P Selected for a Viewpoint in *Physics*
PHYSICAL REVIEW LETTERS

Minimally Entangled Typical Quantum States at Finite Temperature

Steven R. White

Department of Physics and Astronomy, University of California, Irvine, California 92697, USA
(Received 2 March 2009; revised manuscript received 24 March 2009; published 11 May 2009)

We introduce a class of states, called minimally entangled typical thermal states, designed to resemble a typical state of a quantum system at finite temperature with a bias towards classical (minimally entangled) properties. These states reveal in an intuitive way properties such as short-range order which may often be hidden. A finite-T density matrix renormalization group algorithm is presented which is only modestly slower than the $T = 0$ density matrix renormalization group.

What is a typical wave function of a quantum system at a finite temperature? The fundamental proposition of statistical mechanics is that the density matrix of a system at inverse temperature β with Hamiltonian H is $\rho = \exp(-\beta H)$. One can regard $\rho = \exp(-\beta H)$ as arising from several different physical situations: from an ensemble average of pure states, from the long time average of one system, from quantum mechanical entanglement with a heat bath which produces mixed states, or from some combination of these effects. The resulting predictions of statistical mechanics depend only on ρ. On the other hand, statistical mechanics is an idealization; a real physical system has a specific history and environment which may favor thinking about it in one way over another. Here we will focus on the ensemble-of-pure-states point of view. We have in mind equilibrating the system with weak coupling to a heat bath and then moving the heat bath far from the system, removing any couplings. From this viewpoint, our question is a natural one. In this Letter, we propose a set of idealized states which we argue are useful to think of as "typical" and whose ensemble generates $\rho = \exp(-\beta H)$. In addition, the algorithm we introduce to generate them provides a substantially more efficient route to determining finite temperature properties of lattice models when using diagonalization, density matrix renormalization group (DMRG) [1], and tensor product wave function approaches [2,3].

What do we mean by typical? We mean that there is a set of states $\{|\phi(i)\rangle\}$ with unnormalized probabilities $P(i)$, from which we can select states, with

$$\sum_i P(i)|\phi(i)\rangle\langle\phi(i)| = e^{-\beta H}. \qquad (1)$$

Then the expectation value of any Hermitian operator A can be determined by an unweighted average of $\langle\phi(i)|A|\phi(i)\rangle$, with each $|\phi(i)\rangle$ chosen at random according to $P(i)$. We also impose looser criteria based on physics: that one can imagine some physical thermalization process which might generate the $\{|\phi(i)\rangle\}$ and that the $\{|\phi(i)\rangle\}$ do not exhibit special "atypical" physical charac-

teristics. We do not require that every state in the Hilbert space be included in the $\{|\phi(i)\rangle\}$.

For classical systems on a lattice, the only reasonable typical states are the classical product states (CPSs) $|i\rangle = \prod_{\text{sites } \ell}|i_\ell\rangle$, where i_ℓ labels the states of a site. For example, for an Ising model, a CPS is a spin configuration, e.g., $|i\rangle = |\uparrow\downarrow\downarrow\uparrow\ldots\rangle$. These states are often generated numerically and provide an intuitive understanding of a system's properties which would be difficult to obtain from the system's density matrix. For quantum spin systems, one can generate CPSs, but they are not typical wave functions; e.g., at $T = 0$, the typical wave function should be the ground state, not a CPS.

The energy eigenvalues E_s and eigenstates $|s\rangle$ satisfy $\rho = \sum_s e^{-\beta E_s}|s\rangle\langle s|$ and thus Eq. (1). However, they should not be thought of as typical states. Schrödinger called this idea "altogether wrong" and "irreconcilable with the very foundations of quantum mechanics" [4]. For a large system, excluding very low temperature, equilibration processes do not drive the system to any single eigenstate. Any such process would take an exponentially long time (in the number of particles N) because of the exponentially small energy level spacing. The eigenstates are also exponentially sensitive to uncertainties in the Hamiltonian. Nevertheless, more recent introductions to statistical mechanics than Schrödinger's often give the impression that the typical thermal wave function is an eigenstate of the Hamiltonian [5].

It is easy to construct other states satisfying Eq. (1). Let $\{|i\rangle\}$ be any complete orthonormal basis of the system. Define the normalized (but not orthogonal) set of typical states

$$|\phi(i)\rangle \equiv P(i)^{-1/2} \exp(-\beta H/2)|i\rangle, \qquad (2)$$

$$P(i) \equiv \langle i| \exp(-\beta H)|i\rangle = \text{Tr}\{\rho|i\rangle\langle i|\}. \qquad (3)$$

Note that the partition function is given by $Z = \text{Tr}\rho = \sum_i P(i)$. We see Eq. (1) immediately follows, and

$$\langle A\rangle \equiv Z^{-1}\text{Tr}\{\rho A\} = Z^{-1}\sum_i P(i)\langle\phi(i)|A|\phi(i)\rangle. \qquad (4)$$

190601-1

Similar results are obtained if the states $\{|i\rangle\}$ are not orthonormal, provided there exists a positive set of weights $p(i)$ such that $\sum_i p(i)|i\rangle\langle i| = 1$, and similarly for a continuous distribution of states.

The energy eigenstates can serve as the set $|i\rangle$, in which case $|\phi(i)\rangle = |i\rangle$. Another choice is to select the $\{|i\rangle\}$ as random normalized vectors in the Hilbert space, selected using the Haar measure. Both of these approaches are intractable computationally except on the smallest systems. The first would be unsuitable even for a Lanczos approach because of the small level spacings—a full diagonalization of H would be required. These choices are also poor from a physical point of view. In a broken symmetry phase, these states would be highly nonclassical superpositions of many states with different values of the order parameter. If the system consisted of two widely separated noninteracting subsystems, the random vector approach would give highly entangled states of the two subsystems. These choices ignore decoherence effects, which tend to eliminate highly nonclassical entanglement.

We take as our favored typical states $\{|\phi(i)\rangle\}$ the ones with the least entanglement, generated by taking $\{|i\rangle\}$ to be a CPS. By entanglement we mean that we consider dividing the system into two parts, say, by taking a dividing plane, and calculating the von Neumann entanglement entropy S between the two parts. A CPS has $S = 0$ for any division. At nonzero β we expect the resulting $|\phi(i)\rangle$ to have minimal entropy within this general class of states, and so we call them minimally entangled typical thermal states (METTS). METTS have a number of nice properties. A METTS for a system with noninteracting subsystems is a product of METTS for the subsystems. For systems with long-range order, METTS break symmetries, choosing an order parameter at random. Even for systems without broken symmetries, METTS reveal underlying short-range order.

We give two approaches to the calculation of METTS. The ancilla method [6,7] uses a set of auxiliary sites acting as a heat bath and has been used to simulate finite temperatures using time-dependent DMRG methods [8]. Its adaptation to produce METTS resembles a highly idealized physical thermalization process. The pure-state method does not use a heat bath and is more efficient.

To describe the ancilla method [6], let us take the system A to be composed of N spins with $S = 1/2$. The heat bath B is also composed of N $S = 1/2$ spins (called ancilla), and we pair up the spins in A and B—for a 1D system, think of a ladder. Label the sites by ℓ, and let $|i_\ell\rangle_A$ label the local states of the system site at ℓ, and similarly for B. The initial unnormalized pure state of $A + B$ which describes infinite temperature is

$$|\psi(\beta = 0)\rangle = \sum_{i_1} \ldots \sum_{i_N} |i_1\rangle_A |i_1\rangle_B \ldots |i_N\rangle_A |i_N\rangle_B. \quad (5)$$

This state is a product of site-ancilla pair states, with each pair maximally entangled. If one traces out the ancilla from $|\psi\rangle\langle\psi|$, one obtains the infinite temperature density matrix 1. The Hamiltonian of $A + B$ is that of A alone: there are no A-B or B-B terms. Let

$$|\psi(\beta)\rangle = \exp(-\beta H/2)|\psi(0)\rangle. \quad (6)$$

We find $\mathrm{Tr}_B|\psi(\beta)\rangle\langle\psi(\beta)| = \exp(-\beta H)$. Alternatively, one can measure any operator O of A as $\langle O \rangle = \langle\psi(\beta)|O|\psi(\beta)\rangle$. The calculation of $|\psi(\beta)\rangle$ for 1D systems is easily performed using imaginary time-dependent DMRG [8], with initial state $|\psi(\beta = 0)\rangle$.

To obtain a pure state for A from the entangled state of $A + B$, we perform a physical measurement of all of the spins of B. A physical measurement projects the wave function into one eigenstate of the measured operator, with the appropriate probability. Specifically, to measure one particular spin in the z direction, compute $P(\uparrow) = \langle S_z \rangle + \frac{1}{2}$, let $P(\downarrow) = 1 - P(\uparrow)$, and set

$$|\psi\rangle \rightarrow \begin{cases} P(\uparrow)^{-1/2}|\uparrow\rangle\langle\uparrow|\psi\rangle & \text{prob } P(\uparrow), \\ P(\downarrow)^{-1/2}|\downarrow\rangle\langle\downarrow|\psi\rangle & \text{prob } P(\downarrow). \end{cases} \quad (7)$$

We get the same probability distribution whether we measure the sites sequentially or jointly all at once, but sequentially is much more convenient numerically, taking one half-sweep in DMRG [9]. We are free to measure each spin with respect to any axis, all the same or different, randomly or predetermined. The probability of the final CPS $|i\rangle_B$ is given by Eq. (3). The measurement puts the combined system into the product state

$$P(i)^{-1/2}|i\rangle_B\langle i|_B\psi(\beta)\rangle = |i\rangle_B|\phi(i)\rangle_A. \quad (8)$$

At this point, one can ignore B. Note that the initial perfect entanglement takes the place of the coupling one would have in a real thermalization process.

In the pure-state method, we start with any CPS $|i\rangle$ and apply $\exp(-\beta H/2)$. We then physically measure a new CPS $|i'\rangle$ from this state and apply $\exp(-\beta H/2)$ to it, etc. We call one iteration a "thermal step." This process resembles Monte Carlo sweep, but with the quantum measurement process taking the place of the usual spin flips. The set of METTS are a fixed point of this process: consider an infinite ensemble of such systems, initially with $|i\rangle$ distributed with probability $P(i)$. Then by Eq. (4), the ensemble of $|\phi(i)\rangle$ correctly reproduces all thermodynamic measurements, so a set of $|i'\rangle$ determined from it is correctly distributed with probability $P(i')$.

To study METTS in more detail, we consider the one-dimensional $S = \frac{1}{2}$ Heisenberg model, with Hamiltonian

$$H = \sum_\ell \vec{S}_\ell \cdot \vec{S}_{\ell+1} \quad (9)$$

and with open boundary conditions. We implement the algorithms using time-dependent DMRG with a second-order breakup and with a time step of 0.05. The physical

measurements generating the $|i\rangle$ were done at different random orientations for each spin and thermal time step.

In Fig. 1, we show that the two algorithms give the same (numerically exact) results for the energy on a 100-site Heisenberg chain. The main figure shows that, for the pure-state method, one reaches the equilibrium distribution for the energy very precisely after 5–10 thermal steps, starting from a random configuration. This suggests that the thermal-step autocorrelation time (similar to a Monte Carlo autocorrelation time) is very short, the key to efficient sampling. In practical calculations (inset), one does one long run with many thermal steps, discarding the first results as being a warm-up, say, about 10 steps. Averaging over only 100 METTS obtained in $N_\tau = 100$ thermal steps ($+10$ for the warm-up), we obtain the total energy to a relative accuracy of about 10^{-5}. The fluctuations in the total energy are quite small; one can obtain reasonable results with only one METTS.

With DMRG, particularly for low temperatures and modest accuracies, the pure-state METTS method is much faster than the ancilla method for obtaining thermal averages. (For averages, there is no point in generating METTS with the ancilla method, since averages can be measured directly.) Suppose for a specified accuracy a system requires m_0 states per block for a $T = 0$ standard DMRG calculation. In pure-state METTS we solve the imaginary-time-dependent Schrödinger equation from 0 to $\beta/2$. We find that, for pure-state METTS, the m required starts at 1 for small imaginary time and saturates to m_0 for very large imaginary times, which are needed only for large β. The calculation time scales as $Nm_0^3\beta N_\tau$, where N is the number of sites. In the ancilla method, in the limit of low temperatures, the heat bath and the system both independently encode the ground state, as a product state but with their sites intermingled. This means that DMRG requires m_0^2 states, and the calculation time scales as $Nm_0^6\beta$, bigger by a factor of m_0^3/N_τ compared to the pure-state METTS approach. Typical values of m_0 are 50–5000 for systems ranging from simple 1D spin chains to 2D clusters with width 8–10. Consequently, taking $N_\tau = 10$–100, the pure-state METTS method is faster by a factor of 10^3–10^{10}.

In Fig. 2, we show properties of some METTS for a Heisenberg chain. All of the measurements show substantial randomness, which diminishes at lower temperatures as the METTS approach the ground state. Since the model is antiferromagnetic, we multiply the spin measurements by $(-1)^\ell$ to make twisting of the antiferromagnetic order more apparent. For example, for $\beta = 8$, pronounced twisting is visible near $\ell = 105$–110. The values of $\langle \vec{S} \cdot \vec{S} \rangle$ show an increase in dimerization in the same region. Similar twisting and dimerization is visible at $\beta = 3$ near $\ell = 100$–108. At a finite temperature, the system has a finite spin-spin correlation length; this could come about (we imagine) via random twisting of the spin order, by regions with strong dimerization, or by some combination. In these METTS both effects occur, with twisting being somewhat more pronounced. The open squares measure
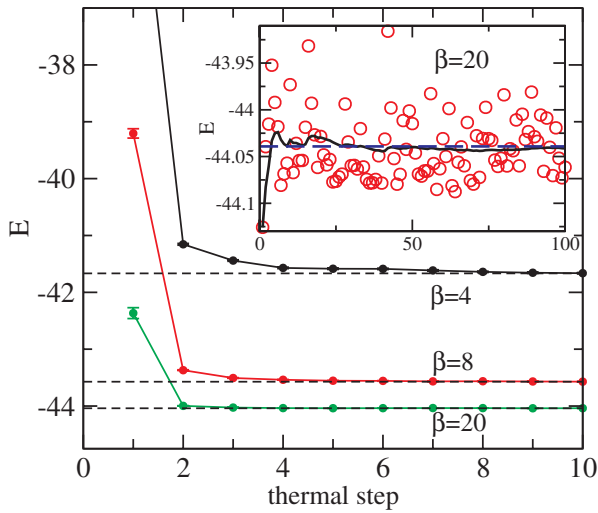

FIG. 1 (color online).   Energy of a 100-site Heisenberg chain at various temperatures. The dashed lines label the results of the ancilla approach, considered exact. The symbols are derived from repeated use of the pure-state method, with each use starting from a new completely random state $|i\rangle$ and proceeding 10 thermal steps. For a fixed thermal step, we averaged over the ensemble. The inset shows one long pure-state calculation. The open circles show energies of individual METTS, while the solid line shows the moving average.
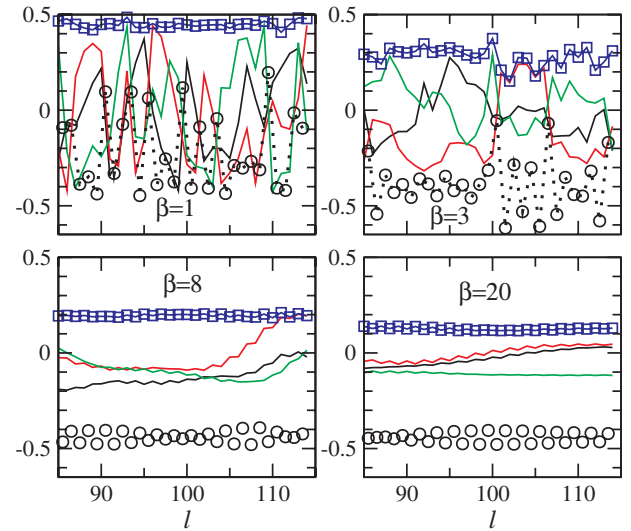

FIG. 2 (color online).   Properties of METTS for a 200-site Heisenberg chain, central 30 sites. Each panel shows properties of a METTS generated for the indicated temperature (the METTS in the different panels are unrelated). The three solid lines (red, black, and green) without symbols show $(-1)^\ell\langle S^\alpha \rangle$, for $\alpha = x, y, z$ (which line is x, etc., is arbitrary). The open squares at the top show C_ℓ. The open circles show $\langle \vec{S} \cdot \vec{S} \rangle$ on each bond.
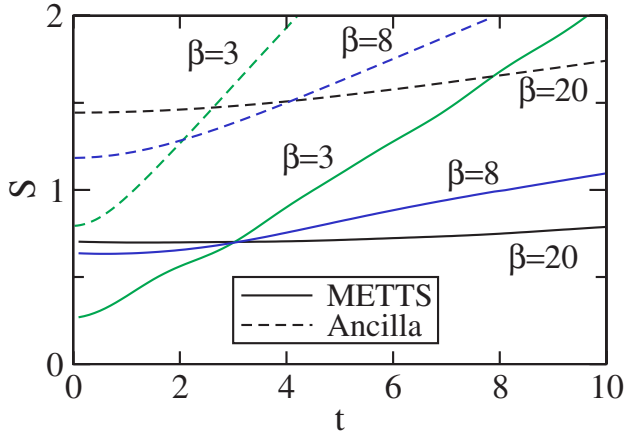
FIG. 3 (color online). Entanglement entropy at the center bond of a 40-site Heisenberg chain as a function of real time. Each solid line is for a single METTS, while the dashed lines are for an ancilla system with 40 sites and 40 ancilla.

$\mathcal{C}_\ell \equiv (\langle S_\ell^x \rangle^2 + \langle S_\ell^y \rangle^2 + \langle S_\ell^z \rangle^2)^{1/2}$. $\mathcal{C}_\ell$ measures how classical a spin is—how entangled it is within the METTS. For an isolated $S = 1/2$ in any pure state, $\mathcal{C}_\ell = 1/2$. Any total $S = 0$ wave function would give $\mathcal{C}_\ell = 0$ for every ℓ. The METTS are biased to be classical, which makes $\mathcal{C}_\ell$ meaningful. It is surprising how little variation there is in $\mathcal{C}_\ell$ from site to site.

The METTS can be evolved in real time (say, with real-time DMRG). The ensemble averages of METTS states are time-independent, but the METTS themselves are not. Much as a single particle with a narrowly peaked wave function would spread out in time, METTS evolve to states with much higher entanglement entropy. In Fig. 3, we show the growth of S with time for several different temperatures. In the higher temperature cases, the entropy starts smaller but grows more rapidly. The same effect is seen in the results for an ancilla system, for which the typical entropy is roughly twice that of the METTS, in agreement with the behavior of m, with $m \sim \exp(S)$. The behavior of the entropy as a function of time determines the effectiveness of the real-time DMRG to calculate finite temperature spectral functions [10]. Our results show that METTS will be able to reach longer times than the ancilla approach [10].

The rapid growth of entanglement with time for METTS raises the question of whether METTS (at $t = 0$) really are typical wave functions of real systems. The answer is very likely no; typical wave functions have more entanglement than METTS, with eventual entanglement growth perhaps limited by decoherence. One can evolve an ensemble of METTS to some fixed time t; the resulting set of states also satisfy Eq. (1), would exhibit more entanglement, and thus could be considered as being more realistic physically. However, the METTS themselves are more useful computationally.

We briefly note several other approaches to finite temperatures. A quite different (but powerful) finite temperature DMRG approach for infinite, translationally invariant 1D systems is the transfer matrix DMRG [11,12]. More closely related to our work are two approaches adapted for Lanczos calculations [13,14] and one recent DMRG approach [15]. METTS has a more rigorous foundation than these approaches, e.g., not dependent on any completeness properties of the Lanczos or DMRG basis and applicable to any temperature. The ground state DMRG can currently treat 2D clusters with fermions or frustration with about 100–200 sites. Because the computational effort with METTS is modest, of order (β/time step) times that of the ground state DMRG, we expect to be able to treat such systems at a finite temperature for the first time. METTS may also reveal short- or long-range order which was not expected and not looked for in correlation functions even at zero temperature.

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992); S. R. White, Phys. Rev. B **48**, 10 345 (1993); see also U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[2] Y. Nishio *et al.*, arXiv:cond-mat/0401115; F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.

[3] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).

[4] E. Schrödinger, *Statistical Thermodynamics* (Dover, New York, 1989) (orig. ed. 1952), pp. 5 and 89.

[5] R. P. Feynman, *Statistical Mechanics: A Set of Lectures* (Benjamin, Reading, MA, 1972).

[6] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004); see also M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).

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

[8] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003); S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004); A. J. Daley *et al.*, J. Stat. Mech. (2004) P04005.

[9] For a similar sequential determination of a CPS in a classical model, see K. Ueda *et al.*, J. Phys. Soc. Jpn. Suppl. **74**, 111 (2005).

[10] T. Barthel, U. Schollwöck, and S. R. White, arXiv:0901.2342.

[11] T. Nishino, J. Phys. Soc. Jpn. **64**, L3598 (1995); R. J. Bursill *et al.*, J. Phys. Condens. Matter **8**, L583 (1996); X. Wang and T. Xiang, Phys. Rev. B **56**, 5061 (1997).

[12] J. Sirker and A. Klümper, Phys. Rev. B **71**, 241101 (2005); J. Sirker, Phys. Rev. B **73**, 224424 (2006).

[13] J. Jaklič and P. Prelovšek, Phys. Rev. B **49**, 5065 (1994).

[14] M. W. Long *et al.*, Phys. Rev. B **68**, 235106 (2003).

[15] S. Sota and T. Tohyama, Phys. Rev. B **78**, 113101 (2008).