7 The Finite Temperature Lanczos Method and its Applications

Peter Prelovšek

J. Stefan Institute

Jamova 39, SI-1000 Ljubljana, Slovenia

Contents

1   Introduction

Models of strongly correlated systems have been one of the most intensively studied theoretical subjects in the last two decades, stimulated at first by the discovery of compounds superconducting at high-temperatures and ever since by the emergence of various novel materials and phenomena which could be traced back to strongly correlated electrons in these systems. Recently, cold atoms in optical lattices offer a different realization of strongly correlated quantum entities, whereby these systems can be even tuned closer to theoretical models.

One of the most straightforward methods to numerically deal with the lattice (discrete) models of correlated particles, which are inherently many-body (MB) quantum systems, is exact diagonalization (ED) of small-size systems. In view of the absence of well-controlled analytical methods, the ED method has been employed intensively to obtain results for static and dynamical properties of various models with different aims: a) to search and confirm novel phenomena specific to strongly correlated systems, b) to test theoretical ideas and analytical results, c) to get reference results for more advanced numerical techniques.

MB quantum lattice models of interacting particles are characterized by the dimension of the Hilbert space given by the number of basis states $N_{st} \propto K^N$ that is in turn exponentially increasing with the lattice size N, where K is the number of local quantum states. It is therefore clear that ED methods can treat fully only systems with limited N_{st}, i.e., both K and N must be quite modest.

Among the ED approaches the full ED within the Hilbert space of the model Hamiltonian, yielding all eigenenergies and eigenfunctions, is the simplest to understand, most transparent, and easy to implement. In principle it allows the evaluation of any ground state (g.s.) property as well as finite temperature $T > 0$ static or dynamic quantities, at the expense of a very restricted N_{st}. In spite of that, it represents a very instructive approach and also remains essentially the only practical method when all exact levels are needed, e.g., for studies of level statistics.

Lanczos-based ED methods have already a long history of applications since Cornelius Lanczos [1] proposed the diagonalization of sparse matrices using the iterative procedure, allowing for much bigger Hilbert spaces N_{st} relative to full ED. The Lanczos diagonalization technique is at present a part of standard numerical linear algebra procedures [2, 3] and as such in solid state physics mainly used to obtain the g.s. energy and wavefunction and the corresponding expectation values. The approach has been quite early-on extended to calculation of the dynamical response functions within the g.s. [4]. The method has been in the last 20 years extensively used in connection with models related to high-T_c materials, for which we can refer to an earlier overview [5].

Here we focus on recent developments of ED-based and Lanczos-based methods. The basics of the Lanczos method are presented in Sec. 2 and its application for g.s. properties in Sec. 3. One of the already established generalizations is the finite-temperature Lanczos method (FTLM) [6–8], reviewed in Sec. 4, which allows for the evaluation of $T > 0$ static and dynamic properties within simplest models. Several extensions and modifications of the latter have been introduced more recently, in particular the low-temperature Lanczos method (LTLM) [9] and

the microcanonical Lanczos method (MCLM) [10], particularly applicable within the high-T regime. Recently, there is also quite an intensive activity on studies of real-time evolution of correlated systems, both under equilibrium and non-equilibrium conditions that can be simulated using the ED and Lanczos-based methods, as discussed in Sec. 5.

2   Exact diagonalization and Lanczos method

2.1   Models, geometries, and system sizes

ED-based methods are mostly restricted to simple models with only few local quantum states K per lattice site in order to reach reasonable system sizes N. Consequently, there are only few classes of MB models that so far exhaust the majority of ED and Lanczos-method studies, clearly also motivated and influenced by the challenging physics and relevance to novel materials and related experiments.

To get some feeling for the available sizes reachable within ED-based approaches, it should be kept in mind that in full ED routines the CPU time scales with the number of operations $Op \propto N_{st}^3$, while the memory requirement is related to the storage of the whole Hamiltonian matrix and all eigenvectors, i.e., $Mem \propto N_{st}^2$. This limits, at the present stage of computer facilities, the full ED method to $N_{st} < 2. \, 10^4$ MB states. On the other hand, using Lanczos-based iterative methods for the diagonalization of sparse matrices (Hamiltonians), CPU and memory requirements scale as $Op, Mem \propto N_{st}$, at least in their basic application, to calculate the g.s. and its wavefunction. In present-day applications this allows the consideration of much larger basis sets, i.e., $N_{st} < 10^9$. Still, lattice sizes N reached using the Lanczos technique remain rather modest, compared to some other numerical approaches such as DMRG and quantum-Monte-Carlo QMC methods, if the full Hilbert basis space relevant for the model is used.

The simplest nontrivial class of MB lattice models are spin models, the prototype being the anisotropic Heisenberg model for coupled $S = 1/2$ spins,

$$H = \sum_{\langle ij \rangle \alpha} J_{ij}^{\alpha\alpha} S_i^\alpha S_j^\alpha, \tag{1}$$

where the sum $\langle ij \rangle$ runs over pairs of lattice sites with an arbitrary interaction $J_{ij}^{\alpha\alpha}$ (being in principle anisotropic) and S_i^α are the components of the local $S = 1/2$ operator. The model has just $K = 2$ quantum states per lattice site and therefore allows for biggest possible N in the ED-based approaches, where $N_{st} \propto 2^N$ basis states. To reduce N_{st} as many symmetries and good quantum numbers as practically possible are used to decompose the Hamiltonian into separate blocks. Evident choices are sectors with the (z-component of) total spin S_{tot}^z and the wavevector **q** for systems with periodic boundary conditions, but also rotational symmetries of particular lattices have been used. In this way system sizes up to $N \sim 36$ (for the largest and most interesting sector $S_{tot}^z = 0$) have been reached so far using the Lanczos technique without any basis reduction.

On the basis of this simple model one can already discuss the feasibility of Lanczos-based methods with respect to other numerical quantum MB methods. For the g.s. of 1D spin systems

more powerful methods allowing for much bigger systems are DMRG and related approaches. For unfrustrated models in D > 1 QMC methods are superior for the evaluation of static quantities at any T. Still, Lanczos-based methods become competitive or at least are not superseded for frustrated spin models (where QMC can run into the minus-sign problem) or for dynamical properties at $T > 0$.

Next in complexity and very intensively studied is the t-J model, representing strongly correlated itinerant electrons with an antiferromagnetic (AFM) interaction between their spins

$$H = -\sum_{\langle ij \rangle s} \left(t_{ij}\, \tilde{c}^\dagger_{js} \tilde{c}_{is} + \text{H.c.} \right) + J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j \,, \tag{2}$$

where due to the strong on-site repulsion doubly occupied sites are forbidden and one is dealing with projected fermion operators $\tilde{c}_{is} = c_{is}(1 - n_{i,-s})$. The model can be considered as a good microscopic model for superconducting cuprates, which are doped Mott insulators, and has therefore been one of the most studied models using the Lanczos method [5]. For a theoretical and experimental overview of Mott insulators and metal-insulator transitions see Ref. [11]. It has $K = 3$ quantum states per lattice site and, besides S^z_{tot} and $\mathbf{q}$, also the number of electrons N_e (or more appropriately the number of holes $N_h = N - N_e$) are the simplest quantum numbers to implement. Since the model reveals an interesting physics in D > 1, the effort was in connection with high-T_c cuprates mostly concentrated on the 2D square lattice. Here the alternative numerical methods have more drawbacks (e.g., the minus sign problem in QMC methods due to the itinerant character of the fermions) so that Lanczos-based methods are still competitive, in particular for getting information on $T > 0$ dynamics and transport. The largest systems considered with the Lanczos method so far are 2D square lattices with $N = 32$ sites and $N_h = 4$ holes [12].

Clearly, one of the most investigated problems within the MB community is the standard single-band Hubbard model, which has $K = 4$ states per lattice site. Due to the complexity $N_{st} \propto 4^N$ the application of ED and Lanczos-based method is already quite restricted reaching so far $N = 20$ sites [13] requiring already $N_{st} \sim 10^9$ basis states. The model is also the subject of numerous studies using more powerful QMC method and various cluster dynamical-mean-field-theory (DMFT) methods for much larger lattices so Lanczos-based approaches have here more specific goals.

Since reachable lattices sizes for the above mentioned models are rather small it is important to properly choose their geometries. This is not a problem for 1D models, but becomes already essential for 2D lattices, analyzed in connection with novel materials, in particular high-T_c cuprates and related materials. In order to keep periodic boundary conditions for 2D square lattices the choice of Pythagorean lattices with $N = \lambda_x^2 + \lambda_y^2$ with λ_x, λ_y [14] has significantly extended available sizes. Some frequently used ones are presented in Fig. 1. Taking into account only even N, such lattices include $N = 8, 10, 16, 18, 20, 26, 32,$ and 36 sites. While the unit cells of such lattices are squares, it has been observed that they are not always optimal with respect to the number of next-nearest and further nearest neighbors. It has been claimed and partly tested that better result are obtained with slightly deformed lattices (still with periodic boundary conditions) which at the same time offer an even larger choice of sizes [15].
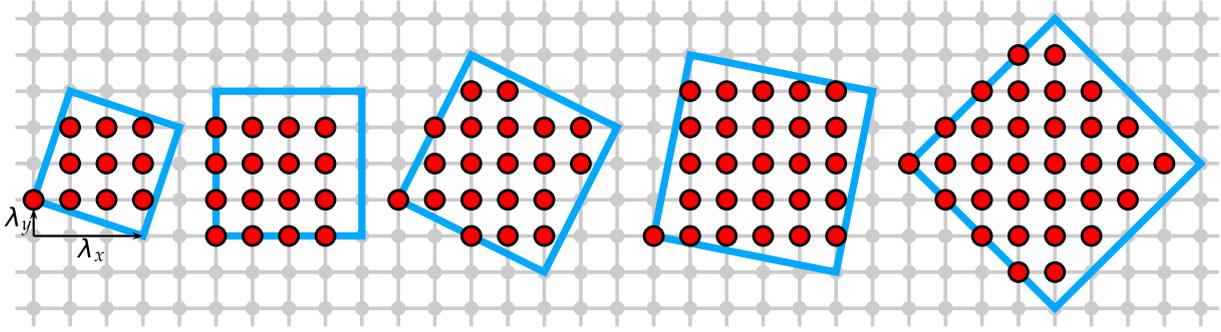
Fig. 1: *Some tilted clusters used in 2D square-lattice studies*

2.2 Lanczos diagonalization technique

The Lanczos technique is a general procedure to transform and reduce a symmetric $N_{st} \times N_{st}$ matrix A to a symmetric $M \times M$ tridiagonal matrix T_M. From the chosen initial N_{st}-dimensional vector $\mathbf{v_1}$ one generates an orthogonal basis of $\{\mathbf{v_1}, \cdots, \mathbf{v_M}\}$ vectors which span the Krylov space $\{\mathbf{v_1}, \mathbf{Av_1}, \cdots, \mathbf{A^{M-1}v_1}\}$ [1–3, 16].

In usual applications for the quantum MB system defined with Hamiltonian H the Lanczos algorithm starts with a normalized vector $|\phi_0\rangle$, chosen as a random vector in the relevant Hilbert space with N_{st} basis states. The procedure generates orthogonal Lanczos vectors $L_M = \{|\phi_m\rangle \mid m = 0 \ldots M\}$ spanning the Krylov space $\{|\phi_0\rangle, H|\phi_0\rangle, \cdots, H^M|\phi_0\rangle\}$. Steps are as follows: H is applied to $|\phi_0\rangle$ and the resulting vector is split in components parallel to $|\phi_0\rangle$, and normalized $|\phi_1\rangle$ orthogonal to it, respectively,

$$H|\phi_0\rangle = a_0|\phi_0\rangle + b_1|\phi_1\rangle. \tag{3}$$

Since H is Hermitian, $a_0 = \langle\phi_0|H|\phi_0\rangle$ is real, while the phase of $|\phi_1\rangle$ can be chosen so that b_1 is also real. In the next step H is applied to $|\phi_1\rangle$,

$$H|\phi_1\rangle = b_1'|\phi_0\rangle + a_1|\phi_1\rangle + b_2|\phi_2\rangle, \tag{4}$$

where $|\phi_2\rangle$ is orthogonal to $|\phi_0\rangle$ and $|\phi_1\rangle$. It follows also $b_1' = \langle\phi_0|H|\phi_1\rangle = b_1$. Proceeding with the iteration one gets in i steps

$$H|\phi_i\rangle = b_i|\phi_{i-1}\rangle + a_i|\phi_i\rangle + b_{i+1}|\phi_{i+1}\rangle, \qquad 1 \leq i \leq M, \tag{5}$$

where in Eq. (5) by construction there are no terms involving $|\phi_{i-2}\rangle$ etc. By stopping the iteration at $i = M$ and setting $b_{M+1} = 0$, the Hamiltonian can be represented in the basis of orthogonal Lanczos functions $|\phi_i\rangle$ as the tridiagonal matrix H_M with diagonal elements a_i, $i = 0 \ldots M$, and off-diagonal ones b_i, $i = 1 \ldots M$. Such a matrix is easily diagonalized using standard numerical routines to obtain approximate eigenvalues ε_j and corresponding orthonormal eigenvectors $|\psi_j\rangle$,

$$|\psi_j\rangle = \sum_{i=0}^{M} v_{ji}|\phi_i\rangle, \qquad j = 0 \ldots M. \tag{6}$$

It is important to realize that $|\psi_j\rangle$ are (in general) not exact eigenfunctions of H, but show a remainder. On the other hand, it is evident from the diagonalization of H_M that matrix elements

$$\langle \psi_i | H | \psi_j \rangle = \varepsilon_j \delta_{ij}, \qquad i,j = 0 \ldots M, \tag{7}$$

are diagonal independently of L_M (but provided $i, j \le M$), although the values ε_j can be only approximate.

If in the equation (5) $b_{M+1} = 0$, we have found an $(M+1)$-dimensional eigenspace where H_M is already an exact representation of H. This inevitably happens when $M = N_{st} - 1$, but for $M < N_{st} - 1$ it can only occur if the starting vector is orthogonal to some invariant subspace of H which we avoid by choosing the input vector $|\phi_0\rangle$ as a random one.

It should be recognized that the Lanczos approach is effective only for sparse Hamiltonians, characterized by the connectivity of each basis state with $K_n \ll N_{st}$ basis states. All prototype discrete tight-binding models discussed in Sec. 2.1 are indeed of such a type in the local MB basis. Estimating the computation requirements, the number of operations Op needed to perform M Lanczos iterations scales as $Op \propto K_n M N_{st}$. The main restriction is still in memory requirements due to the large N_{st}. A straightforward application of Eq. (5) would require the fast storage of all $|\phi_i\rangle$, $i = 0 \ldots M$, i.e., also the memory capacity $Mem \propto M N_{st}$. However, for the evaluation of the eigenvalues alone during the iteration, Eq. (5), only three $|\phi_i\rangle$ are successively required, so this leads to $Mem \propto 3 N_{st}$. If the Hamiltonian matrix is not evaluated on the fly, then also $Mem \propto K_n N_{st}$ for the nonzero Hamilton matrix elements is needed.

The Lanczos diagonalization is in essence an iterative power method which is known to converge fast for the extreme lower and upper eigenvalues [2, 3]. In physical application most relevant is the search for the g.s. energy E_0 and the corresponding wavefunction $|\Psi_0\rangle$. Typically, $M > 50$ are enough to reach very high accuracy for both. It is evident that for such modest $M \ll N_{st}$ one cannot expect any reliable results for eigenstates beyond the few at the bottom and the top of the spectrum. On the other hand, the Lanczos procedure is subject to roundoff errors, introduced by the finite-precision arithmetics which usually only becomes severe at larger $M > 100$ after the convergence of extreme eigenvalues, and is seen as the loss of orthogonality of the vectors $|\phi_i\rangle$. It can be remedied by successive reorthogonalization [2, 3, 16] of new states $|\phi_i'\rangle$, plagued with errors, with respect to previous ones. However this procedure requires $Op \sim M^2 N_{st}$ operations, and can become computationally more demanding than the Lanczos iterations themselves. This effect also prevents one from using the Lanczos method, e.g., to efficiently tridiagonalize large dense matrices [3].

3   Ground state properties and dynamics

After $|\Psi_0\rangle$ is obtained, the static properties of the g.s. can be evaluated in principle for any operator A as

$$\bar{A}_0 = \langle \Psi_0 | A | \Psi_0 \rangle. \tag{8}$$

Clearly, the procedure (8) is effective for large a basis only if the operator A is sparse in the same basis, as is the case for most operators of interest.

It is an important advantage of the Lanczos procedure that dynamical g.s. functions can easily be calculated [4]. Let us consider the dynamical (autocorrelation) response function

$$C(\omega) = \langle \Psi_0 | A^\dagger \frac{1}{\omega^+ + E_0 - H} A | \Psi_0 \rangle \tag{9}$$

for the observable given by the operator A, where $\omega^+ = \omega + i\delta$ with $\delta > 0$. To calculate $C(\omega)$ one has to run a second Lanczos procedure with a new initial function $|\tilde{\phi}_0\rangle$,

$$|\tilde{\phi}_0\rangle = \frac{1}{\alpha} A | \Psi_0 \rangle, \qquad \alpha = \sqrt{\langle \Psi_0 | A^\dagger A | \Psi_0 \rangle}. \tag{10}$$

Starting with $|\tilde{\phi}_0\rangle$ one generates another Lanczos subspace $\tilde{L}_{\tilde{M}} = \{|\tilde{\phi}_j\rangle, \ j = 0, \tilde{M}\}$ with (approximate) eigenvectors $|\tilde{\psi}_j\rangle$ and eigenenergies $\tilde{\varepsilon}_j$. The matrix for H in the new basis is again a tridiagonal one with $\tilde{a}_j$ and $\tilde{b}_j$ elements, respectively. Terminating the Lanczos procedure at a given $\tilde{M}$, one can evaluate Eq. (9) as a resolvent of the $H_{\tilde{M}}$ matrix expressed in the continued-fraction form [17, 4, 5],

$$C(\omega) = \cfrac{\alpha^2}{\omega^+ + E_0 - \tilde{a}_0 - \cfrac{\tilde{b}_1^2}{\omega^+ + E_0 - \tilde{a}_1 - \cfrac{\tilde{b}_2^2}{\omega^+ + E_0 - \tilde{a}_2 - \dots}}}, \tag{11}$$

terminating with $\tilde{b}_{\tilde{M}+1} = 0$, although other termination functions can also be employed and can be well justified.

We note that frequency moments of the spectral function

$$\mu_l = -\frac{1}{\pi} \int_{-\infty}^{\infty} \omega^l \operatorname{Im} C(\omega) \, d\omega = \langle \Psi_0 | A^\dagger (H - E_0)^l A | \Psi_0 \rangle = \alpha^2 \langle \tilde{\phi}_0 | (H - E_0)^l | \tilde{\phi}_0 \rangle \tag{12}$$

are exact for given $|\Psi_0\rangle$ provided $l \leq \tilde{M}$, since the operator H^l, $l < \tilde{M}$, is exactly reproduced within the Lanczos (or corresponding Krylov) space $\tilde{L}_{\tilde{M}}$.

Finally, $C(\omega)$ (11) can be presented as a sum of $j = 0, \tilde{M}$ poles at $\omega = \tilde{\varepsilon}_j - E_0$ with corresponding weights w_j. As a practical matter we note that in analogy to Eq. (6)

$$w_j = |\langle \tilde{\psi}_j | A | \Psi_0 \rangle|^2 = \alpha^2 |\langle \tilde{\psi}_j | \tilde{\phi}_0 \rangle|^2 = \alpha^2 \tilde{v}_{j0}^2, \tag{13}$$

hence no matrix elements need to be evaluated within this approach. In contrast to the autocorrelation function (11), the procedure allows also the treatment of general correlation functions $C_{AB}(\omega)$, with $B \neq A^\dagger$. In this case matrix elements $\langle \Psi_0 | B | \tilde{\psi}_j \rangle$ have to be evaluated explicitly. It should be also mentioned that at least the lowest poles of $C(\omega)$, Eq. (11), should coincide with eigenenergies $\omega = E_i - E_0$ if $|\tilde{\phi}_0\rangle$ is not orthogonal to $|\Psi_0\rangle$. However, using $\tilde{M} > 50$, spurious poles can emerge (if no reorthogonalization is used) which, however, carry no weight as is evident from exact moments (12).

In this chapter we do not intend to present an overview of applications of the full ED and Lanczos-type studies of g.s. static and dynamical properties of correlated systems. There have been numerous such investigations even before the high-T_c era, intensified strongly with studies of prototype models relevant for high-T_c cuprates [5] and other novel materials with correlated electrons. Although a variety of models has been investigated they are still quite restricted in the number of local degrees and sizes.

4 Static properties and dynamics at $T > 0$

Before describing the finite temperature Lanczos method (FTLM) we should note that the Lanczos basis is a very useful and natural basis for evaluating matrix elements of the type

$$W_{kl} = \langle n|H^k B H^l A|n\rangle, \tag{14}$$

where $|n\rangle$ is an arbitrary normalized vector, and A, B are general operators. One can calculate this expression exactly by performing two Lanczos procedures with $M = \max(k, l)$ steps. The first one, starting with the vector $|\phi_0\rangle = |n\rangle$, produces the Lanczos basis L_M along with approximate eigenstates $|\psi_j\rangle$ and ε_j. The second Lanczos procedure is started with the normalized vector $|\tilde{\phi}_0\rangle \propto A|\phi_0\rangle = A|n\rangle$, Eq. (10), and generates $\tilde{L}_M$ with corresponding $|\tilde{\psi}_j\rangle$ and $\tilde{\varepsilon}_j$. We can now define projectors onto limited subspaces

$$P_M = \sum_{i=0}^{M} |\psi_i\rangle\langle\psi_i|, \quad \tilde{P}_M = \sum_{i=0}^{M} |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|. \tag{15}$$

Provided that $(l, k) < M$ projectors P_M and $\tilde{P}_M$ span the whole relevant basis for the operators H^k and H^l, respectively, so that one can rewrite W_{kl} in Eq. (14) as

$$W_{kl} = \langle\phi_0|P_M H P_M H \dots H P_M B \tilde{P}_M H \dots \tilde{P}_M H \tilde{P}_M A|\phi_0\rangle. \tag{16}$$

Since H is diagonal in the basis $|\psi_j\rangle$ and $|\tilde{\psi}_j\rangle$, respectively, one can write finally

$$W_{kl} = \sum_{i=0}^{M} \sum_{j=0}^{M} \langle\phi_0|\psi_i\rangle\langle\psi_i|B|\tilde{\psi}_j\rangle\langle\tilde{\psi}_j|A|\phi_0\rangle \ (\varepsilon_i)^k (\tilde{\varepsilon}_j)^l. \tag{17}$$

It is important to note that expression (17) for the matrix element is exact, independently of how (in)accurate the representation $|\psi_i\rangle$, ε_i and $|\tilde{\psi}_j\rangle$, ε_j, respectively, are for true system eigenvalues. The only condition is that number of Lanczos steps is sufficient, i.e., $M > (l, k)$.

4.1 Finite-temperature Lanczos method: Static quantities

A straightforward calculation of the canonical thermodynamic average of an operator A at finite temperature $T > 0$ (in a finite system) requires the knowledge of all eigenstates $|\Psi_n\rangle$ and corresponding energies E_n, obtained, e.g., by the full ED of H

$$\langle A\rangle = \sum_{n=1}^{N_{st}} e^{-\beta E_n} \langle\Psi_n|A|\Psi_n\rangle \ \Big/ \ \sum_{n=1}^{N_{st}} e^{-\beta E_n}, \tag{18}$$

where $\beta = 1/k_B T$. Such a direct evaluation is both CPU time and storage demanding for larger systems and is at present accessible only for $N_{st} \sim 20000$.

In a general orthonormal basis $|n\rangle$ for finite system with N_{st} basis states one can express the canonical expectation value $\langle A\rangle$ as

$$\langle A\rangle = \sum_{n=1}^{N_{st}} \langle n|e^{-\beta H} A|n\rangle \ \Big/ \ \sum_{n=1}^{N_{st}} \langle n|e^{-\beta H}|n\rangle, \tag{19}$$

The FTLM for $T > 0$ is based on the evaluation of the expectation value in Eq. (19) for each starting $|n\rangle$ using the Lanczos basis. We note that such a procedure guarantees the correct high-T expansion series (for given finite system) to high order. Let us perform the high-T expansion of Eq. (19),

$$
\begin{aligned}
\langle A \rangle &= Z^{-1} \sum_{n=1}^{N_{st}} \sum_{k=0}^{\infty} \frac{(-\beta)^k}{k!} \langle n|H^k A|n\rangle, \\
Z &= \sum_{n=1}^{N_{st}} \sum_{k=0}^{\infty} \frac{(-\beta)^k}{k!} \langle n|H^k|n\rangle.
\end{aligned}
\tag{20}
$$

Terms in the expansion $\langle n|H^k A|n\rangle$ can be calculated exactly using the Lanczos procedure with $M \geq k$ steps (using $|\phi_0^n\rangle = |n\rangle$ as the starting function) since this is a special case of the expression (14). Using relation (17) with $l = 0$ and $B = 1$, we get

$$
\langle n|H^k A|n\rangle = \sum_{i=0}^{M} \langle n|\psi_i^n\rangle\langle\psi_i^n|A|n\rangle (\varepsilon_i^n)^k.
\tag{21}
$$

Working in a restricted basis $k \leq M$, we can insert the expression (21) into sums (20), extending them to $k > M$. The final result can be expressed as

$$
\begin{aligned}
\langle A \rangle &= Z^{-1} \sum_{n=1}^{N_{st}} \sum_{i=0}^{M} e^{-\beta\varepsilon_i^n} \langle n|\psi_i^n\rangle\langle\psi_i^n|A|n\rangle, \\
Z &= \sum_{n=1}^{N_{st}} \sum_{i=0}^{M} e^{-\beta\varepsilon_i^n} \langle n|\psi_i^n\rangle\langle\psi_i^n|n\rangle,
\end{aligned}
\tag{22}
$$

and the error of the approximation is $O(\beta^{M+1})$.

Evidently, within a finite system Eq. (22), expanded as a series in β, reproduces exactly the high-T series to the order M. In addition, in contrast to the usual high-T expansion, Eq. (22) remains accurate also for $T \to 0$. Let us assume for simplicity that the g.s. $|\Psi_0\rangle$ is nondegenerate. For initial states $|n\rangle$ not orthogonal to $|\Psi_0\rangle$, already at modest $M \sim 50$ the lowest eigenstate $|\psi_0^n\rangle$ converges to $|\Psi_0\rangle$. We thus have for $\beta \to \infty$,

$$
\langle A \rangle = \sum_{n=1}^{N_{st}} \langle n|\Psi_0\rangle\langle\Psi_0|A|n\rangle \Big/ \sum_{n=1}^{N_{st}} \langle n|\Psi_0\rangle\langle\Psi_0|n\rangle = \langle\Psi_0|A|\Psi_0\rangle/\langle\Psi_0|\Psi_0\rangle,
\tag{23}
$$

where we have taken into account the completeness of the set $|n\rangle$. Thus we obtain just the usual g.s. expectation value of operator A.

The computation of static quantities (22) still involves the summation over the complete set of N_{st} states $|n\rangle$, which is clearly not feasible in practice. To obtain a useful method, a further essential approximation replaces the full summation over $|n\rangle$ by a partial one over a much smaller set of random states [18, 19]. Such an approximation is analogous to Monte Carlo methods and leads to a statistical error which can be well estimated and is generally quite small.

Let us first consider only the expectation value (19) with respect to a single random state $|r\rangle$, which is a linear combination of basis states

$$|r\rangle = \sum_{n=1}^{N_{st}} \eta_{rn} |n\rangle, \tag{24}$$

where the η_{rn} are assumed to be distributed randomly. Then the random quantity can be expressed as

$$\tilde{A}_r = \frac{\langle r|e^{-\beta H} A|r\rangle}{\langle r|e^{-\beta H}|r\rangle} = \sum_{n,m=1}^{N_{st}} \eta_{rn}^* \eta_{rm} \langle n|e^{-\beta H} A|m\rangle \bigg/ \sum_{n,m=1}^{N_{st}} \eta_{rn}^* \eta_{rm} \langle n|e^{-\beta H}|m\rangle. \tag{25}$$

Assuming that due to the random sign (phase), offdiagonal terms with $\eta_{rn}^* \eta_{rm}$, $m \neq n$ cancel on average for large N_{st}, we remain with

$$\bar{A}_r = \sum_{n=1}^{N_{st}} |\eta_{rn}|^2 \langle n|e^{-\beta H} A|n\rangle \bigg/ \sum_{n=1}^{N_{st}} |\eta_{rn}|^2 \langle n|e^{-\beta H}|n\rangle. \tag{26}$$

We can express $|\eta_{rn}|^2 = 1/N_{st} + \delta_{rn}$. The random deviations δ_{rn} should not be correlated with matrix elements $\langle n|e^{-\beta H}|n\rangle = Z_n$ and $\langle n|e^{-\beta H} A|n\rangle = Z_n A_n$, therefore $\bar{A}_r$ is close to $\langle A\rangle$ with an statistical error related to the effective number of terms $\bar{Z}$ in the thermodynamic sum, i.e.

$$\bar{A}_r = \langle A\rangle (1 + \mathcal{O}(1/\sqrt{\bar{Z}})), \tag{27}$$

$$\bar{Z} = e^{\beta E_0} \sum_n Z_n = \sum_{n=1}^{N_{st}} \langle n|e^{-\beta(H-E_0)}|n\rangle. \tag{28}$$

Note that for $T \to \infty$ we have $\bar{Z} \to N_{st}$ and therefore at large N_{st} a very accurate average (28) can be obtained even from a single random state [18, 19]. On the other hand, at finite $T < \infty$ the statistical error of $\tilde{A}_r$ increases with decreasing $\bar{Z}$.

To reduce statistical errors, in particular at modest $T > 0$, within the FTLM we sum in addition over R different randomly chosen $|r\rangle$, so that in the final application Eq. (22) leads to

$$\langle A\rangle = \frac{N_{st}}{ZR} \sum_{r=1}^{R} \sum_{j=0}^{M} e^{-\beta \varepsilon_j^r} \langle r|\psi_j^r\rangle \langle \psi_j^r|A|r\rangle,$$

$$Z = \frac{N_{st}}{R} \sum_{r=1}^{R} \sum_{j=0}^{M} e^{-\beta \varepsilon_j^r} |\langle r|\psi_j^r\rangle|^2. \tag{29}$$

Random states $|r\rangle = |\phi_0^r\rangle$ serve as initial functions for the Lanczos iteration, resulting in M eigenvalues ε_j^r with corresponding $|\psi_j^r\rangle$. The relative statistical error is reduced by sampling (both for $\langle A\rangle$ and Z) and behaves as

$$\delta\langle A\rangle / \langle A\rangle = \mathcal{O}(1/\sqrt{R\bar{Z}}). \tag{30}$$

For a general operator A the calculation of $|\psi_j^r\rangle$ and the corresponding matrix elements $\langle\psi_j^r|A|r\rangle$ is needed. On the other hand, the computational effort is significantly reduced if A is a conserved quantity, i.e., $[H, A] = 0$, and can be diagonalized simultaneously with H. Then

$$\langle A \rangle = \frac{N_{st}}{ZR} \sum_{r=1}^{R} \sum_{j=0}^{M} e^{-\beta\varepsilon_j^r} |\langle r|\psi_j^r\rangle|^2 A_j^r. \tag{31}$$

In this case the evaluation of eigenfunctions is not necessary since the element $\langle r|\psi_j^r\rangle = v_{j0}^r$, Eq. (6), is obtained directly from the eigenvectors of the tridiagonal matrix H_M^r. There are several quantities of interest which can be evaluated in this way, in particular thermodynamic properties such as internal energy, specific heat, entropy, as well as uniform susceptibilities etc. [7,8].

Taking into account all mentioned assumptions, the approximation $\langle A \rangle$ (29) yields a good estimate of the thermodynamic average at all T. For low T the error is expected to be of the order of $\mathcal{O}(1/\sqrt{R})$, while for high T the error is expected to scale even as $\mathcal{O}(1/\sqrt{N_{st}R})$. Since arguments leading to these estimates are not always easy to verify, it is essential to test the method for particular cases.

4.2 Finite-temperature Lanczos method: Dynamical response

The essential advantage of the FTLM with respect to other methods is in the calculation of dynamical quantities. Let us consider the dynamical susceptibility as given by the autocorrelation function $C(\omega)$ (the procedure for a general correlation function $C_{AB}(\omega)$ is given in Ref. [7]),

$$\chi''(\omega) = \pi(1 - e^{-\beta\omega})C(\omega), \qquad C(\omega) = \frac{1}{\pi}\mathrm{Re} \int_0^{+\infty} dt\, e^{i\omega t} C(t), \tag{32}$$

with

$$C(t) = \langle A^\dagger(t) A(0) \rangle = \frac{1}{Z} \sum_n \langle n|e^{(-\beta+it)H} A^\dagger e^{-iHt} A|n\rangle. \tag{33}$$

Expanding the exponentials in analogy to static quantities, Eq. (20), we get

$$C(t) = Z^{-1} \sum_{n=1}^{N_{st}} \sum_{k,l=0}^{\infty} \frac{(-\beta+it)^k}{k!} \frac{(-it)^l}{l!} \langle n|H^k A^\dagger H^l A|n\rangle. \tag{34}$$

The expansion coefficients in Eq. (34) can be again obtained via the Lanczos method, as discussed in Sec. 4.1. Performing two Lanczos iterations with M steps, starting from the normalized vectors $|\phi_0^n\rangle = |n\rangle$ and $|\tilde{\phi}_0^n\rangle \propto A|n\rangle$, respectively, we calculate the coefficients W_{kl} following equation (17). We again note that (within the full basis $|n\rangle$) the series are, via the W_{kl}, exactly evaluated within the Lanczos basis up to order $l, k \leq M$. The latter yields through Eq. (34) a combination of (β, t) expansion, i.e., a combination of a high-T and short-t (in frequency high-ω) expansion to very high order. Extending and resumming the series in k and l into exponentials, we get in analogy with Eq. (22)

$$C(t) = Z^{-1} \sum_{n=1}^{N_{st}} \sum_{i,j=0}^{M} e^{-\beta\varepsilon_i^n} e^{it(\varepsilon_i^n - \tilde{\varepsilon}_j^n)} \langle n|\psi_i^n\rangle \langle\psi_i^n|A^\dagger|\tilde{\psi}_j^n\rangle \langle\tilde{\psi}_j^n|A|n\rangle. \tag{35}$$

Finally replacing the full summation with the random sampling from the FTLM recipe for the correlation function, we obtain

$$C(\omega) = \frac{N_{st}}{ZR} \sum_{r=1}^{R} \sum_{i,j=1}^{M} e^{-\beta\varepsilon_i} \langle r|\psi_i^r\rangle \langle\psi_i^r|A^\dagger|\tilde{\psi}_j^r\rangle \langle\tilde{\psi}_j^r|r\rangle \delta(\omega - \tilde{\varepsilon}_j^r + \varepsilon_i^r). \tag{36}$$

We check the nontrivial $T = 0$ limit of the above expression. If the $|n\rangle$ are not orthogonal to the g.s., $|\Psi_0\rangle$, then for large enough M the lowest-lying state converges to $\varepsilon_0^n \sim E_0$ and $|\psi_0^n\rangle \sim |\Psi_0\rangle$, respectively. In this case we have

$$C(\omega, T = 0) \approx \frac{N_{st}}{R} \sum_{r=1}^{R} \sum_{j=0}^{M} \langle\Psi_0|A^\dagger|\tilde{\psi}_j^n\rangle \langle\tilde{\psi}_j^n|A|r\rangle \langle r|\Psi_0\rangle \delta(\omega + E_0 - \tilde{\varepsilon}_j^n) \tag{37}$$

At $T \sim 0$ one needs in general $M \gg 100$ so that at least the low-lying states relevant to Eq. (37) approach $|\tilde{\psi}_j^n\rangle \to |\Psi_j\rangle$ and $\tilde{\varepsilon}_j^n \to E_j$. Also a considerable number of samples $R > 1$ is required to get correct also amplitudes of separate peaks in the spectrum of Eq. (37), which are a subject to statistical errors due to the incomplete projection on the different random $|r\rangle$ in $\langle\tilde{\psi}_j^n|A|r\rangle\langle r|\Psi_0\rangle$. Similar statistical errors can in fact appear also for static quantities in Eq. (29).

4.3 Finite temperature Lanczos method: Implementation

Most straightforward is the implementation of the FTLM for static quantities, Eq. (29). In particular for conserved quantities, Eq. (31), the computation load is essentially that of the g.s. Lanczos iteration, repeated R times, and only minor changes are needed within the usual g.s. Lanczos code.

On the other hand, for the dynamical correlation function (36) the memory requirement as well as the CPU time is dominated mostly by the evaluation of the matrix element $\langle\psi_i^r|A^\dagger|\tilde{\psi}_j^r\rangle$ where the operations scale as $Op \propto RM^2N_{st}$ and memory as $Mem \propto MN_{st}$. This effectively limits the application of the FTLM to $50 < M < 500$ where the lower bound is determined by the convergence of the g.s. $|\Psi_0\rangle$. Still, it should be noted that the calculation can be done simultaneously (without any additional cost) for all desired T, since matrix elements are evaluated only once. Evidently, one should use as much as possible symmetries of the Hamiltonian, e.g., N_e, S_{tot}^z, q to reduce the effective N_{st} by splitting the sampling over different symmetry sectors.

The effect of finite M is less evident. Since $M \sim 100$ is enough to converge well a few lowest levels, it is also generally satisfactory for reliable dynamical correlation functions at low T. At high T, however, one can observe very regular oscillations which are an artifact of the Lanczos iterations with $M \ll N_{st}$. Namely, the procedure generates between the extreme eigenvalues a spectrum of quasi-states with quite equidistant level spacing $\Delta\varepsilon \sim \Delta E/M$, where ΔE is the full energy span of MB eigenstates. The effect is well visible in Fig. 2 where the high-T result for the spin structure factor $S(q = \pi, \omega)$ os the 1D Heisenberg model, Eq. (1), is presented for various M. It is evident that for the presented case ($N = 24$ and $\Delta E \sim 16J$) $M > 200$ is sufficient to obtain smooth spectra even for high $T \gg J$. However, larger M are advisable if sharper structures persist at high T.
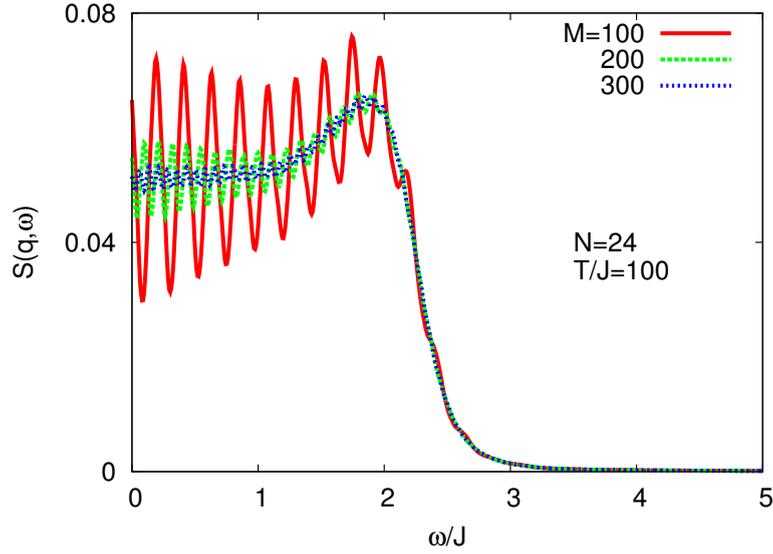
Fig. 2: *High-T spin structure factor $S(q = \pi, \omega)$ for the 1D Heisenberg model, as calculated with different numbers of Lanczos steps M.*

The role of random sampling R is less important for intermediate and high T since the relative error is largely determined via $\bar{Z}$ as evident from Eq. (28). Larger $R \gg 1$ is necessary only for the correct limit $T \to 0$ (for given system size) and for off-diagonal operators A.

One can claim that the FTLM in general obtains for all reachable systems results which are at any T very close to exact (full ED) results for the same finite (given N) system and the accuracy can be improved by increasing M and R. Still, it remains nontrivial but crucial to understand and have in control finite size effects.

At $T = 0$ both static and dynamical quantities are calculated from the g.s. $|\Psi_0\rangle$, which can be quite dependent on the size and the shape of the system. At least in 1D for static quantities the finite-size scaling $N \to \infty$ can be performed in a controlled way, although in this case more powerful methods as, e.g., DMRG are mostly available. In higher dimensional lattices, e.g., in 2D systems, finite-size scaling is practically impossible due to the very restricted choice of small sizes and different shapes. Also g.s. ($T = 0$) dynamical quantities are often dominated by few (typically $N_p < M$) peaks which are finite-size dependent [5]. On the other hand, $T > 0$ generally introduces the thermodynamic averaging over a large number of eigenstates. This directly reduces finite-size effects for static quantities, whereas for dynamical quantities spectra become denser. From Eq. (36) it follows that we get in spectra at elevated $T > 0$ typically $N_p \propto RM^2$ different peaks resulting in nearly continuous spectra. This is also evident from the example of a high-T result in Fig. 2.

It is plausible that finite-size effects at $T > 0$ become weaker. However, it should be recognized that there could exist several characteristic length scales in the physical (and model) system, e.g. the antiferromagnetic (AFM) correlation length ξ, the transport mean free path l_s etc. These lengths generally decrease with increasing T and results for related quantities get a macroscopic relevance provided that $\xi(T), l_s(T) < L$ where $L \propto N^{1/D}$ is the linear size of the system. However, there exist also anomalous cases, e.g., in an integrable system l_s can remain infinite even at $T \to \infty$ [20, 21].
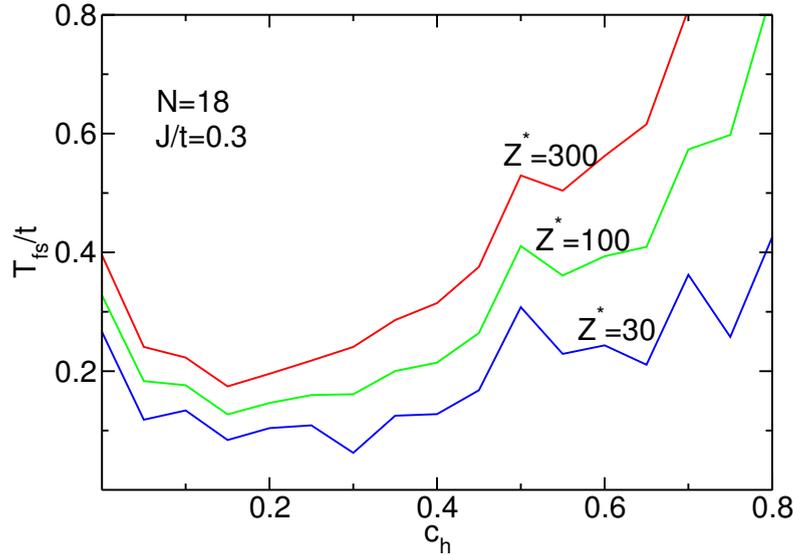
Fig. 3: *Finite-size temperature T_{fs} vs. hole doping c_h in the 2D t-J model with $J/t = 0.3$, as calculated with the FTLM in a system of $N = 18$ sites [7].*

As a simple criterion for finite size effects one can use the normalized thermodynamic sum $\bar{Z}(T)$, Eq. (28), which provides the effective number of MB states contributing at chosen T (note that for a system with a nondegenerate g.s. $\bar{Z}(T = 0) = 1$). A finite-size temperature T_{fs} can be thus defined with the relation $\bar{Z}(T_{fs}) = Z^*$ where in practice the range $10 < Z^* < 50$ is reasonable. Clearly, the FTLM is best suited just for systems with a large density of low lying MB states, i.e., for large $\bar{Z}$ at low T.

Since $\bar{Z}(T)$ is directly related to the entropy density s and the specific heat C_v of the system, large $\bar{Z}$ at low T is the signature of frustrated quantum MB systems, which are generally difficult to cope with using other methods (e.g., the QMC method). Typically examples of such strongly correlated electrons with an inherent frustration are the doped AFM and the t-J model, Eq. (2), in the strong correlation regime $J < t$. As an example, we present in Fig. 3 the variation of T_{fs} in the 2D t-J model with the hole doping $c_h = N_h/N$, as calculated for different $Z^* = 30$–300 for the fixed system of $N = 18$ sites and $J/t = 0.3$ as relevant for high-T_c cuprates. It is indicative that T_{fs} reaches a minimum for intermediate (optimum) doping $c_h = c_h^* \sim 0.15$, where we are able to reach $T_{fs}/t \sim 0.1$. Away from the optimum doping T_{fs} is larger, i.e., low-energy spectra are quite sparse both for the undoped AFM and even more so for the effectively noninteracting electrons far away from half-filling (for nearly empty or full band).

4.4   Low-temperature Lanczos method

The standard FTLM suffers at $T \to 0$ from statistical errors due to finite sampling R, both for static quantities, Eqs. (29), (30), as well as for dynamical correlations, Eqs. (36), (37). The discrepancy can be easily monitored by the direct comparison with the g.s. Lanczos method, Eqs. (8), (11). To avoid this problem, a variation of the FTLM method, called Low-temperature Lanczos method (LTLM) has been proposed [9] which obtains correct g.s. result (for finite systems) independent of the sampling R.

The idea of LTLM is to rewrite Eq. (19) in a symmetric form

$$\langle A \rangle = \frac{1}{Z} \sum_{n=1}^{N_{st}} \langle n | e^{-\beta H/2} A e^{-\beta H/2} | n \rangle, \tag{38}$$

and insert the Lanczos basis in analogy with the FTLM, Eq. (19), now represented with a double sum

$$\langle A \rangle = \frac{N_{st}}{ZR} \sum_{r=1}^{R} \sum_{j,l=0}^{M} e^{-\beta(\varepsilon_j^r + \varepsilon_l^r)/2} \langle r | \psi_j^r \rangle \langle \psi_j^r | A | \psi_l^r \rangle \langle \psi_l^r | r \rangle, \tag{39}$$

The advantage of the latter form is that it satisfies the correct $T = 0$ limit provided that the g.s. is well converged, i.e., $|\psi_0^r\rangle \sim |\Psi_0\rangle$. It then follows from Eq. (39),

$$\langle A \rangle = \sum_{r=1}^{R} \langle r | \Psi_0 \rangle \langle \Psi_0 | A | \Psi_0 \rangle \langle \Psi_0 | r \rangle \Big/ \sum_{r=1}^{R} \langle \Psi_0 | r \rangle \langle r | \Psi_0 \rangle = \langle \Psi_0 | A | \Psi_0 \rangle, \tag{40}$$

for any chosen set of $|r\rangle$. For the dynamical correlations $C(t)$ one can in a straightforward way derive the corresponding expression in the Lanczos basis

$$C(\omega) = \frac{N_{st}}{ZR} \sum_{r=1}^{R} \sum_{i,j,l=0}^{M} e^{-\beta(\varepsilon_i^r + \varepsilon_l^r)/2} \langle r | \psi_i^r \rangle \langle \psi_i^r | A^\dagger | \tilde{\psi}_j^{rl} \rangle \langle \tilde{\psi}_j^{rl} | A | \psi_l^r \rangle \langle \psi_l^r | r \rangle \, \delta(\omega - \tilde{\varepsilon}_j^{rl} + \frac{1}{2}(\varepsilon_i^r + \varepsilon_l^r)). \tag{41}$$

It is again evident that for $T \to 0$ the sampling does not influence results, being correct even for $R = 1$ if the g.s. $|\Psi_0\rangle$ is well converged for all starting $|r\rangle$. The payoff is in an additional summation over the new Lanczos basis sets $|\tilde{\psi}_j^{rl}\rangle$, which needs to be started from each $A|\psi_l^r\rangle$ in Eq. (41) separately. Since the LTLM is designed for lower T, one can effectively restrict summations in (i, l) in Eq. (41) to much smaller $M' \ll M$, where only lowest states with $\varepsilon_i^r, \varepsilon_l^r \sim E_0$ contribute [9], and in addition use smaller $M_1 \ll M$ for the basis $|\tilde{\psi}_j^{rl}\rangle$.

An alternative version for a Lanczos-type approach [22] to dynamical quantities is not to start the second Lanczos run from $A|r\rangle$ [7] or from $A|\psi_l^r\rangle$ [9], but from

$$|\widetilde{Ar}\rangle = \sum_{l=0}^{M} A|\psi_l^r\rangle e^{-\beta \varepsilon_l^r/2} \langle \psi_l^r | r \rangle. \tag{42}$$

In this way one obtains with the second Lanczos run the Lanczos eigenstates $|\tilde{\psi}_k^r\rangle$, which cover the relevant Hilbert space for starting random vector $|r\rangle$ and the inverse temperature β. The resulting dynamical autocorrelation function is

$$C(\omega) = \frac{N_{st}}{RZ} \sum_{r=1}^{R} \sum_{i,k=0}^{M} e^{-\beta \varepsilon_i^r/2} \langle r | \psi_i^r \rangle \langle \psi_i^r | A^\dagger | \tilde{\psi}_k^r \rangle \langle \tilde{\psi}_k^r | \widetilde{Ar} \rangle \, \delta(\omega - \tilde{\varepsilon}_k^r + \varepsilon_i^r). \tag{43}$$

In this way the sufficiency of only one random vector in the $T = 0$ limit is reproduced, while at $T > 0$ the algorithm has the same time efficiency as the FTLM, but with much smaller random sampling needed to reach the same accuracy (at least for low T). However, the price paid is that results for each T need to be calculated separately, while within the FTLM all T (or T up to a certain value within the LTLM) are evaluated simultaneously.

4.5   Microcanonical Lanczos method

While most investigations in strongly correlated systems focus on the low-T regime, there are systems where dynamical properties are nontrivial even at high T. A well known such case is the spin diffusion constant $D_s(T)$ in the isotropic Heisenberg model, Eq. (1), whose value is not known, and even its existence at any $T > 0$ is uncertain. Similar although somewhat less controversial is the case of transport quantities, both for integrable or generic nonintegrable models. Whereas the FTLM seems well adapted for studies of transport response functions, oscillations due to a limited M can compromise the crucial low-ω resolution, cf. Fig. 2.

At elevated T it is therefore an advantage to use the microcanonical Lanczos method (MCLM) [10], employing the fact from statistical physics that in the thermodynamic limit (for large system) the microcanonical ensemble should yield the same results as the canonical one. Short-comings of the MCLM are due to the fact that in finite systems statistical fluctuations are much larger within the microcanonical ensemble. Still, reachable finite-size systems have a very high density of states in the core of the MB spectrum as probed by high T. Hence, statistical fluctuations are at high T effectively smoothed out in contrast to low-T properties dominated by a small number of low lying MB states.

The implementation of the MCLM is quite simple and straightforward. One first determines the target energy $\lambda = \langle H \rangle(T)$ which represents the microcanonical energy equivalent to the canonical average energy for chosen T and system size N. Since λ is a parameter within the MCLM, one can relate it to T by performing either FTLM (simplified due to H being a conserved quantity) on the same system, or extrapolating full ED results (with linear dependence on N) on small lattices. Next we find a representative microcanonical state $|\Psi_\lambda\rangle$ for the energy λ. One convenient way within the Lanczos-type approach is to use the new operator

$$V = (H - \lambda)^2. \tag{44}$$

Performing Lanczos iterations with the operator V yields again the extremum eigenvalues, in particular the lowest one close to $V \sim 0$. In contrast to the g.s. procedure, the convergence to a true eigenstate cannot be reached in system sizes of interest even with $M_1 \gg 100$. The reason is the extremely small eigenvalue spacing of the operator V, scaling as $\Delta V_n \propto (\Delta E/N_{st})^2$, with ΔE being the whole energy span within the given system. Fortunately such a convergence is not necessary (nor even desired) since the essential parameter is the small energy uncertainty σ_E, given by

$$\sigma_E^2 = \langle \Psi_\lambda | V | \Psi_\lambda \rangle. \tag{45}$$

For small energy spread $\sigma_E/\Delta E < 10^{-3}$ typically $M_1 \sim 1000$ is needed. Again, to avoid storing M_1 Lanczos wavefunctions $|\phi_i\rangle$ the Lanczos procedure is performed twice as described in Sec. 2.2, i.e., the second time with known tridiagonal matrix elements to calculate finally $|\Psi_\lambda\rangle$ in analogy with Eq. (6). The latter is then used to evaluate any static expectation average $\langle A \rangle$ or the dynamical correlation function as in Eq. (9),

$$C(\omega, \lambda) = \langle \Psi_\lambda | A^\dagger \frac{1}{\omega^+ + \lambda - H} A | \Psi_\lambda \rangle. \tag{46}$$

The latter is evaluated again using Lanczos iterations with M_2 steps starting with the initial wavefunction $|\tilde{\phi}_0\rangle \propto A|\Psi_\lambda\rangle$ and $C(\omega, \lambda)$ is represented in terms of a continued fraction. Since the MB levels are very dense and correlation functions smooth at $T \gg 0$, large $M_2 \gg 100$ are needed but as well easily reachable to achieve high-ω resolution in $C(\omega, \lambda)$.

It is evident that the computer requirement for the MCLM both regarding the CPU and memory are essentially the same as for the g.s. dynamical calculations except that typically $M_1, M_2 \gg 100$. In particular, requirements are less demanding than using the FTLM with $M > 100$. A general experience is that for systems with large $N_{st} \gg 10000$ the MCLM dynamical results agree very well with FTLM results for the same system. It should also be noted that the actual frequency resolution $\delta\omega$ in $C(\omega, \lambda)$, Eq. (46), is limited by $\delta\omega \sim \sigma_E$ which is, however, straightforward to improve by increasing M_1, M_2 with typical values $M_1, M_2 > 1000$. One can also improve the MCLM results for any T by performing an additional sampling over initial random starting $|\phi_0\rangle$ as well as over λ with a probability distribution $p(\lambda)$ simulating the canonical ensemble in a finite-size system, i.e., by replacing Eq. (46) with

$$C(\omega) = \sum_\lambda p(\lambda) C(\omega, \lambda). \tag{47}$$

4.6 Statical and dynamical quantities at $T > 0$: Applications

The FTLM has been designed to deal with the simplest tight-binding models of strongly correlated electrons, at the time mostly with challenging microscopic electronic models of high-T_c superconductors [6, 7], where besides superconductivity there is a variety of anomalous non-Fermi-liquid-like properties even in the normal state. Clearly of interest in this connection are prototype MB models as the Heisenberg model, Eq. (1), the t-J model, Eq. (2), and the Hubbard model on the 2D square lattice. The unfrustrated Heisenberg model can be numerically studied on much bigger lattices with QMC and related methods. The 2D Hubbard model was and still is mostly subject of DMFT and QMC studies, since at half-filling or close to it the Lanczos methods are quite restricted due to the large N_{st} even for modest sizes $N \sim 16$. Therefore one focus of Lanczos-based approaches was on the t-J model being, with some generalizations, a microscopic representation of electronic properties of high-T_c cuprates.

Thermodynamic quantities such as chemical potential μ, entropy density s, specific heat C_v are the easiest to implement within the FTLM. Their T- and (hole) doping c_h-dependence within the t-J model on a 2D square lattice (calculated for up to $N = 26$ sites) reveal the very anomalous behavior of doped Mott insulators [23] (as evident already from Fig. 3), confirmed also by results for the more complete Hubbard model [24].

The advantages of the FTLM and also its feasibility for the 2D t-J model are even more evident in numerous studies of spin- and charge-dynamics at $T > 0$ [7], which show good agreement with neutron scattering and NMR [25, 26], optical conductivity $\sigma(\omega)$ and resistivity $\rho(T)$ [27], as well as some other anomalous properties of the cuprates [8]. As an example of a transport quantity hardly accessible by other methods we present in Fig. 4 the universal planar resistivity $\rho(T)$, as extracted from the dynamical conductivity $\sigma(\omega \to 0) = 1/\rho$, within the t-J model
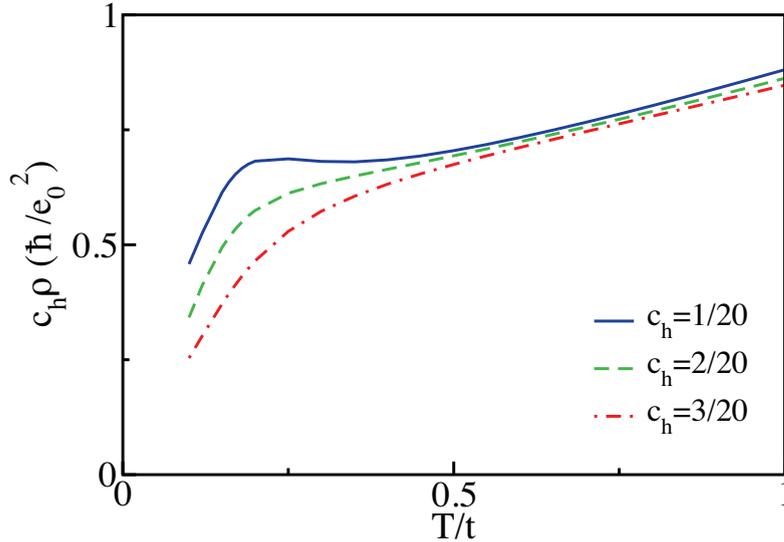
Fig. 4: *Normalized 2D resistivity $c_h\rho$ vs. T/t within the t-J model with $J/t = 0.3$ for different hole concentrations c_h [27].*

for different doping levels c_h [27]. The result in Fig. 4 clearly show a linear dependence below the pseudogap temperature T^* dependent on doping c_h. Another characteristic signature is a saturation (plateau) of $\rho(T)$ at low doping and the universal trend at high T.

Spectral properties as manifested in the single-particle spectral functions $A(\mathbf{k}, \omega)$ are at the core of the understanding of cuprates, as well as of strongly correlated electrons in general. Here, even g.s. and low-T properties are a challenge for numerical studies whereby the FTLM can be viewed as a controlled way to get reliable (macroscopic-like) $T \to 0$ result, in contrast to quite finite-size plagued results obtained via the g.s. Lanczos procedure [5]. Using the FTLM at $T \sim T_{fs}$ with twisted boundary conditions we can simulate a continuous wavevector $\mathbf{k}$. Using in addition coarse graining averaging one can reach results for $A(\mathbf{k}, \omega)$ [28,29] giving insights into electron vs. hole doped angle-resolved photoemission experiments, quasiparticle relaxation, and waterfall-like effects. A characteristic result of such studies is in Fig. 5 for the single-particle density of states $\mathcal{N}(\omega) = \sum_{\mathbf{k}} A(\mathbf{k}, \omega)$ [28]. Here, the strength of the FTLM is visible in the high ω resolution within the most interesting low-ω window. Interesting and reproducible are also nontrivial spectral shapes as the sharp peak close to $\omega < 0$ and a broad shoulder for $\omega \ll 0$. Most important is, however, the evident pseudogap (observed also experimentally in cuprates) visible at $\omega \sim 0$ in the low-doping regime.

Besides the challenging models for cuprates there have been also studies of static and dynamical properties of multiband and multiorbital models which either reduce to the generalized t-J model [30] or to Kondo lattice models [31, 32]. While the increasing number of local basis states K clearly limits the applicability of ED-based methods, they are competitive in treating nontrivial frustrated spin models less suitable for QMC and other methods, however closely related to the physics of novel materials. Moreover, frustrated models are characterized by a large entropy density s and related low T_{fs}, essential conditions for the feasibility of FTLM results. Examples of such systems are the Shastry-Sutherland model [33, 34], the 2D J_1-J_2 model [35], and properties of frustrated magnetic molecules [36–38].
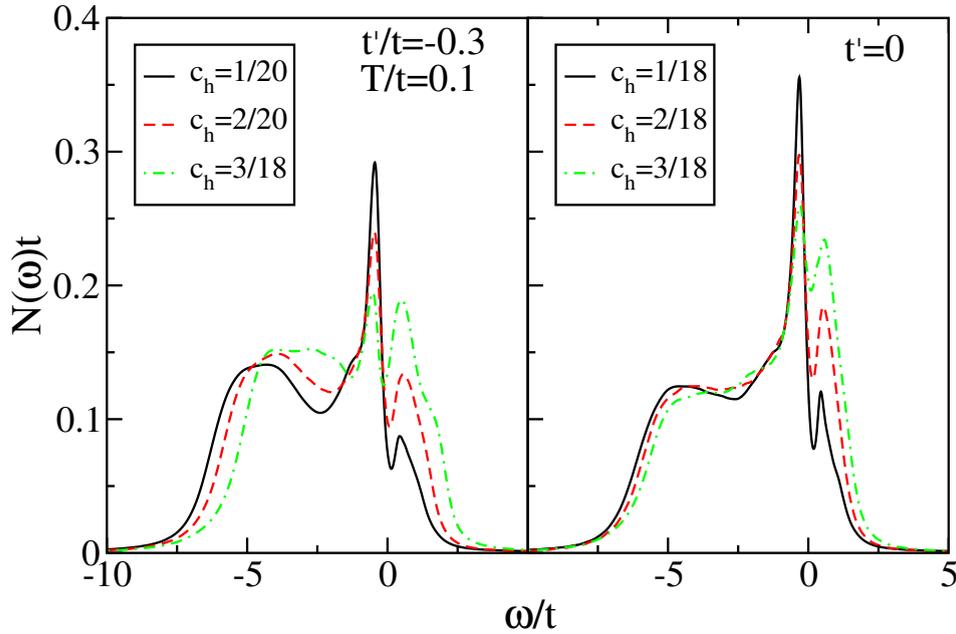
Fig. 5: *Density of states $\mathcal{N}(\omega)$ for different dopings c_h within the extended t-J model with n.n.n. hopping $t' = -0.3t$ and $t' = 0$, respectively [28].*

Another class of problems which can be quite effectively dealt with using the FTLM and MCLM approaches is the fundamental as well as experimentally relevant problem of transport in 1D systems of interacting fermions as realized, e.g., in quasi-1D spin-chain materials [39]. It has been recognized that the transport response at any $T > 0$ crucially differs between integrable and nonintegrable systems. Since the 1D isotropic as well as anisotropic Heisenberg model, Eq. (1), is integrable it opens a variety of fundamental questions of anomalous transport in such systems, the effects of perturbative terms and impurities. Such fundamental questions on transport and low-ω dynamic response remain nontrivial even at high T [20, 21], hence the MCLM is the most feasible and straightforward method. It has been in fact first probed on the anomalous transport in 1D insulators [40] but furtheron used to study interaction-induced transport at $T > 0$ in disordered 1D systems [41, 42], in particular in relation to challenging problems of many-body localization [43, 44] being inherently the question of low-frequency dynamics at $T \to \infty$.

In Fig. 6 we present as an example MCLM result for the dynamical spin conductivity in the anisotropic Heisenberg model, Eq. (1), where $J^{zz} \neq J^{xx} = J^{yy} = J$ in the Ising-like (with the spin gap in the g.s.) regime $\Delta = J^{zz}/J > 1$. Results for the high-T dynamical spin conductivity $T\sigma(\omega)$ are shown for various next-neighbor (anisotropic) couplings $\alpha = J_2^{zz}/J$. The first message is that the MCLM is well adapted for the high ω resolution (here using $M_1 = M_2 = 2000$) and reaching large $N = 30$ ($N_{st} \sim 5.10^6$ in a single $S^z = 0, q$ sector). Another conclusion is that the dynamics of such systems is very anomalous. For the integrable case $\alpha = 0$ we find $\sigma_0 = \sigma(0) \sim 0$ but also an anomalous finite-size peak at $\omega_p \propto 1/N$ [40]. At the same time breaking integrability with $\alpha > 0$ appears to lead to $\sigma_0 > 0$ still approaching an 'ideal' insulator (insulating at all T) for a weak perturbation $\sigma_0(\alpha \to 0) \to 0$ [45].
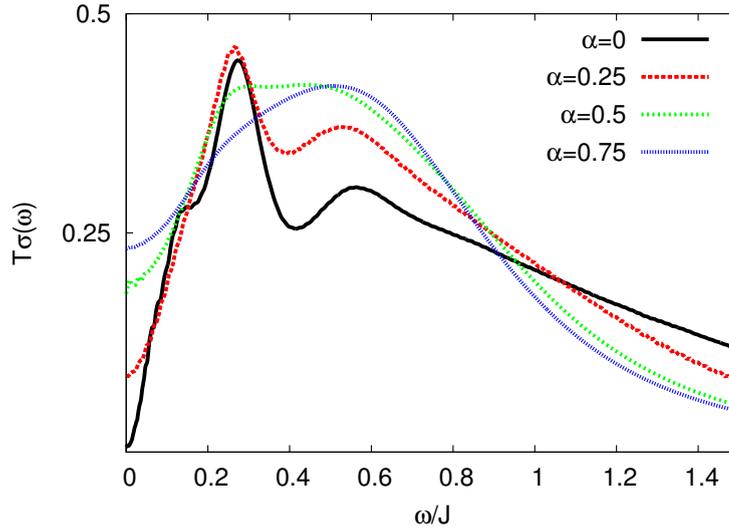
Fig. 6: *High-T dynamical spin conductivity $T\sigma(\omega)$ within the anisotropic Heisenberg model in the Ising-like regime, $\Delta = 1.5$, and various next-neighbor interactions $\alpha = J_2^{zz}/J$ as calculated with the MCLM on a chain of $N = 30$ sites.*

5   Real-time dynamics using the Lanczos method

Research in the field of non-equilibrium dynamics of complex quantum systems constitutes a formidable theoretical challenge. When dealing with ED approaches or calculations in a reduced basis, the time evolution of the time-dependent Shrödinger equation,

$$i\frac{\partial \Psi(t)}{\partial t} = H(t)\Psi(t), \tag{48}$$

can be efficiently obtained using the time-dependent Lanczos technique, as originally described in Ref. [46] and later applied and analyzed in more detail [47]. One of the straightforward reasons is that most commonly the Lanczos method is used to compute the g.s. of MB Hamiltonian. Generalizing the method to time-dependent calculation represents only a minor change to already existing codes. Even though the method is most suitable for the time evolution of the time-independent Hamiltonian, it can nevertheless be applied even to the time-dependent case. The time evolution of $|\Psi(t)\rangle$ is then calculated by a step-wise propagation in time t by small time increments δt, generating at each step a Lanczos basis of dimension M (typically $M < 10$), to follow the evolution

$$|\Psi(t + \delta t)\rangle \simeq e^{-iH(t)\delta t}|\Psi(t)\rangle \simeq \sum_{l=1}^{M} e^{-i\varepsilon_l \delta t}|\psi_l\rangle\langle\psi_l|\Psi(t)\rangle, \tag{49}$$

where $|\psi_l\rangle, \varepsilon_l, l = 0 \ldots M$ are Lanczos eigenfunctions and eigenvalues, respectively, obtained via the Lanczos iteration started with $|\phi_0\rangle = |\Psi(t)\rangle$. The advantage of the time-evolution method following Eq. (49) is that it preserves the normalization of $|\Psi(t + \delta t)$ for arbitrarly large δt. The approximation of finite M in Eq. (49) is also correct at least to the M-th Taylor-expansion order in δt. It is, however, important to stress that δt should be chosen small enough to properly take into account the time-dependence of $H(t)$. E.g., when driving the system with a constant external electric field, $\delta t/t_B \sim 10^{-3}$ where t_B is the Bloch oscillation period [48,45].

So far, investigations of correlated systems under the influence of a driving electric field in 1D using Lanczos time-evolution focused on generic systems, like the metallic and Mott-insulating regime of interacting spinless fermions [48, 45]. Even though rather small systems can be studied it has been established that a steady state can be reached without any additional coupling to a heat bath, provided that the Joule heating of the system is properly taken into account.

6 Discussion

Exact diagonalization based methods, both the full ED and the Lanczos-type ED approach, are very extensively employed in the investigations of strongly correlated MB quantum systems in solids and elsewhere. The reason for their widespread use are several: a) unbiased approach to the MB problem without any simplifications or approximations, independent of the complexity of the MB system, b) relative simplicity of generating the codes for various models and observables, c) easy and straightforward testing of codes, d) direct interpretation of the obtained quantum MB states and their possible anomalous structure and properties, e) high pedagogical impact as a quick and at the same time very nontrivial introduction into the nature of MB quantum physics. Also the Lanczos-based methods described in this review, i.e., the g.s. Lanczos method for static and dynamic quantities, and the somewhat more elaborate FTLM, MCLM, LTLM and EDLFS, require rather modest programming efforts in comparison with more complex numerical methods, e.g., QMC- and DMRG-based methods, as described in other chapters. Clearly, the main drawback of ED methods is the smallness of lattice sizes N limited by the number of basis states (at present $N_{st} < 10^9$) that can be treated with a Lanczos iteration procedure. The achievable N with ED methods appears quite modest in comparison with some established and recently developed numerical methods, such as QMC, DMRG, matrix-product-states methods, etc. Still, in spite of the intensive developments and advances of novel numerical methods in last two decades, there are several aspects of strong-correlation physics, where ED-based methods are so far either the only feasible or at least superior ones. In this chapter we have focused mostly on Lanczos-based methods and applications where they are competitive and get nontrivial results with a macroscopic validity:

a) MB g.s. and its properties: for frustrated and complex models mostly so far do not offer alternative powerful methods at least beyond $D = 1$ systems, where DMRG is efficient.

b) $T > 0$ static properties evaluated with as the FTLM and the LTLM are most powerful and reliable for frustrated and complex system, in particular in systems with high degeneracies of MB states and large entropy at low T,

c) $T > 0$ Lanczos methods for dynamical quantities, such as FTLM and MCLM, yield for many models and geometries results superior to other methods or, in several cases, even the only accessible results. Particular advantages are the high ω resolution at all T beyond the finite size limit $T > T_{fs}$, macroscopic-like results at low T with proper $T \to 0$ scaling, and the possibility of detailed studies of systems with nontrivial dynamics at any, in particular high T.

d) The Lanczos technique is the natural application for methods with a restricted MB basis sets and DMRG-type targeting, as well as for the real-time evolution of strongly correlated systems.

References

[1] C. Lanczos, J. Res. Nat. Bur. Stand., **45**, 255 (1950)

[2] B.N. Parlett: *The Symmetric Eigenvalue Problem* (Prentice Hall, Eaglewood Cliffs, 1980)

[3] J.W. Demmel: *Applied Numerical Linear Algebra* (SIAM, Philadelphia, 1997)

[4] R. Haydock, V. Heine, and M.J. Kelly, J. Phys. C: Solid State Phys. **8**, 2591 (1975)

[5] for a review, see E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994)

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

[7] for a review, see J. Jaklič and P. Prelovšek, Adv. Phys. **49**, 1 (2000)

[8] for a review, see P. Prelovšek and J. Bonča, in A. Avella and F. Mancini (eds.):
     Strongly Correlated Systems: Numerical Methods (Springer, 2013)

[9] M. Aichhorn, M. Daghofer, H.G. Evertz, and W. von der Linden,
     Phys. Rev. B **67**, 161103(R) (2003)

[10] M.W. Long, P. Prelovšek, S. El Shawish, J. Karadamoglou, and X. Zotos,
      Phys. Rev. B **68**, 235106 (2003)

[11] for a review, see M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)

[12] P.W. Leung, Phys. Rev. B **73**, 14502 (2006)

[13] T. Tohyama, Y. Inoue, K. Tsutsui, and S. Maekawa, Phys. Rev. B **72**, 045113 (2005)

[14] J. Oitmaa and D.D. Betts, Can. J. Phys. **56**, 897 (1978)

[15] D.D. Betts, H.Q. lin, and J.S. Flynn, Can. J. Phys. **77**, 3535 (1999);
      P.R.C. Kent, M. Jarrell, T.A. Maier, and Th. Pruschke Phys. Rev. B **72**, 060411 (2005)

[16] J. Cullum and R.A. Willoughby, J. Comp. Phys. **44**, 329 (1981)

[17] H. Mori, Prog. Theor. Phys. **34**, 423 (1965)

[18] M. Imada and M. Takahashi, J. Phys. Soc. Jpn. **55**, 3354 (1986)

[19] R.N. Silver and H. Röder, Int. J. Mod. Phys. C **5**, 735 (1995)

[20] X. Zotos and P. Prelovšek, Phys. Rev. B **53**, 983 (1996)

[21] for a review, see F. Heidrich-Meisner, A. Honecker, and W. Brenig,
      Eur. Phys. J. Special Topics **151**, 135 (2007)

[22] J. Kokalj, Ph.D. Thesis, University of Ljubljana, 2010, unpublished.

[23] J. Jaklič and P. Prelovšek, Phys. Rev. Lett. **77**, 892 (1996)

[24] J. Bonča and P. Prelovšek, Phys. Rev. B **67**, 180502(R) (2003)

[25] J. Jaklič and P. Prelovšek, Phys. Rev. Lett. **75**, 1340 (1995)

[26] P. Prelovšek, I. Sega, and J. Bonča, Phys. Rev. Lett. **92**, 027002 (2004)

[27] M.M. Zemljič and P. Prelovšek, Phys. Rev. B **72**, 075108 (2005)

[28] M.M. Zemljič and P. Prelovšek, Phys. Rev. B **75**, 104514 (2007)

[29] M.M. Zemljič, P. Prelovšek, and T. Tohyama, Phys. Rev. Lett. **100**, 036402 (2008)

[30] P. Horsch, J. Jaklič and F. Mack, Phys. Rev. B **59**, 6217 (1999)

[31] P. Horsch, J. Jaklič and F. Mack, Phys. Rev. B **59**, 14149(R) (1999)

[32] K. Haule, J. Bonča, and P. Prelovšek, Phys. Rev. B **61**, 2482 (2000)

[33] S. El Shawish, J. Bonča, and I. Sega, Phys. Rev. B **72**, 184409 (2005)

[34] S. El Shawish, A. Ramšak, and J. Bonča, Phys. Rev. B **75**, 205442 (2007)

[35] B. Schmidt, P. Thalmeier, and N. Shannon, Phys. Rev. B **76**, 125113 (2007)

[36] J. Schnack and O. Wendland, Eur. Phys. J. B **78**, 535 (2010)

[37] for a review, see J. Schnack and J. Ummethum, Polyhedron **66**, 28 (2103)

[38] M. Haertel, J. Richter, D. Ihle, J. Schnack, and S.-L. Drechsler,
     Phys. Rev. B **84**, 104411 (2011)

[39] for a review, see C. Hess, Eur. Phys. J. Special Topics **151**, 73 (2007)

[40] P. Prelovšek, S. El Shawish, X. Zotos, and M. Long, Phys. Rev. B **70**, 205129 (2004)

[41] A. Karahalios, A. Metavitsiadis, X. Zotos, A. Gorczyca, and P. Prelovšek,
     Phys. Rev. B **79**, 024425 (2009)

[42] O.S. Barišić and P. Prelovšek, Phys. Rev. B **82**, 161106 (2010)

[43] O.S. Barišić, J. Kokalj, I. Balog, and P. Prelovšek, Phys. Rev. B **94**, 045126 (2016)

[44] P. Prelovšek, O.S. Barišić, and M. Žnidarič, Phys. Rev. B **94**, 241104 (2016)

[45] M. Mierzejewski, J. Bonča, and P. Prelovšek, Phys. Rev. Lett. **107**, 126601 (2011)

[46] T.J. Park and J.C. Light, J. Chem. Phys. **85**, 5870 (1986)

[47] N. Mohankumar and S.M. Auerbach, Comp. Phys. Comm. **175**, 473 (2006)

[48] M. Mierzejewski and P. Prelovšek, Phys. Rev. Lett. **105**, 186405 (2010)