{I}\rangle, \tag{31}$$

where $0 \leqslant \sigma \leqslant 1$, such that the thermal expectation value of an observable $\mathcal{O}$ can be computed as

$$\langle \mathcal{O} \rangle = \langle \Psi_\sigma(\beta)|\mathcal{O}|\Psi_\sigma(\beta)\rangle \tag{32a}$$

$$= \langle \mathbb{I}|e^{-(1-\sigma)\beta H}\mathcal{O}e^{-\sigma\beta H}|\mathbb{I}\rangle. \tag{32b}$$

Note that the thermal state defined in Eq. (2a) is simply $|\Psi_{\sigma=1/2}(\beta)\rangle$. Using the fact that $\langle \mathbb{I}|X|\mathbb{I}\rangle = \text{Tr}X$, combined with the cyclic property of a trace, it is easy to show that Eq. (32) yields the correct ensemble average for the physical observable $\mathcal{O}$.

In the exact theory as well as for mean-field and CI approximations, thermal properties can be computed as symmetric ($\sigma = 1/2$) or asymmetric ($\sigma \neq 1/2$) expectation values. On the other hand, thermal CC expectation values are inherently asymmetric, even for $\sigma = 1/2$, because a simultaneous exponential parametrization of both the bra and the ket is computationally unfeasible.
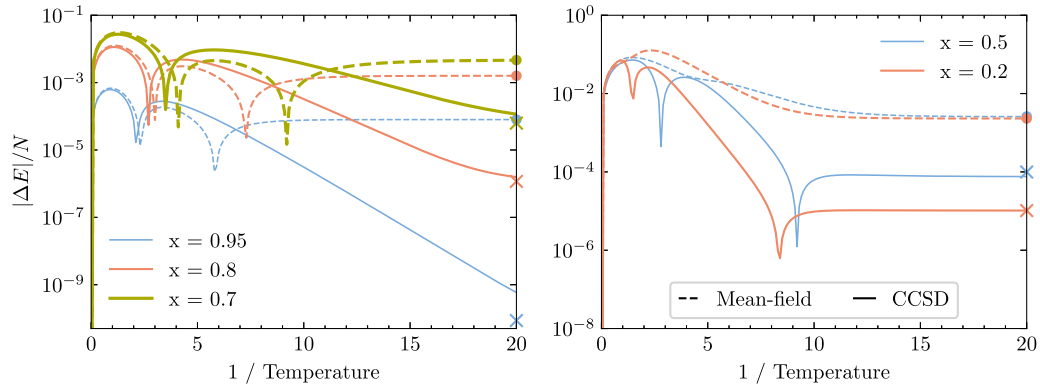
FIG. 1. Absolute error in internal energy per site as a function of the inverse temperature for the 32-site Lipkin model in both the weakly interacting regime (left), where the mean-field Hamiltonian H_0 preserves the symmetry of the ground state, and strongly interacting regime (right), where the ground-state symmetry is broken at the level of mean field. The plots compare mean-field theory and thermal CCSD against exact results. The colored "o" and "x" markers on the y axis indicate the corresponding energy error per site for ground-state RHF/UHF and CCSD, respectively.

V. RESULTS

We study the Lipkin model and the TFIM to assess the performance of our thermal CC theory. As indicated in Eqs. (12), we consider a cluster operator that is quadratic in $su(4)$ generators. For the Lipkin model, we present results for error in internal energy, while we consider both the energetics and properties for TFIM.

A. Lipkin-Meshkov-Glick model

The Lipkin-Meshkov-Glick [2–4] model describes a closed-shell nucleus with schematic monopole interactions. The system consists of N spins on a lattice in the presence of an external magnetic field in the z direction. Any two spin-up (spin-down) states can flip to spin-down (spin-up) states and lower the energy of the system. The Hamiltonian can be expressed as

$$H = xJ_z - \frac{1-x}{N}(J_+J_+ + J_-J_-), \quad (33)$$

where $J_\mu = \sum_{p=1}^{N} J_{\mu p}$ for $\mu \in \{\pm, z\}$ are the global SU(2) operators. The parameter x tunes the interaction strength. The system is noninteracting for $x = 1$ while correlation strength grows as x is reduced, and becomes extremely correlated at $x = 0$. In this paper, we only consider the parameter regime $0 \leqslant x \leqslant 1$.

The Lipkin model is exactly solvable within the Richardson-Gaudin ansatz [64,65]. Exact diagonalization (or full configuration interaction) is also straightforward as the size of the Hilbert space grows linearly with the number of spins. Despite its seemingly simple structure, the Lipkin model exhibits nontrivial physics, particularly near the transition from weakly to strongly correlated regimes; mean-field theory predicts that the parity symmetry, $P = e^{i\pi J_z}$, breaks spontaneously for $x < x_c = (2N-2)/(3N-2)$. For this reason, the Lipkin model serves as an ideal test bed for new computational methods and theories in many-body physics and chemistry (cf. Refs. [25,26]).

In an exact theory, spontaneous symmetry breaking occurs in the thermodynamic limit (TDL) but not for finite systems. We can, however, artificially break the symmetry to obtain

an energetically lower mean-field solution. This is known as the unrestricted mean-field theory. In our thermal CC implementation, we use an unrestricted mean-field Hamiltonian H_0 that allows for ground-state symmetry breaking whenever possible. For Lipkin, this happens for $x \leqslant x_c$. Appendix A contains more details on symmetry breaking and mean-field theory for the Lipkin model.

Finally, it is noteworthy that the Lipkin model can also be considered as a system of spinless fermions (see, e.g., Ref. [25]). The fermionic Lipkin model has been studied as a benchmark system in thermal cluster cumulant theory, one of the earliest thermal generalizations of CC [44]. While the spin and fermionic representations come with their own merits and demerits, for the purpose of this article the $su(2)$ version merely acts as an exactly solvable benchmark model.

Figure 1 shows the absolute error in the internal energy per site, computed at the level of mean-field and CCSD approximations [Eq. (20)], as a function of the inverse temperature and for various values of the interaction parameter x in a 32-site Lipkin model. The left (right) panel in the figure presents results for theories constructed from mean-field reference states built with a symmetry-preserving (symmetry-breaking) mean-field Hamiltonian. Thermal CC improves significantly over mean field, particularly for low temperatures. Most of the loss in accuracy occurs at intermediate temperatures and near $x = x_c$, where a single-reference description of the system is inadequate. Both the symmetry-adapted as well as symmetry-broken thermal mean-field and CCSD results approach the respective ground-state theories in the limit of zero temperature. The ground-state mean-field (CCSD) energy errors are indicated by the colored "o" and "x" markers on the y axis in the figure. The spikes observed in the logarithmic plots for the absolute error in the internal energy occur due to the accidental crossing between the exact and the approximate internal energy curves. This is an effect of the nonvariational nature of the internal energy itself combined with the nonvariational character of the CCSD approximation. Total internal energy results for the 32-site Lipkin model with $x = 0.7$ are presented in Appendix C, and shed more light on this issue. We would also like to mention that for the Lipkin model with $x = 0.5$ (in the right panel), while it may appear otherwise, thermal

CC does go to the correct ground-state limit as we evolve the system to a very large value of inverse temperature, $\beta \simeq 1000$.

Implementation details

The global SU(2) symmetry in the Lipkin model allows us to significantly reduce the structure of the cluster operator. We can drop the summation over the lattice indices and use

$$T(\beta) = t_0 + t_1 S^+ + \tilde{t}_1 \tilde{S}^+ + \alpha_1 Y^{++} + m S^+ \tilde{S}^+$$
$$+ t_2 S^+ S^+ + \tilde{t}_2 \tilde{S}^+ \tilde{S}^+ + \alpha_2 Y^{++} Y^{++}. \quad (34)$$

Here, the $su(4)$ operators represent the global operators; i.e., for any generator X,

$$X \equiv \sum_p X_p. \quad (35)$$

The operators Y^{++} and $S^+ \tilde{S}^+$ have overlapping contributions to the thermal state since

$$S^+ \tilde{S}^+ = \sum_{pq} S_p^+ \tilde{S}_q^+ \quad \text{and} \quad Y^{++} = \sum_p S_p^+ \tilde{S}_p^+. \quad (36)$$

To avoid making the imaginary-time evolution complicated, we choose either Y^{++} or $S^+ \tilde{S}^+$ (but not both) in our cluster operator. We find that these choices lead to very similar results. Therefore, for all the work presented here, we use Y^{++} but not the $S^+ \tilde{S}^+$ term in the cluster operator.

The number of parameters in the cluster operator is independent of the system size and depends only on the order of approximation. Hence, the computational scaling of thermal CC in Lipkin depends only on the number of grid points over which we perform the integration. While this implies that we can study large Lipkin models without any added computational cost, we consider a model with merely 32 spins for the ease of generating exact finite-temperature results.

B. Transverse field Ising model

The one-dimensional (1D) TFIM is a canonical model to study quantum criticality and phase transitions. The reason behind its popularity is that the model exhibits a quantum phase transition between ordered and disordered phases while also being tractable both analytically and numerically. Consequently, it serves as an ideal model to benchmark new computational theories like our thermal CC. The Hamiltonian is given by

$$H = -4 \sum_i J_i^z J_{i+1}^z + 2g \sum_i J_i^x, \quad (37)$$

where g (chosen to be positive) quantifies the strength of the transverse magnetic field. In our work, we only consider 1D chains with periodic boundary conditions. In the absence of the transverse field, we have a ferromagnetic Ising model that breaks the $\mathbb{Z}_2$ symmetry in the TDL. On the other hand, for large g, the model has a disordered paramagnetic ground state. The one-dimensional chains exhibit a quantum phase transition from a ferromagnetically ordered phase to a disordered phase at $g = 1$. It is well known that the mean-field theory overestimates the magnetic order and predicts a transition at $g = 2$ instead. For our thermal CC theory, once again, we artificially break the $\mathbb{Z}_2$ symmetry to obtain an energetically
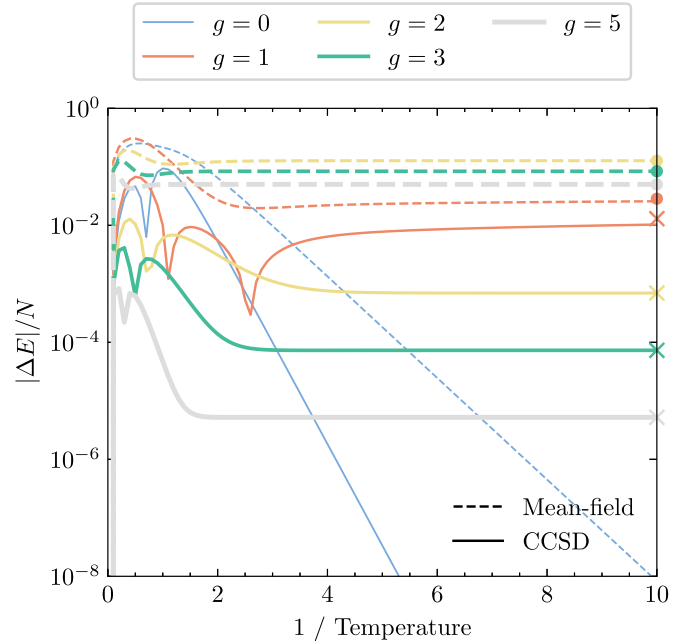

FIG. 2. Absolute error in internal energy per site for ten-site transverse field Ising models at various values of the transverse field g. The colored "o" and "x" markers on the y axis indicate the corresponding errors for the ground-state mean-field and CCSD energies, respectively.

lower mean-field solution when possible. Appendix B contains further details about the mean-field theory and the choice of H_0 for thermal mean field.

Figure 2 plots absolute error in internal energy per site for thermal mean-field and CCSD as a function of inverse temperature for a ten-site TFIM at various values of the transverse field. Here, we have used Eq. (20) to compute the thermal CCSD internal energy. As for Lipkin, thermal CC significantly improves over mean field. We also observe similar spikes in the logarithmic error plots. To corroborate this nonvariational nature of the results, we plot the total mean-field, CCSD, and exact internal energies for the 1D TFIM at $g = 1$ in Appendix C. We should note that for $g = 0$ and $g \gg 1$, both the mean-field theory and thermal CCSD are exact for the ground state. However, even when ground-state CC or mean-field theory are exact, the finite-temperature theories may not be so. As we can observe from the $g = 0$ results in Fig. 2, both thermal mean field and CC give nonzero errors for the internal energy, which decrease exponentially as we evolve towards zero temperature.

For the TFIM, we also compute properties, namely, magnetization density and spin-spin correlation functions. We first consider the magnetization density, which we calculate using both the CC expectation value and λ derivative of free energy (as already noted in Sec. IV). To make a sensible comparison, we first introduce an external field f in the z direction; i.e., we redefine the Hamiltonian as

$$H = -4I \sum_i J_i^z J_{i+1}^z + 2g \sum_i J_i^x + 2f \sum_i J_i^z. \quad (38)$$

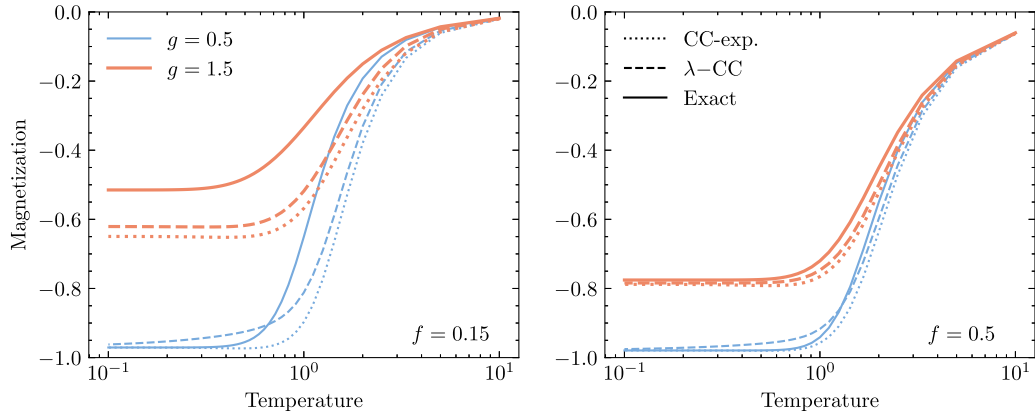We have also introduced the Ising coupling constant I for bookkeeping. The magnetization density can then be

FIG. 3. Magnetization curves for ten-site transverse field Ising models for transverse fields $g = 0.5$ and 1.5, indicated by red and blue, respectively, and magnetizing fields $f = 0.15$ (left panel) and $f = 0.5$ (right panel). The plots compare magnetization calculated using CC expectation value [Eq. (20)] and λ-derivative CC [Eq. (26)] against exact magnetization, which was calculated as the f derivative of the exact free energy. Different colors indicate results for different g values, while the line styles distinguish between various approximations.

calculated as

$$M_z = \lim_{f \to 0^+} \frac{1}{N} \frac{\partial F}{\partial f}. \qquad (39)$$

While this definition holds well in the TDL, there are some caveats when working with finite N. Consider the ferromagnetic regime near $g = 0$ and $f > 0$ so that the ground state consists of all spins pointing downwards. Starting from this ground state, we can have two different excitations: single spin flip, for which the excitation energy is $\sim 2I$, and all spin flips, i.e., the ferromagnetic state with all spins pointing up. For the latter, the excitation energy is $\sim 2fN$. For a conventional ferromagnetic phase, single spin flips constitute the low-energy excitations. Therefore, the limit for f in Eq. (39) should be carefully defined, and we should have $2fN > 2I$, or $f > I/N$. Using $I = 1$, magnetization at $g = 0$ should be redefined as

$$M_z = \lim_{f \to (1/N)^+} \frac{1}{N} \frac{\partial F}{\partial f}. \qquad (40)$$

As we go away from $g = 0$, the correct limiting value of the external field for the ferromagnetic phase, f_c, will change.

To avoid such discrepancies while comparing our benchmark calculations and exact results, we only consider magnetizing fields that are sufficiently large. Figure 3 shows the temperature dependence of magnetization as we compare the results for CC expectation value [Eq. (20)] and λ-derivative [Eq. (26)] approaches against the exact results which were obtained by taking the λ derivative of the exact free energy. In both approximate results, we use a λ-dependent reference. Since magnetization is a one-body operator, we define the field-dependent thermal reference as

$$|\Phi(\beta, \lambda)\rangle = e^{-\beta(H_0 + 2f \sum_i J_i^z)/2} |\mathbb{I}\rangle. \qquad (41)$$

While CC expectation values use a fixed value of the external field (at which the properties are calculated), the λ-derivative approach considers f dependence of the mean field at every grid point before the derivative is evaluated. Both the CC expectation value and λ derivative perform reasonably well in the weakly correlated regime ($g = 0.5$) for all the values of external field. Near the phase transition region ($g = 1.5$), the overall quality of the CC wave function drops. However,
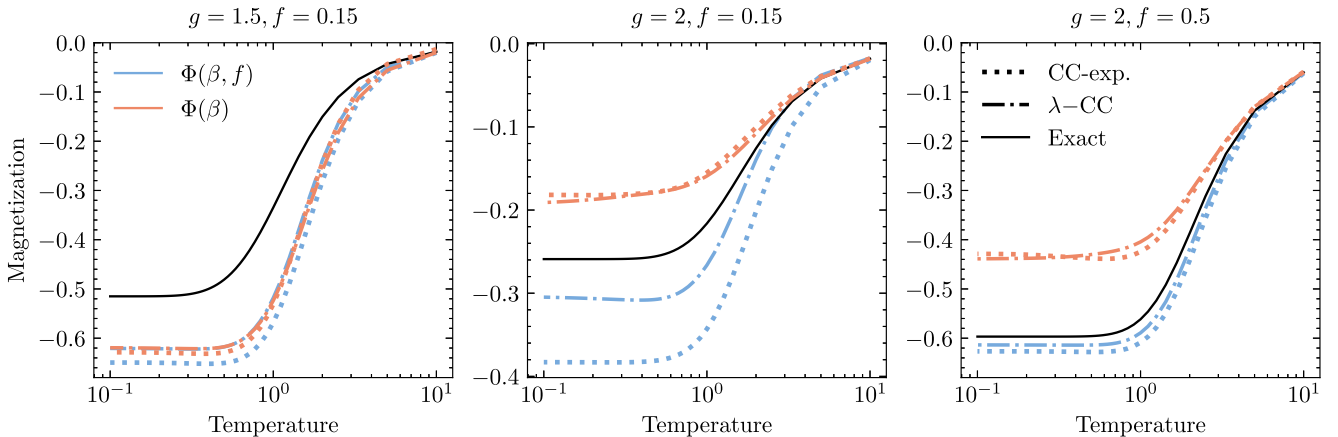

FIG. 4. Magnetization densities for ten-site TFIMs with $g = 1.5$, $f = 0.15$ on the left, $g = 2$, $f = 0.15$ in the center, and $g = 2$, $f = 0.5$ on the right, respectively. We compare CC expectation value and λ-derivative methods, with and without the external field or f dependence in the mean-field thermal state, against the exact results.
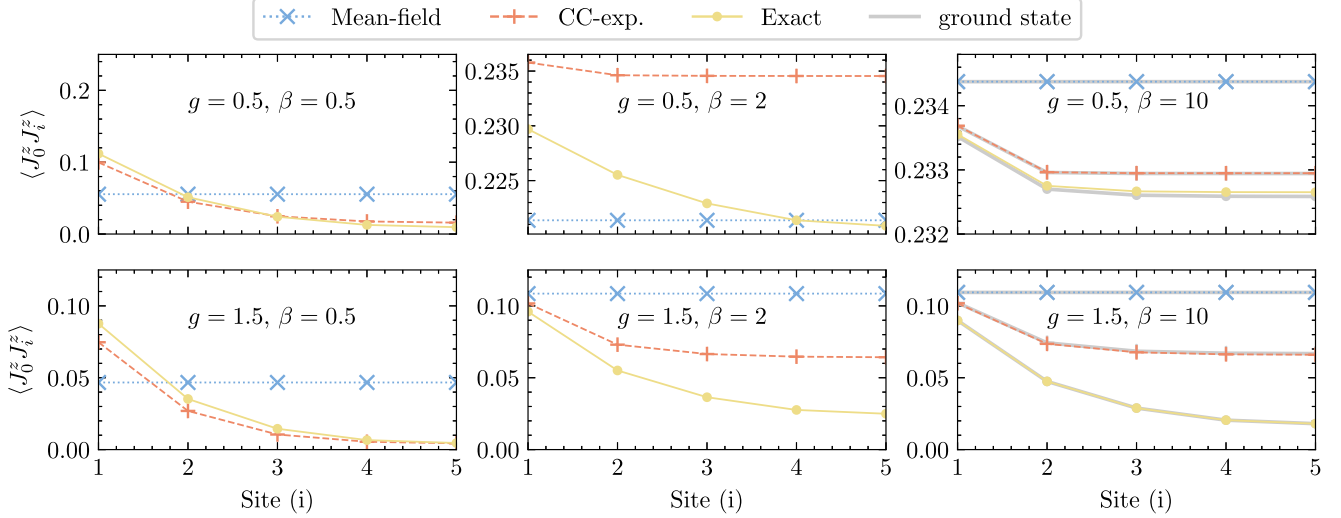
FIG. 5. Spin-spin correlation plots for ten-site transverse field Ising models for $g = 0.5$ (top row) and $g = 1.5$ (bottom row) and $f = 0$. We compare the mean-field and CC expectation value [Eq. (20)] estimates against exact results, which were calculated as ensemble averages, at various values of β. Results for the corresponding ground-state theories are also plotted in grey in the third column ($\beta = 10$) and demonstrate convergence of thermal theories as $\beta \to \infty$. For the ten-site model, with periodic boundary conditions, the correlation function $\langle J_0^z J_i^z \rangle$ is symmetric about $i = 5$. Therefore we plot data only for $i = 1$–5.

the λ-derivative results are consistent with the CC expectation values.

In order to illustrate the importance of an f-dependent mean field, we present magnetization density results for a ten-site TFIM at $(g, f) = (1.5, 0.15)$, $(g, f) = (2, 0.15)$, and $(g, f) = (2, 0.5)$. Figure 4 compares CC expectation value and λ-derivative estimates to the magnetization density, computed with and without an external field or f dependence in the mean-field reference, against exact results. For small external field, as in the case for $(g, f) = (1.5, 0.15)$, expectation values with f-dependent and f-independent mean-field references are similar, with the latter being marginally better. This is because mean-field theory in 1D TFIM already overestimates the ferromagnetic order. Including the effect of an external field only increases the strength of the magnetization. On the other hand, as the external field strength is increased, including f dependence in the mean-field theory becomes crucial. This is corroborated by the results for $(g, f) = (2.0, 0.15)$ and $(g, f) = (2.0, 0.5)$.

In Fig. 5, we also compare mean-field and CC expectation value results for the spin-spin correlation function $\langle J_0^z J_i^z \rangle$ against exact ensemble averages for this ten-site model. Once again, the mean-field approach performs poorly and gives a flat, featureless correlation function. On the other hand, thermal CCSD adds significant corrections, quantitative and qualitative, for both $g = 0.5$ and $g = 1.5$, particularly at low and high temperatures. Near the thermal phase transition region, i.e., at $\beta = 2$, where thermal CCSD is the least accurate, the correlation function, despite exhibiting the right qualitative structure, is not quantitatively accurate (see second column of Fig. 5). For large β, i.e., as we approach zero temperature, the correlation curves converge to the corresponding ground-state properties, as we can deduce from the grey-colored ground-state curves in the third panels of each row in Fig. 5. We also observe one of the side effects of the asymmetric expectation values in CC theory. For $g = 0.5$,

going from $\beta = 2$ to $\beta = 10$, we find that the strength of the thermal CCSD correlation function decreases while the exact correlation increases. While the decrease in correlation is negligible in this case ($\sim 10^{-3}$), the erratic behavior may become severe when CC is not a good approximation to the thermal state. Higher-order approximation to the bra state can be used to address such problems. Finally, we note that $g = 1.5$ results do not exhibit similar problems.

Implementation details

We can exploit the symmetries of the system to simplify the structure of the cluster operator for TFIM, just as we did for the Lipkin model. For the periodic chains under consideration, all sites are equivalent. Therefore, we can express the cluster operator as

$$T(\beta) = t_0 + s \sum_p S_p^+ + \tilde{s} \sum_p \tilde{S}_p^+ + \alpha_0 \sum_p Y_p^{++}$$

$$+ \frac{1}{2} \sum_{pq} (d_{pq} S_p^+ S_q^+ + \tilde{d}_{pq} \tilde{S}_p^+ \tilde{S}_q^+ + m_{pq} Y_p^{++} Y_q^{++})$$

$$+ \sum_{pq} x_{pq} S_p^+ \tilde{S}_q^+, \tag{42}$$

where we enforce that the tensors d, $\tilde{d}$, m, and x are symmetric and have zero diagonals. With these simplifications, we can perform the imaginary-time evolution with a computational scaling of $\mathcal{O}(N^4 N_{grid})$. This scaling can be further brought down by realizing that the tensors d_{pq}, $\tilde{d}_{pq}$, m_{pq}, and x_{pq} depend only on $|p - q|$.

VI. CONCLUSION

We have demonstrated a generalization of coupled cluster theory to finite-temperature SU(2) systems with the help of thermofield dynamics. The CCSD framework, which has been

applied successfully to study ground-state properties of spin systems, performs equally well at finite temperatures. At the same time, thermal CC experiences the same problems as its ground-state counterpart: it is insufficient for strongly correlated systems, and the expectation values are asymmetric. The single-reference nature of the CC ansatz fails to capture the underlying multiconfigurational physics in strongly correlated systems. This, in conjunction with the fact that CC is a nonvariational theory, leads to poor performance, both for the internal energy and properties. However, a key advantage of coupled cluster is that it is systematically improvable. Including higher-order excitations (triples, quadruples, etc.) in both the cluster operator and the configuration interaction approximation to the bra state can help alleviate some of the associated issues. In fact, for spin lattices, where CCSD scales as $\mathcal{O}(N^3)$ or $\mathcal{O}(N^4)$, higher-order excitations can be added without making the theory computationally intractable. Another alternative is to use multiconfigurational wave functions (such as nonorthogonal CI) or unconventional mean-field states (such as spin antisymmetrized geminal power state, dimer mean field, resonating valence bond, etc.) as reference states to build the thermal CC ansatz. Such strategies have been explored in the context of ground-state CC theory and have shown promise. Their thermal generalizations provide pathways to expand upon our current work.

APPENDIX A: SYMMETRY BREAKING IN LIPKIN MODEL

The Lipkin-Meshkov-Glick Hamiltonian, given by

$$H = xJ_z - \frac{1-x}{n}(J_+J_+ + J_-J_-), \quad (A1)$$

commutes with the Casimir operator,

$$J^2 = \frac{J_+J_- + J_-J_+}{2} + J_zJ_z. \quad (A2)$$

The J^2 eigenvalue is determined by the total number of spins; for n spins, the eigenstates of H reside in the $j = n/2$ sector. The Hamiltonian also commutes with the parity operator, $\hat{P} = e^{i\pi J_z}$, the symmetry of interest. For a given spin configuration, the parity eigenvalue is 1 (or -1) if the difference in the number of up- and down-spins is even (or odd). Therefore, simultaneous eigenstates of H and $\hat{P}$ are made out of spin configurations, all of which have only even or only odd parity. As a result, the symmetry-preserving mean-field [or the restricted Hartree-Fock (RHF)] state that optimizes the energy is simply

the configuration with all down-spins, i.e.,

$$|\phi_{\mathrm{RHF}}\rangle = \bigotimes_i |\downarrow\rangle_i. \quad (A3)$$

The corresponding mean-field Hamiltonian and the mean-field ground-state energy are

$$H_0^{\mathrm{RHF}} = xJ_z, \quad (A4a)$$

$$E_{\mathrm{RHF}} = -\frac{nx}{2}. \quad (A4b)$$

While working in the symmetry-adapted basis, we use H_0^{RHF} to define the mean-field thermal state.

We can also consider a rotated product state, that is not an eigenvector of the parity operator, as the ground-state mean-field reference. This unrestricted Hartree-Fock (UHF) state can be expressed as

$$|\phi_{\mathrm{UHF}}\rangle = \frac{1}{(1+\kappa^2)^{n/2}} e^{\kappa J_+} |\phi_{\mathrm{RHF}}\rangle, \quad (A5)$$

where κ parametrizes the UHF state. Using a UHF state as an approximation to the ground state is equivalent to a rotation of the underlying $su(2)$ algebra such that the UHF wave function has all down-spins in the new basis. The $su(2)$ operators in the new basis are related to the original operators via the following transformation:

$$J_+ = \frac{K_+ - \kappa^2 K_- - 2\kappa K_z}{1 + \kappa^2}, \quad (A6a)$$

$$J_- = \frac{K_- - \kappa^2 K_+ - 2\kappa K_z}{1 + \kappa^2}, \quad (A6b)$$

$$J_z = \frac{\kappa(K_+ + K_-) + (1 - \kappa^2)K_z}{1 + \kappa^2}. \quad (A6c)$$

The Hamiltonian, expressed in terms of the K operators, becomes

$$H = h_z K_z + h_\pm(K_+ + K_-) + v_\pm(K_+K_+ + K_-K_-)$$
$$+ v_z K_z^2 + v_\times K_+K_- + v_{\pm z}(K_+K_z + K_zK_-). \quad (A7)$$

Exact expressions for the Hamiltonian parameters in the above expression can be found in the Appendix of Ref. [25]. The UHF mean-field Hamiltonian and the corresponding mean-field energy are

$$H_0^{\mathrm{UHF}} = h_z K_z - \frac{v_z n}{2} K_z, \quad (A8a)$$

$$E_{\mathrm{UHF}} = -\frac{h_z n}{2} + \frac{v_z n^2}{4}. \quad (A8b)$$

The UHF energy is minimized to find the optimal value of the rotation parameter κ. As we have mentioned in the main text, $\kappa \neq 0$ only for $x \leqslant x_c$. While working in the broken-symmetry regime (i.e., $x \leqslant x_c$), we use H_0^{UHF} to define the mean-field thermal state.

APPENDIX B: SYMMETRY BREAKING IN TFIM

The TFIM has a $\mathbb{Z}_2$ symmetry; i.e., if all the spins (in z basis) are flipped, then the energy remains unchanged. In the Ising limit, i.e., for $g = 0$, the model has a doubly degenerate ferromagnetic ground state. In the thermodynamic limit, the
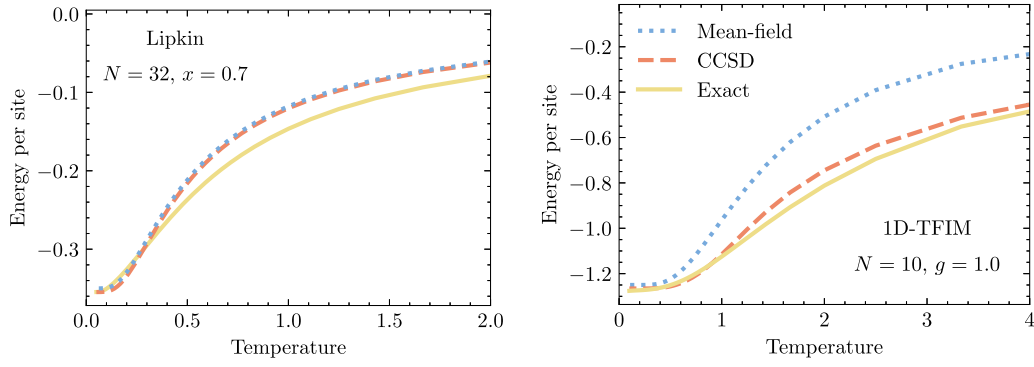
FIG. 6. Total internal energy per site for 32-site Lipkin model with $x = 0.7$ (left) and 10-site transverse field Ising model at $g = 1.0$ (right). The figure compares mean-field and CCSD approximations against the exact results.

$\mathbb{Z}_2$ symmetry breaks spontaneously. For finite systems, we can break this symmetry artificially by introducing the following symmetry-broken mean-field state to approximate the ground state:

$$|\phi_{\text{UHF}}\rangle = \bigotimes_p \left(-\sin\frac{\theta}{2}|\uparrow\rangle_p + \cos\frac{\theta}{2}|\downarrow\rangle_p \right). \quad \text{(B1)}$$

The mean-field Hamiltonian and its corresponding energy depend on the rotation parameter θ, and for the one-dimensional periodic case with N spins, in the presence of an external magnetizing z field f [see Eq. (38) for the Hamiltonian], they are given by

$$H_0 = 2 \sum_i \left((\cos\theta + f)J_i^z + gJ_i^x \right), \quad \text{(B2a)}$$

$$E_{\text{UHF}} = -N(\cos^2\theta + g\sin\theta + f\cos\theta). \quad \text{(B2b)}$$

The UHF energy is minimized with respect to the rotation parameter θ, the optimal values of which (at $f = 0$) are

$$\theta = \begin{cases} \arcsin g/2, & g \leqslant 2 \\ \pi/2, & g > 2. \end{cases} \quad \text{(B3)}$$

We use this H_0 (with optimized θ) that breaks the $\mathbb{Z}_2$ symmetry, except when $\theta = \pi/2$, to construct the mean-field thermal state for the TFIM.

In the same way as for the Lipkin model, we can introduce a rotation in the $su(2)$ algebra so that, in the new basis, the UHF state corresponds to all down-spins. The full Hamiltonian in the rotated basis then becomes

$$H = -4 \sum_i \left[\cos^2\theta K_i^z K_{i+1}^z + \sin^2\theta K_i^x K_{i+1}^x \right.$$
$$- \frac{1}{2}\sin 2\theta \left(K_i^x K_{i+1}^z + K_i^z K_{i+1}^x \right)$$
$$\left. + \frac{g}{2}\left(\cos\theta K_i^x + \sin\theta K_i^z \right) \right]. \quad \text{(B4)}$$

Similarly, the mean-field Hamiltonian, written in the new basis, becomes simply

$$H_0 = -2 \sum_i (-\cos^2\theta + g\sin\theta)K_i^z. \quad \text{(B5)}$$

While the results do not depend on the choice of the basis, it is generally convenient to construct CC-like wave functions

starting with a mean-field state that corresponds to all down-spins. Our description of the mean-field theory for the TFIM closely follows Ref. [22] and we recommend the reader to refer to this article for further details.

APPENDIX C: TOTAL INTERNAL ENERGY FOR 1D TFIM

The logarithmic plots for absolute internal energy errors (Figs. 1 and 2) show several spikes that correspond to zero error. These occur when the approximate and exact internal energy curves intersect. These are a result of the nonvariational nature of the internal energy and may be worsened by the nonvariational character of the CC theory. In Fig. 6, we plot the total internal energy per site for the 32-site Lipkin model with $x = 0.7$ in the left panel, and the 10-site 1D TFIM at $g = 1$ in the right panel. It is clear that neither the mean-field nor the CCSD internal energies provide a variational upper bound to the exact internal energy. For example, in the 1D TFIM plot, the CCSD internal energy crosses the exact curve at two different temperatures. These correspond to the two spikes that are observed in Fig. 2.

APPENDIX D: THERMAL PROPERTIES

In Sec. IV, we described two different ways to compute thermal properties: CC expectation values and λ derivatives. Here, we will discuss a few details about the equivalence, or lack thereof, between the two approaches.

1. Exact theory

For a first-derivative property (e.g., the dipole moment or magnetization density), the property can be computed as the derivative of the free energy. We thus define a λ-dependent thermal state,

$$|\Psi(\beta, \lambda)\rangle = e^{-\beta H(\lambda)/2}|\mathbb{I}\rangle, \quad \text{(D1)}$$

where $H(\lambda) = H + \lambda\mathcal{O}$. The norm of $|\Psi(\beta, \lambda)\rangle$ gives us the partition function, $\mathcal{Z}(\beta, \lambda) = \langle\Psi(\beta, \lambda)|\Psi(\beta, \lambda)\rangle$, from which we can extract the free energy and write the thermal average of $\mathcal{O}$ as

$$\langle\mathcal{O}\rangle = -\frac{1}{\beta}\frac{\partial}{\partial\lambda}\ln\mathcal{Z}(\beta, \lambda). \quad \text{(D2)}$$

The λ derivative of the partition function can be simplified using the Wilcox identity [60],

$$\frac{\partial \mathcal{Z}(\beta, \lambda)}{\partial \lambda} = \frac{\partial}{\partial \lambda} \langle \mathbb{I} | e^{-\beta H(\lambda)} | \mathbb{I} \rangle \tag{D3a}$$

$$= -\beta \int_0^1 d\alpha \, \langle \mathbb{I} | e^{-(1-\alpha)\beta H(\lambda)} \mathcal{O} e^{-\alpha\beta H(\lambda)} | \mathbb{I} \rangle . \tag{D3b}$$

Using the fact that for a physical operator X, $\langle \mathbb{I} | X | \mathbb{I} \rangle = \text{Tr} X$, combined with the cyclic property of trace, we get

$$\frac{\partial \mathcal{Z}(\beta, \lambda)}{\partial \lambda} = -\beta \int_0^1 d\alpha \, \text{Tr}(e^{-(1-\alpha)\beta H(\lambda)} \mathcal{O} e^{-\alpha\beta H(\lambda)}) \tag{D4a}$$

$$= -\beta \int_0^1 d\alpha \, \text{Tr}(e^{-\beta H(\lambda)/2} \mathcal{O} e^{-\beta H(\lambda)/2}) \tag{D4b}$$

$$= -\beta \langle \Psi(\beta, \lambda) | \mathcal{O} | \Psi(\beta, \lambda) \rangle . \tag{D4c}$$

Therefore, in the exact theory, the λ derivative of the free energy is equivalent to the expectation value of $\mathcal{O}$ over the thermal state, i.e.,

$$\langle \mathcal{O} \rangle = -\frac{1}{\beta} \frac{\partial}{\partial \lambda} \ln \mathcal{Z}(\beta, \lambda) = \frac{\langle \Psi(\beta, \lambda) | \mathcal{O} | \Psi(\beta, \lambda) \rangle}{\langle \Psi(\beta, \lambda) | \Psi(\beta, \lambda) \rangle}. \tag{D5}$$

2. Mean-field theory

Unlike the exact theory, the expectation value and the free-energy derivative do not necessarily lead to the same results in the mean-field theory. The expectation value of a physical operator $\mathcal{O}$ over the mean-field thermal state is defined as

$$\langle \mathcal{O} \rangle_{0,\text{exp}} = \langle 0(\beta) | \mathcal{O} | 0(\beta) \rangle . \tag{D6}$$

On the other hand, we can also define $\langle \mathcal{O} \rangle_0$ as the λ derivative of the mean-field free energy, i.e.,

$$\langle \mathcal{O} \rangle_{0,\lambda} = -\frac{1}{\beta} \lim_{\lambda \to 0} \frac{\partial}{\partial \lambda} \ln \langle \Phi(\beta, \lambda) | \Phi(\beta, \lambda) \rangle , \tag{D7}$$

where

$$| \Phi(\beta, \lambda) \rangle = e^{-\beta(H_0 + \lambda \mathcal{O}_0)/2} | \mathbb{I} \rangle , \tag{D8}$$

with $\mathcal{O}_0$ being the mean-field (or one-body) part of $\mathcal{O}$. For a general two-body $\mathcal{O}$, Eqs. (D6) and (D7) are not equivalent. The thermal expectation value of only the mean-field component of $\mathcal{O}$ is identical to the λ derivative, i.e.,

$$\langle \mathcal{O} \rangle_{0,\lambda} = \langle \mathcal{O}_0 \rangle_{0,\text{exp}} , \tag{D9}$$

which implies that only when the observable $\mathcal{O}$ is a one-body operator, the mean-field expectation value and λ derivative of the mean-field free energy are identical.

Note that in the foregoing equations, we have ignored the orbital optimization of the mean-field free energy in the presence of $\lambda \mathcal{O}_0$, which, if included, will result in a mean-field Hamiltonian $H_0(\beta, \lambda)$ that is nonlinear in λ.

3. Coupled cluster theory

It follows directly from mean-field theory that for approximate wave-function theories, expectation values over thermal state and free-energy derivatives do not necessarily produce identical results. Let us see an explicit proof for the coupled cluster theory. The CC expectation value and λ-derivative expressions are given in Eqs. (20) and (26),

$$\langle \mathcal{O} \rangle_{CC,\text{exp}} = \frac{\langle \Psi_L | \mathcal{O} | \Psi_R \rangle}{\langle \Psi_L | \Psi_R \rangle}, \tag{D10a}$$

$$\langle \mathcal{O} \rangle_{CC,\lambda} = \lim_{\lambda \to 0} \frac{\partial F}{\partial \lambda}, \tag{D10b}$$

where F is the free energy of the system and is defined as

$$F(\beta, \lambda) = -\frac{1}{\beta} \ln \mathcal{Z}_{CC}(\beta, \lambda) \tag{D11a}$$

$$= \frac{1}{\beta} \int_0^\beta d\tau E_{CC}(\tau, \lambda), \tag{D11b}$$

with $E_{CC}(\tau, \lambda) = \langle \Psi_L | H(\lambda) | \Psi_R \rangle / \langle \Psi_L | \Psi_R \rangle$. Recall that $\langle \Psi_L |$, $| \Psi_R \rangle$, and the CC partition function $\mathcal{Z}_{CC}$ are given by

$$\langle \Psi_L | = \langle \Phi(\beta, \lambda) | (1 + Z) e^{-T} e^{z_0}, \tag{D12a}$$

$$| \Psi_R \rangle = e^{t_0 + T} | \Phi(\beta, \lambda) \rangle , \tag{D12b}$$

$$\mathcal{Z}_{CC}(\beta, \lambda) = \langle \Phi(\beta, \lambda) | \Phi(\beta, \lambda) \rangle e^{t_0 + z_0}. \tag{D12c}$$

To show that CC expectation value and λ-derivative approaches are different, we will show that

$$\frac{\partial}{\partial \beta} (\beta \langle \mathcal{O} \rangle_{CC,\text{exp}}) \neq \frac{\partial}{\partial \beta} (\beta \langle \mathcal{O} \rangle_{CC,\lambda}). \tag{D13}$$

The left-hand side (LHS) can be simplified as

$$\text{LHS} = \left(1 + \beta \frac{\partial}{\partial \beta}\right) \frac{\langle \Psi_L | \mathcal{O} | \Psi_R \rangle}{\langle \Psi_L | \Psi_R \rangle} \tag{D14a}$$

$$= \langle \mathcal{O} \rangle_{CC,\text{exp}} + \frac{\beta}{2} \left(2E(\beta) \langle \mathcal{O} \rangle_{CC,\text{exp}} - \frac{\langle \Psi_L | (H\mathcal{O} + \mathcal{O}H) | \Psi_R \rangle}{\langle \Psi_L | \Psi_R \rangle} \right). \tag{D14b}$$

On the other hand, the right-hand side (RHS), before taking the limit of $\lambda \to 0$, simplifies into

$$\text{RHS} = \frac{\partial}{\partial \beta} \int_0^\beta d\tau \frac{\partial E_{CC}(\tau, \lambda)}{\partial \lambda} \tag{D15a}$$

$$= \frac{\partial E(\beta, \lambda)}{\partial \lambda} = \frac{\partial}{\partial \lambda} \langle 0(\beta, \lambda) | (1 + Z) e^{-T} H(\lambda) e^T | 0(\beta, \lambda) \rangle \tag{D15b}$$

$$
\begin{aligned}
&= \langle \mathcal{O} \rangle_{CC,\exp} + \langle 0(\beta,\lambda)| \frac{\partial Z}{\partial \lambda} e^{-T} H(\lambda) e^{T} |0(\beta,\lambda)\rangle + \langle 0(\beta,\lambda)|(1+Z)\left(\frac{\partial e^{-T}}{\partial \lambda} H e^{T} + e^{-T} H \frac{\partial e^{T}}{\partial \lambda} \right)|0(\beta,\lambda)\rangle \\
&\quad + \frac{\partial \langle 0(\beta,\lambda)|}{\partial \lambda}(1+Z) e^{-T} H e^{T} |0(\beta,\lambda)\rangle + \langle 0(\beta,\lambda)| (1+Z) e^{-T} H e^{T} \frac{\partial |0(\beta,\lambda)\rangle}{\partial \lambda}.
\end{aligned}
\tag{D15c}
$$

At first glance, the left- and right-hand-side expressions do not show any resemblance beyond the $\langle \mathcal{O} \rangle_{CC,\exp}$ term. Even if we ignore the λ dependence in the mean-field reference, and therefore the quasiparticle excitation and deexcitation operators, we get

$$
\text{RHS} = \langle \mathcal{O} \rangle_{CC,\exp} + \sum_{\mu} \frac{\partial z_{\mu}}{\partial \lambda} \langle 0(\beta)|\tau_{\mu}^{\dagger} e^{-T} H(\lambda) e^{T} |0(\beta)\rangle + \sum_{\mu} \frac{\partial t_{\mu}}{\partial \lambda} \langle 0(\beta)|(1+Z)[e^{-T} H(\lambda) e^{T}, \tau_{\mu}]|0(\beta)\rangle,
\tag{D16}
$$

which is clearly different from the LHS. It is well established in ground-state CC theory that, in the absence of orbital relaxation effects, i.e., using a λ-independent mean-field reference, CC expectation value and λ-derivative approaches result in identical results. However, the same is not obviously true at nonzero temperatures.

---

[1] P. Pfeuty, Ann. Phys. **57**, 79 (1970).

[2] H. J. Lipkin, N. Meshkov, and A. J. Glick, Nucl. Phys. **62**, 188 (1965).

[3] N. Meshkov, A. J. Glick, and H. J. Lipkin, Nucl. Phys. **62**, 199 (1965).

[4] A. J. Glick, H. J. Lipkin, and N. Meshkov, Nucl. Phys. **62**, 211 (1965).

[5] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[6] S. T. Belyaev, Kgl. Danske Videnskab. Selskab. Mat.-Fys. Medd. **31** (1959), https://www.osti.gov/biblio/4262925-effect-pairing-correlations-nuclear-properties.

[7] W. Ogle, S. Wahlborn, R. Piepenbring, and S. Fredriksson, Rev. Mod. Phys. **43**, 424 (1971).

[8] A. W. Sandvik and J. Kurkijärvi, Phys. Rev. B **43**, 5950 (1991).

[9] A. W. Sandvik, Phys. Rev. B **59**, R14157(R) (1999).

[10] S. R. White, Phys. Rev. B **48**, 10345 (1993).

[11] J. I. Cirac and F. Verstraete, J. Phys. A: Math. Theor. **42**, 504004 (2009).

[12] P. Kopietz, L. Bartosch, and F. Schütz, *Introduction to the Functional Renormalization Group*, Lecture Notes in Physics (Springer-Verlag, Berlin, 2010).

[13] S. Li, J. Chem. Phys. **120**, 5017 (2004).

[14] L. Isaev, G. Ortiz, and J. Dukelsky, Phys. Rev. B **79**, 024409 (2009).

[15] Z. Fan and Q.-l. Jie, Phys. Rev. B **91**, 195118 (2015).

[16] K. Gunst, S. Wouters, S. De Baerdemacker, and D. Van Neck, Phys. Rev. B **95**, 195127 (2017).

[17] A. Papastathopoulos-Katsaros, C. A. Jiménez-Hoyos, T. M. Henderson, and G. E. Scuseria, arXiv:2102.10043.

[18] T. D. Crawford and H. F. Schaefer, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (Wiley, New York, 2000), pp. 33–136.

[19] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[20] R. F. Bishop, Theor. Chim. Acta **80**, 95 (1991).

[21] J. Rosenfeld, N. E. Ligterink, and R. F. Bishop, Phys. Rev. B **60**, 4030 (1999).

[22] J. Rosenfeld and N. E. Ligterink, Phys. Rev. B **62**, 308 (2000).

[23] R. F. Bishop, N. Ligterink, and N. R. Walet, Int. J. Mod. Phys. B **20**, 4992 (2006).

[24] T. M. Henderson, G. E. Scuseria, J. Dukelsky, A. Signoracci, and T. Duguet, Phys. Rev. C **89**, 054305 (2014).

[25] J. M. Wahlen-Strothman, T. M. Henderson, M. R. Hermes, M. Degroote, Y. Qiu, J. Zhao, J. Dukelsky, and G. E. Scuseria, J. Chem. Phys. **146**, 054110 (2017).

[26] G. Harsha, T. Shiozaki, and G. E. Scuseria, J. Chem. Phys. **148**, 044107 (2018).

[27] R. F. Bishop, P. H. Y. Li, O. Götze, and J. Richter, Phys. Rev. B **100**, 024401 (2019).

[28] T. M. Henderson and G. E. Scuseria, J. Chem. Phys. **153**, 084111 (2020).

[29] A. Khamoshi, G. P. Chen, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **154**, 074113 (2021).

[30] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).

[31] A. E. Feiguin and S. R. White, Phys. Rev. B **72**, 220401(R) (2005).

[32] P. Czarnik, M. M. Rams, and J. Dziarmaga, Phys. Rev. B **94**, 235142 (2016).

[33] N. Niggemann, B. Sbierski, and J. Reuther, Phys. Rev. B **103**, 104431 (2021).

[34] B. M. Rubenstein, S. Zhang, and D. R. Reichman, Phys. Rev. A **86**, 053606 (2012).

[35] J. Claes and B. K. Clark, Phys. Rev. B **95**, 205109 (2017).

[36] H. R. Petras, S. K. Ramadugu, F. D. Malone, and J. J. Shepherd, J. Chem. Theory Comput. **16**, 1029 (2020).

[37] Y. Liu, T. Shen, H. Zhang, and B. Rubenstein, J. Chem. Theory Comput. **16**, 4298 (2020).

[38] T. Shen, Y. Liu, Y. Yu, and B. M. Rubenstein, J. Chem. Phys. **153**, 204108 (2020).

[39] S. R. White, Phys. Rev. Lett. **102**, 190601 (2009).

[40] E. M. Stoudenmire and S. R. White, New J. Phys. **12**, 055026 (2010).

[41] G. Sanyal, S. H. Mandal, and D. Mukherjee, Chem. Phys. Lett. **192**, 55 (1992).

[42] G. Sanyal, S. H. Mandal, S. Guha, and D. Mukherjee, Phys. Rev. E **48**, 3373 (1993).

[43] S. H. Mandal, G. Sanyal, and D. Mukherjee, in *Microscopic Quantum Many-Body Theories and Their Applications*, Lecture Notes in Physics, edited by J. Navarro and A. Polls (Springer, Berlin, 1998), pp. 93–117.

[44] S. H. Mandal, R. Ghosh, G. Sanyal, and D. Mukherjee, Int. J. Mod. Phys. B **17**, 5367 (2003).

[45] M. R. Hermes and S. Hirata, J. Chem. Phys. **143**, 102818 (2015).

[46] A. F. White and G. K.-L. Chan, J. Chem. Theory Comput. **14**, 5690 (2018).

[47] F. Hummel, J. Chem. Theory Comput. **14**, 6505 (2018).

[48] S. Hirata and P. K. Jha, in *Annual Reports in Computational Chemistry*, edited by D. A. Dixon (Elsevier, Amsterdam, 2019), Vol. 15, pp. 17–37.

[49] G. Harsha, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **150**, 154109 (2019).

[50] G. Harsha, T. M. Henderson, and G. E. Scuseria, J. Chem. Theory Comput. **15**, 6127 (2019).

[51] G. Harsha, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **153**, 124115 (2020).

[52] H. Matsumoto, Y. Nakano, H. Umezawa, F. Mancini, and M. Marinaro, Prog. Theor. Phys. **70**, 599 (1983).

[53] G. W. Semenoff and H. Umezawa, Nucl. Phys. B **220**, 196 (1983).

[54] H. Umezawa, Prog. Theor. Phys. Suppl. **80**, 26 (1984).

[55] T. S. Evans, I. Hardman, H. Umezawa, and Y. Yamanaka, J. Math. Phys. **33**, 370 (1992).

[56] M. Suzuki, J. Stat. Phys. **42**, 1047 (1986).

[57] T. Hatsuda, Nucl. Phys. A **492**, 187 (1989).

[58] N. R. Walet and A. Klein, Nucl. Phys. A **510**, 261 (1990).

[59] J. Zhao, Ph.D. thesis, Rice University, 2018, https://scholarship.rice.edu/handle/1911/105668.

[60] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).

[61] J. R. Dormand and P. J. Prince, J. Comput. Appl. Math. **6**, 19 (1980).

[62] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems*, 2nd ed., Springer Series in Computational Mathematics (Springer-Verlag, Berlin, 1993).

[63] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey *et al.*, Nat. Methods **17**, 261 (2020).

[64] G. Ortiz, R. Somma, J. Dukelsky, and S. Rombouts, Nucl. Phys. B **707**, 421 (2005).

[65] S. Lerma H. and J. Dukelsky, Nucl. Phys. B **870**, 421 (2013).