



# CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Time-evolving a matrix product state with long-ranged interactions

Michael P. Zaletel, Roger S. K. Mong, Christoph Karrasch, Joel E. Moore, and Frank Pollmann

Phys. Rev. B **91**, 165112 — Published 7 April 2015

DOI: [10.1103/PhysRevB.91.165112](https://doi.org/10.1103/PhysRevB.91.165112)

# Time-evolving a matrix product state with long-ranged interactions

Michael P. Zaletel,<sup>1,2</sup> Roger S. K. Mong,<sup>3</sup> Christoph Karrasch,<sup>1,4</sup> Joel E. Moore,<sup>1,4</sup> and Frank Pollmann<sup>5</sup>

<sup>1</sup>*Department of Physics, University of California, Berkeley, California 94720, USA*

<sup>2</sup>*Department of Physics, Stanford University, Stanford, California 94305, USA*

<sup>3</sup>*Walter Burke Institute for Theoretical Physics and Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

<sup>4</sup>*Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

<sup>5</sup>*Max-Planck-Institut für Physik komplexer Systeme, 01187 Dresden, Germany*

(Dated: February 17, 2015)

We introduce a numerical algorithm to simulate the time evolution of a matrix product state under a long-ranged Hamiltonian in moderately entangled systems. In the effectively one-dimensional representation of a system by matrix product states, long-ranged interactions are necessary to simulate not just many physical interactions but also higher-dimensional problems with short-ranged interactions. Since our method overcomes the restriction to short-ranged Hamiltonians of most existing methods, it proves particularly useful for studying the dynamics of both power-law interacting, one-dimensional systems, such as Coulombic and dipolar systems, and quasi two-dimensional systems, such as strips or cylinders. First, we benchmark the method by verifying a long-standing theoretical prediction for the dynamical correlation functions of the Haldane-Shastry model. Second, we simulate the time evolution of an expanding cloud of particles in the two-dimensional Bose-Hubbard model, a subject of several recent experiments.

## I. INTRODUCTION

The ability to study dynamical properties in and out of equilibrium is essential for understanding the physics of strongly interacting systems. Following the success of the density-matrix renormalization group (DMRG) for finding one-dimensional (1D) ground states [1], a number of closely related techniques have been developed to explore the dynamical properties of short-ranged 1D systems [2–6]. This exciting development has given access to experimentally relevant observables, such as dynamical correlation functions which can be compared with data from neutron scattering and ultracold atomic gasses, and non-equilibrium dynamics, providing insight into long standing questions about thermalization [7]. Simultaneously, large-scale DMRG has begun to study ground-state properties of quasi-two dimensional (2D) quantum systems, such as strips and cylinders, allowing one to probe much larger systems than accessible to exact diagonalization [8]. The 2D-DMRG method orders the sites of a 2D lattice into a 1D chain with long-ranged interactions. Truly 2D tensor network methods can capture more entanglement, so may eventually supplant this approach [9, 10], but 2D-DMRG is currently a standard tool due to its reliability.

It is desirable to combine these two developments in order to evaluate dynamical properties of quasi-2D systems (e.g., the time evolution of bosons in a 2D optical trap as shown in Fig. 1). However, the existing DMRG-based time-evolution methods cannot be easily applied to a quasi-2D system. This is mainly due to the long-ranged interactions that occur when representing a 2D system as a 1D chain; a similar difficulty exists for 1D systems with power-law Coulombic and dipolar interactions.

In this work we address this problem by providing a method to time-evolve long-ranged Hamiltonians. The unique advantage of the method is that it simultaneously (a) can be applied to any long-ranged Hamiltonian while preserving all symmetries, (b) has a constant error per site in the thermodynamic

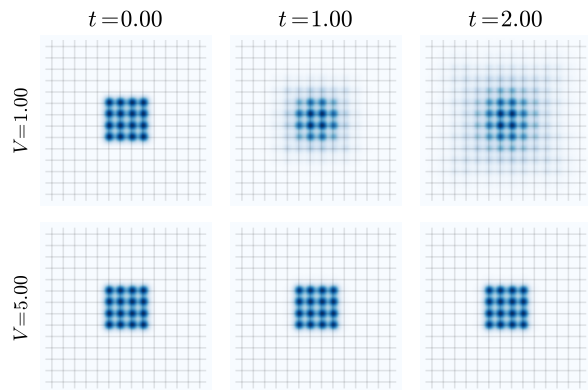


FIG. 1. Quasi-exact time evolution of interacting hard-core bosons in a  $14 \times 14$  lattice trap. The bosons hop with bandwidth  $J = 1$  and interact with nearest-neighbor repulsion  $V$ . 16 bosons begin in an un-entangled product state, and evolve in time  $t$  from left to right. In the top row,  $V = 1$ , and the bosons expand outward. In the bottom row,  $J < V = 5$ , the bosons remain trapped in a bound state due to the strong interactions. A similar effect has been observed experimentally in cold-atom optical lattices [11, 12].

limit at fixed computational effort, (c) can be applied to an infinitely long system assuming translation invariance and (d) can be easily implemented using standard DMRG methods.

Like other 1D methods, we work in the framework of matrix product states (MPSs) [13–15], a variational ansatz for finitely entangled states within which we wish to simulate the full many-body dynamics (consequently, the method is practical only for moderately entangled systems). The structure of an MPS can be generalized to operators, called matrix product operators (MPOs) [16]. An MPO can be efficiently applied to an MPS using standard methods [7, 17, 18]. If a long-ranged Hamiltonian  $H$  has a compact MPO approximation for  $e^{tH}$ , then the time evolution can be efficiently simulated by succes-

sively applying the MPO to the MPS. The most naive time-stepper, an Euler step  $1 + tH$ , as well as its Runge-Kutta [19] and Krylov [5, 20, 21] improvements, indeed has an efficient MPO representation. But these global methods have an error per *site* which diverges with the system size  $L$ , for example as  $\mathcal{O}(Lt^2)$  for the Euler step, which eventually renders them impractical as  $L \rightarrow \infty$ . For certain simple  $H$ , such as a nearest neighbor interactions or a sum of commuting terms [17], a compact MPO with *finite* error per site exists, which is the basis behind the highly successful time evolving block decimation (TEBD) [2] and tDMRG [3, 4]. However, these methods do not generalize well for long-ranged Hamiltonians, which is the focus of this work.

The basic insight of this work is that a Hamiltonian which is expressed as a sum of terms  $H = \sum_x H_x$  admits a *local* version of a Runge-Kutta step; for instance we could improve the Euler step by taking

$$1 + t \sum_x H_x \rightarrow \prod_x (1 + tH_x). \quad (1)$$

The error is still at  $\mathcal{O}(t^2)$ , so it is formally a 1<sup>st</sup>-order time stepper. But any set of distant regions all receive the correct 1<sup>st</sup>-order step in parallel. Hence, in contrast to the naive Euler step, the total error scales as  $Lt^2$ , rather than as  $L^2t^2$ : this improvement is absolutely essential in order to apply the method in the thermodynamic limit. The main result of this work is that an improved version of Eq. (1) has a very compact MPO representation which can easily be extended to higher-order approximations in  $\mathcal{O}(t^p)$ .

In Fig. 2, we compare the accuracy of the methods proposed here, called  $W^I$  and  $W^{II}$ , against TEBD and global 2<sup>nd</sup> order Runge-Kutta. TEBD works for short-ranged Hamiltonians, so we compare by quenching from product states into the spin-1/2 nearest-neighbor Heisenberg chain, where a high order TEBD calculation serves as a quasi-exact reference. Runge-Kutta is orders of magnitude less accurate, with an error that scales as  $L^5$  compared to  $L$  for TEBD and  $W^{I/II}$ . Both TEBD and  $W^{I/II}$  are comparable in accuracy; for evolution starting from a Neel state,  $W^{II}$  is slightly more accurate than TEBD, while from a random state TEBD is more accurate. Any such difference can be easily mitigated by a small decrease in time step. But unlike TEBD,  $W^{I/II}$  can be immediately applied to a long-ranged problem without a Trotter decomposition.

To our knowledge, the other existing method which can time-evolve long-ranged interactions with a constant error per site is the time dependent variation principle (TDVP), which projects the exact Schrödinger equation into the MPS variational space and numerically integrates the resulting equations [22, 23]. While the method has yet to be applied to quasi-2D systems, a version was successfully applied to the long-ranged transverse field Ising model [24, 25]. However, in contrast to the proposal here, which involves the standard tensor network technique of applying an MPO, the TDVP requires an entirely distinct and relatively complex set of algorithms, the stability of which can decrease with the desired accuracy. It will be a useful subject for future work to make a detailed comparison between TDVP and the present MPO approach.

The first application presented here is a calculation of a

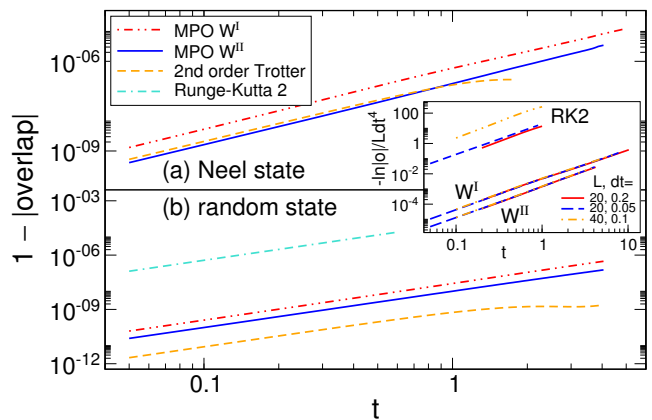


FIG. 2. Comparison of 2<sup>nd</sup>-order MPOs  $W^I$ ,  $W^{II}$ , TEBD, and global Runge-Kutta for the spin-1/2 Heisenberg chain. 4<sup>th</sup>-order TEBD serves as a quasi-exact reference for calculating errors. Panels (a), (b) show quenches starting from a  $L = 20$  Neel state and random state respectively. In the inset, we show the scaling of the errors for system sizes  $L = 20, 40$ . For  $W^{I/II}$  we find perfect collapse to the expected scaling  $Lt^4$ , as the error per site remains constant in the thermodynamic limit. In contrast, for global Runge-Kutta the error increases as  $L^5t^4$ .

dynamical correlation function of the Haldane-Shastry spin chain, which is a 1D spin-half antiferromagnet with power-law long-ranged interactions [26, 27]. Our numerical simulations agree with the analytic exact results [28] up to long times, which serves as a check of the method's accuracy, and show a ballistic spreading of correlations consistent with the model's integrability. The second application is the simulation of dynamics in a 2D Bose-Hubbard model. Here we focus on a class of experiments with ultracold atomic gases that study expansion of a cloud that is initially confined to a small region of the lattice [11, 12]. The main qualitative surprise in the experiments is that even repulsive interactions can lead to self-trapped states, which is reproduced in our model calculation along with several other features, shown in Fig 1. We will further elaborate on these applications later.

## II. MATRIX PRODUCT OPERATORS

In order to understand our main result, we review some basic facts regarding MPOs. An operator  $Z$  acting on a 1D chain with physical sites labeled by  $i$  has an MPO representation

$$Z = \cdots \hat{W}_{(1)} \hat{W}_{(2)} \hat{W}_{(3)} \cdots \quad (2)$$

where each  $\hat{W}_{(i)}$  is a matrix of operators acting on the Hilbert space of site- $i$  (with physical indices  $m_i, m'_i$ ),

$$[\hat{W}_{(i)}]_{a_{i-1}a_i} = \sum_{m_i, m'_i} [W_{(i)}]_{a_{i-1}a_i}^{m_i m'_i} |m_i\rangle \langle m'_i|, \quad (3)$$

with  $[W_{(i)}]_{a_{i-1}a_i}^{m_i m'_i} \in \mathbb{C}$ . In Eq. (2), the matrices are contracted by summing over all indices  $a_i = 1, \dots, \chi_i$ . These indices live in the space between sites  $(i, i + 1)$ , which we refer to as

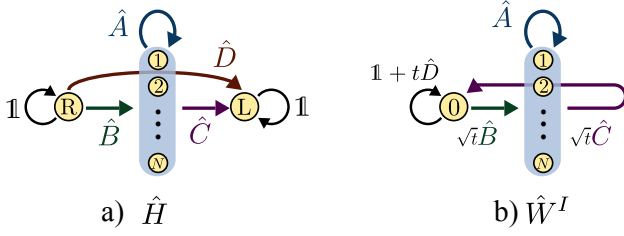


FIG. 3. Graphical depictions of MPOs for (a) the Hamiltonian  $H$  and (b) the time-stepper  $\hat{W}^I(t)$ . As explained in Ref. 30, by analogy to a finite-state-machine the indices of the MPO (labeling rows and columns) are represented as nodes of a graph, while the entries of the MPO are edges.

a *bond*. The  $\chi_i$ 's are called the MPO bond dimensions, and they denote the size of the  $\hat{W}$  matrices. Several algorithms have been developed for efficiently applying an MPO to an MPS, with effort of either  $\mathcal{O}(\chi^2)$  or  $\mathcal{O}(\chi^3)$  [7, 17, 18].

Two classes are of interest to us; sums of local operators (such as a Hamiltonian), and exponentials of such sums (evolution operators). We first review the structure of the former. For the bond between sites  $(i, i + 1)$  that divides the system into regions  $L_i$  and  $R_i$ , any Hamiltonian  $H$  can be decomposed as

$$H = H_{L_i} \otimes \mathbb{1}_{R_i} + \mathbb{1}_{L_i} \otimes H_{R_i} + \sum_{a_i=1}^{N_i} h_{L_i, a_i} \otimes h_{R_i, a_i}. \quad (4)$$

Here  $H_{L_i}/R_i$  are the components of the Hamiltonian localized purely to the left/right of the bond, while the  $h_{L_i, a_i} \otimes h_{R_i, a_i}$  run over  $N_i$  interaction terms which cross the bond. There is a recursion between the decompositions on bond  $(i - 1, i)$  and  $(i, i + 1)$ , which differ by the addition of site  $i$ :

$$\begin{pmatrix} H_{R_{i-1}} \\ h_{R_{i-1}, a_{i-1}} \\ \mathbb{1}_{R_{i-1}} \end{pmatrix} = \begin{matrix} 1 & N_i & 1 \\ N_{i-1} & & 1 \end{matrix} \begin{pmatrix} \hat{\mathbb{1}} & \hat{C} & \hat{D} \\ 0 & \hat{A} & \hat{B} \\ 0 & 0 & \hat{\mathbb{1}} \end{pmatrix}_{(i)} \otimes \begin{pmatrix} H_{R_i} \\ h_{R_i, a_i} \\ \mathbb{1}_{R_i} \end{pmatrix}. \quad (5)$$

Here  $(\hat{A}, \hat{B}, \hat{C}, \hat{D})_{(i)}$  are matrices of operators acting on site  $i$ , with dimensions indicated on the border. This recursion is in fact the MPO: the block matrix in the middle is  $\hat{W}_{(i)}$ , with size  $\chi_i = N_i + 2$ . (See App. A for explicit examples of MPOs.) The optimal  $(\hat{A}, \hat{B}, \hat{C}, \hat{D})_{(i)}$  can be obtained using the block Hankel singular value decomposition, a well known technique in control theory known as balanced model reduction [29].

We can view the recursion relation of Eq. (5) as a finite state machine [30]; the transitions of the machine sequentially place the operators at each site, as illustrated in Fig. 3a. The first/last indices of the MPO, which we denote by L/R respectively, play a special role, as they indicate that no non-trivial operators have been placed to the left/right of the bond. Due to the block-triangular structure of  $\hat{W}$ , once the MPO state transitions into the first index L, it remains there in perpetuity, placing only the identity operator  $\hat{\mathbb{1}}$  with each  $\hat{W}$ . The transition from R to L (not necessarily in one step) places some

local operator  $H_x$ ; the sum over all such paths generates the Hamiltonian.

### III. TIME EVOLUTION OPERATORS

Given the decomposition  $H = \sum_x H_x$ , our goal is to find an efficient MPO for

$$U(t) = 1 + t \sum_x H_x + \frac{1}{2} t^2 \sum_{x,y} H_x H_y + \dots \quad (6)$$

In the most general case, an approximation for  $U(t)$  is necessary, which brings us to our main result.

While the local Euler step defined in Eq. (1) does not have a simple MPO representation, a slight modification does. Let us define  $x < y$  if the sites affected by  $H_x$  are strictly to the left of those affected by  $H_y$ . Consider an evolution operator which keeps all non-overlapping terms:

$$U^I(t) = 1 + t \sum_x H_x + t^2 \sum_{x < y} H_x H_y + t^3 \sum_{x < y < z} H_x H_y H_z + \dots \quad (7)$$

These contributions are a subset of Eqs. (1) and (6). The first error occurs at order  $t^2$ , for terms  $H_x, H_y$  which overlap. For a system of length  $L$ , there are  $\mathcal{O}(L)$  such terms, so the error is  $\mathcal{O}(Lt^2)$ ; a constant error per *site*. Remarkably,  $U^I$  has an exact compact MPO description “ $W^I$ ”, and is trivial to construct from the  $(A, B, C, D)$  of  $H$ , illustrated in Fig. 3b. It has a block structure of total dimension  $\chi_i = N_i + 1$ :

$$\hat{W}_{(i)}^I(t) = \begin{matrix} 1 & N_i \\ N_{i-1} & \end{matrix} \begin{pmatrix} \hat{\mathbb{1}}_{(i)} + t\hat{D}_{(i)} & \sqrt{t}\hat{C}_{(i)} \\ \sqrt{t}\hat{B}_{(i)} & \hat{A}_{(i)} \end{pmatrix}. \quad (8)$$

The MPO bond dimension of  $\hat{W}^I$  is one less than that of  $H$ , so it can be simply constructed and efficiently applied to the state.

While  $\hat{W}^I$  is trivial to construct and performs well, it is not quite optimal. For example, a purely onsite Hamiltonian has a trivial  $\chi = 1$  MPO representation for  $e^{tH}$ . Yet the MPO constructed from  $\hat{W}^I$  would only produce the approximation  $U^I = \prod_x (1 + tH_x)$  in this case.

We can improve Eq. (7), by keeping terms which may overlap by one site. Let  $\langle x, \dots, z \rangle$  denote a collection of terms in which no two cross the same bond. Arbitrarily high powers of a single site term, for example, can appear in these collections. Consider an evolution operator which keeps all such terms:

$$U^{II}(t) = 1 + t \sum_x H_x + \frac{t^2}{2} \sum_{\langle x,y \rangle} H_x H_y + \frac{t^3}{6} \sum_{\langle x,y,z \rangle} H_x H_y H_z + \dots \quad (9)$$

Again, the first error occurs at  $t^2$ , with  $L$  such terms, so the error is still formally  $\mathcal{O}(Lt^2)$ . But for typical interactions far

fewer terms are dropped than in  $U^I$ ; in particular onsite terms are captured to *all* orders. While there is not an exact compact MPO representation for  $U^II$ , we can construct an MPO approximation  $\hat{W}^II$  which differs from  $U^II$  at  $\mathcal{O}(Lt^3)$ . Because the difference is at higher order than the accuracy of  $U^II$ ,  $\hat{W}^II$  still gives a noticeably better approximation than  $\hat{W}^I$ , and retains the feature that an onsite interaction is kept exactly. As  $\hat{W}^II$  is just as compact as  $\hat{W}^I$ , it is always the preferred choice.

The MPO  $\hat{W}^II$  is more complicated to construct, so for a detailed derivation of  $\hat{W}^II$  we refer to App. B. Here we discuss a concrete algorithm to calculate the MPO. As before the MPO takes the block form

$$\hat{W}^II = \begin{matrix} & & 1 & & N_i \\ & & \hat{W}_D^II & \hat{W}_C^II & \\ & & \hat{W}_B^II & \hat{W}_A^II & \\ & & & & \end{matrix} \quad (10)$$

To define the sub-blocks, introduce two vectors of formal parameters  $\phi_a, \bar{\phi}_b$ , with  $a = 1, \dots, N_{i-1}$ ,  $b = 1, \dots, N_i$ . Let  $\phi \cdot \hat{A} \cdot \bar{\phi}$  denote a dot product of these formal parameters into the MPO indices of  $\hat{A}$ , and likewise for  $\phi \cdot \hat{B}$  and  $\hat{C} \cdot \bar{\phi}$ . The sub-blocks are defined by a Taylor expansion of an operator-exponential in terms of  $\phi, \bar{\phi}$ ,

$$\begin{aligned} & e^{\phi \cdot \hat{A} \cdot \bar{\phi} + \phi \cdot \hat{B} \cdot \sqrt{t} + \sqrt{t} \hat{C} \cdot \bar{\phi} + t \hat{D}} \\ &= \hat{W}_D^II + \hat{W}_C^II \cdot \bar{\phi} + \phi \cdot \hat{W}_B^II + \phi \cdot \hat{W}_A^II \cdot \bar{\phi} + \dots \end{aligned} \quad (11)$$

Notice  $\hat{W}_D^II = e^{t\hat{D}}$  is simply the onsite term, which is kept exactly. For certain cases where the Hamiltonian is free, the above expansion can be computed analytically using Pfaffians or permanents for fermionic and bosonic theories respectively. Here we discuss only the most general case, where the result must be obtained numerically.

Dropping the labels *II*, let's compute the block  $\hat{W}_{A;a\bar{a}}$ , where  $a, \bar{a}$  index the rows and columns in correspondence with  $\phi_a, \bar{\phi}_{\bar{a}}$ . Since we only need to compute the Taylor expansion to leading order, we can consider  $\phi_a, \bar{\phi}_{\bar{a}}$  to be formal objects defined by the fact that  $\phi_a^2 = \bar{\phi}_{\bar{a}}^2 = 0$  and that they commute with all other objects. For computational purposes, we can then represent  $\phi_a$  as a hard-core boson creation operator  $\phi_a \rightarrow \hat{c}_a^\dagger$ , and likewise  $\bar{\phi}_{\bar{a}} \rightarrow \hat{c}_{\bar{a}}^\dagger$ , restricted to an occupation of at most 1  $c$ -type and 1  $\bar{c}$ -type boson. We denote the hard-core Hilbert space of the  $c/\bar{c}$  type bosons by  $\mathcal{H}_{c/\bar{c}}$ , and  $\mathcal{H}_{\text{phys}}$  the Hilbert space of the physical site. The desired entries of  $\hat{W}_A$ , which are *operators* in  $\mathcal{H}_{\text{phys}}$ , can be obtained by calculating a vacuum expectation values in the Hilbert space of the  $\mathcal{H}_{c/\bar{c}}$  coupled to the physical site:

$$\begin{aligned} \hat{W}_{A;a\bar{a}} &= \langle 0, \bar{0} | c_a \bar{c}_{\bar{a}} e^{c^\dagger \cdot \hat{A} \cdot \bar{c}^\dagger + c^\dagger \cdot \hat{B} \cdot \sqrt{t} + \sqrt{t} \hat{C} \cdot \bar{c}^\dagger + t \hat{D}} | 0, \bar{0} \rangle \\ &= \langle 0, \bar{0} | c_a \bar{c}_{\bar{a}} e^{c_a^\dagger \hat{A}_{a\bar{a}} + c_a^\dagger \hat{B}_a \sqrt{t} + \sqrt{t} \hat{C}_{\bar{a}} \bar{c}_{\bar{a}}^\dagger + t \hat{D}} | 0, \bar{0} \rangle \end{aligned} \quad (12)$$

where, in the second line, there is no summation over any  $a, \bar{a}$ . To be more explicit, the argument of the exponential is an operator in the space  $\mathcal{H}_c \otimes \mathcal{H}_{\bar{c}} \otimes \mathcal{H}_{\text{phys}}$ . The desired operator-valued entry  $\hat{W}_{A;a\bar{a}}$  is the transition amplitude from the vacuum  $|0, \bar{0}\rangle$  of the  $\mathcal{H}_c \otimes \mathcal{H}_{\bar{c}}$  into the occupied state  $\langle 0, \bar{0} | c_a \bar{c}_{\bar{a}}$ .

Note that the exponential contains only creation operators, but when computing the particular entry  $\hat{W}_{A;a\bar{a}}$  we are interested only in amplitudes where only  $c_a, \bar{c}_{\bar{a}}$  are created, so we can safely truncate to the Hilbert space of two hard-core bosons  $c_a, \bar{c}_{\bar{a}}$  as well as the physical Hilbert space of a single site; if the latter dimension is  $d$ , the total dimension is  $2^2 d$ . Thus the matrix elements can be obtained by exponentiating a matrix of dimension  $4d$ , which is trivial. This is repeated for the  $N^2$  entries of  $\hat{W}_{A;a\bar{a}}$ . Results for  $\hat{W}_{B;a}$  follow as a byproduct by calculating the transition into  $\langle 0, \bar{0} | c_a$ ;  $\hat{W}_{C;\bar{a}}$  for transitions into  $\langle 0, \bar{0} | \bar{c}_{\bar{a}}$ ; and  $\hat{W}_D$  from transitions into  $\langle 0, \bar{0} |$ . All together,  $\hat{W}^II$  can be computed with complexity  $\mathcal{O}(N^2 d^3)$ , where  $N$  is the dimension of the MPO.

We also note that  $H$  has many different MPO representations, and at 2<sup>nd</sup>-order  $\hat{W}^II$  is not invariant under different choices. This choice can be exploited to further reduce errors (cf. App. D). Finally, if  $H$  is a sum of commuting terms, there is an analytic MPO representation for  $e^{tH}$  given in App. B.

#### IV. HIGHER ORDER APPROXIMATIONS

As with TEBD, we want to construct approximations with errors at higher order  $\mathcal{O}(Lt^p)$  in  $t$ , which allow one to use much larger time steps. In fact, simply by cycling through a carefully chosen set of step constants  $\{t_a\}$  we can obtain approximations of arbitrarily high order. MPS compression can be applied between steps so that the over-all complexity of the algorithm increases only linearly in the number of stages, though higher-order approximations presumably depend more sensitively on the accuracy of the intervening compression. In particular, there is a 2<sup>nd</sup>-order approximation which alternates between two complex time steps  $t_1, t_2$ .

Each stage of the approximation should have a compact MPO expression (otherwise the increased complexity cancels the gains of a larger time step), so we consider an ansatz of the form

$$W^II(t_1)W^II(t_2) \dots W^II(t_n) = U(t) + \mathcal{O}(Lt^p), \quad (13)$$

where  $p-1$  is the approximation order. The number of terms  $n$  will depend on the desired order  $p$ , and our goal is to determine a set of step constants  $\{t_a\}$  which produce the desired order. For example, to find a 2<sup>nd</sup>-order step ( $p=3$ ), we expand Eq. (13) order by order and find constraints

$$\sum_{a=1}^n t_a = t, \quad \sum_{a<b}^n t_a t_b = \frac{1}{2} t^2, \quad \sum_{a=1}^n t_a^2 = 0 \quad (14)$$

which can be solved by  $n=2$ ,  $t_1 = \frac{1+i}{2} t$ ,  $t_2 = \frac{1-i}{2} t$ . Thus, by alternating between two compact MPOs,  $W^II(t_1)$  and  $W^II(t_2)$ , we obtain a 2<sup>nd</sup>-order approximation. The same result folds for  $W^I$ . One can continue to arbitrary order; a set of 4  $t_a$ 's is required at 3<sup>rd</sup> order, a set of 7 at 4<sup>th</sup> order. As shown in Fig. 2, the 2<sup>nd</sup>-order behavior is preserved even when truncation to the MPS ansatz intervenes between steps, so the 2<sup>nd</sup>-order time step is no more demanding than the 1<sup>st</sup>-order one.

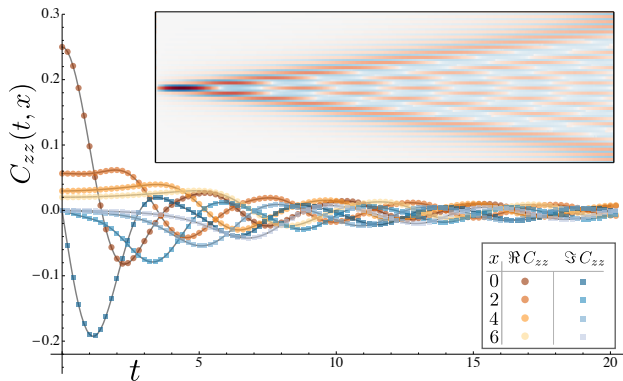


FIG. 4. Time evolution of the response function  $C_{zz}(t, x) = \langle 0|S^z(t, x)S^z(0, 0)|0\rangle$  for the Haldane-Shastry model. Discrete data points are evaluated numerically using the 2<sup>nd</sup>-order MPO time stepper  $W^H$  ( $dt = 0.025$ ), shown here for positions  $x = 0, 2, 4, 6$ . The exact analytic prediction shown in solid curves, giving beautiful agreement with the MPO. The inset shows a density plot of  $C_{zz}(t, x)$  in the  $t$ - $x$  plane.

## V. APPLICATIONS

Our first system beyond the reach of TEBD is the spin-1/2 Haldane-Shastry model, an exactly solvable critical spin chain with long-ranged Hamiltonian

$$H_{\text{HS}} = \sum_{x,r>0} \frac{\mathbf{s}_x \cdot \mathbf{s}_{x+r}}{r^2}. \quad (15)$$

The model can be viewed as a lattice form of the Calogero-Sutherland continuum model of fractional statistics [31, 32] and is connected to the Laughlin fractional quantum Hall wavefunction with an exact MPS representation [33]. The dynamical correlation function  $C_{zz}(t, x)$  was first calculated analytically by Haldane and Zirnbauer [28]. As the system is critical and the Hamiltonian long-ranged, numerically obtaining  $C_{zz}$  is a stringent test of the proposed method. We use an MPO approximation of the Hamiltonian to capture the  $r^{-2}$  power law with high accuracy out to about 200 sites [34]. After using infinite DMRG [34–36] to obtain the ground state with infinite boundary conditions, we act with  $S^z$  and time evolve via  $W^H$ . As described in Fig. 4, the numerically com-

puted  $C_{zz}$  is nearly identical to the analytic prediction (cf. App. E) out to significant time scales.

Finally, one of the most interesting potential applications is time-evolving finitely-entangled 2D systems. We make a preliminary study by considering the 2D Bose-Hubbard model with a hard-core constraint and nearest neighbor repulsion  $V$ . Recently there have been several experimental and theoretical studies of the expansion of strongly-interacting clouds [11, 12, 37–39]. The repulsion  $V$  can generate many-body bound states if it exceeds the bandwidth  $J$ , because there is no way for the interaction energy to transform into kinetic energy. Here we let a 16-boson  $n = 1$  product state expand into a  $14 \times 14$  grid. As shown in Fig. 1, the repulsion  $V$  has a dramatic effect on the expansion, trapping the bosons into a bound state. Because the 2D lattice has been turned into a 1D chain, the errors in  $W^H$  are highly anisotropic. Nevertheless we find that with a time step  $dt = 0.01$ , the density remains rotationally symmetric to within 4% at  $t = 2$ .

## VI. CONCLUSION

We have introduced an MPO based algorithm to simulate the time-evolution of long-ranged Hamiltonians. Our method was benchmarked against existing numerical methods for 1D short ranged models, as well as analytic results for the long-ranged Haldane-Shastry model. We also presented results of a preliminary study of the expansion of interacting bosons in a 2D trap. Given the recent successes of DMRG for investigating gapped 2D ground state and their gapless edges, the techniques presented here could open the door to numerically calculating experimentally relevant dynamic quantities such as spectral functions.

Subsequent to the initial publication of this work, an improved version of the TDVP was proposed.[40]

## ACKNOWLEDGMENTS

We are grateful to J. H. Bardarson, E. M. Stoudenmire, D. Varjas for helpful conversations. The authors wish to thank NSF DMR-1206515 (M.Z. and J.E.M.), the Sherman Fairchild Foundation (R.M.), the Nanostructured Thermoelectrics program of DOE BES (C.K.), and the Simons Foundation (J.E.M.).

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).  
 [2] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).  
 [3] S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).  
 [4] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, Journal of Statistical Mechanics: Theory and Experiment **2004**, P04005 (2004).  
 [5] P. Schmitteckert, Phys. Rev. B **70**, 121302 (2004).  
 [6] E. Jeckelmann, Progress of Theoretical Physics Supplement **176**, 143 (2008).

[7] U. Schollwöck, Annals of Physics **326**, 96 (2011).  
 [8] E. Stoudenmire and S. R. White, Annual Review of Condensed Matter Physics **3**, 111 (2012).  
 [9] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **101**, 250602 (2008).  
 [10] P. Corboz, S. R. White, G. Vidal, and M. Troyer, Phys. Rev. B **84**, 041108 (2011).  
 [11] K. Winkler, G. Thalhammer, F. Lang, R. Grimm, J. Hecker Denschlag, A. J. Daley, A. Kantian, H. P. Büchler, and P. Zoller, Nature (London) **441**, 853 (2006).

- [12] J. P. Ronzheimer, M. Schreiber, S. Braun, S. S. Hodgman, S. Langer, I. P. McCulloch, F. Heidrich-Meisner, I. Bloch, and U. Schneider, *Phys. Rev. Lett.* **110**, 205301 (2013).
- [13] M. Fannes, B. Nachtergaele, and R. W. Werner, *Commun. Math. Phys.* **144**, 443 (1992).
- [14] S. Östlund and S. Rommer, *Phys. Rev. Lett.* **75**, 3537 (1995).
- [15] S. Rommer and S. Östlund, *Phys. Rev. B* **55**, 2164 (1997).
- [16] F. Verstraete, D. Porras, and J. I. Cirac, *Phys. Rev. Lett.* **93**, 227205 (2004).
- [17] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, *New Journal of Physics* **12**, 025012 (2010).
- [18] E. M. Stoudenmire and S. R. White, *New Journal of Physics* **12**, 055026 (2010).
- [19] A. E. Feiguin and S. R. White, *Phys. Rev. B* **72**, 020404 (2005).
- [20] J. J. García-Ripoll, *New Journal of Physics* **8**, 305 (2006).
- [21] M. L. Wall and L. D. Carr, *New Journal of Physics* **14**, 125015 (2012).
- [22] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [23] J. J. Dorando, J. Hachmann, and G. K.-L. Chan, *The Journal of Chemical Physics* **130**, 184111 (2009).
- [24] T. Koffel, M. Lewenstein, and L. Tagliacozzo, *Phys. Rev. Lett.* **109**, 267203 (2012).
- [25] P. Hauke and L. Tagliacozzo, *Phys. Rev. Lett.* **111**, 207202 (2013).
- [26] F. D. M. Haldane, *Phys. Rev. Lett.* **60**, 635 (1988).
- [27] B. S. Shastry, *Phys. Rev. Lett.* **60**, 639 (1988).
- [28] F. D. M. Haldane and M. R. Zirnbauer, *Phys. Rev. Lett.* **71**, 4055 (1993).
- [29] S. Kung and D. Lin, *Automatic Control, IEEE Transactions on* **26**, 832 (1981).
- [30] G. M. Crosswhite and D. Bacon, *Phys. Rev. A* **78**, 012356 (2008).
- [31] F. Calogero, *J. Math. Phys.* **10**, 2191 (1969).
- [32] B. Sutherland, *Phys. Rev. A* **4**, 2019 (1971).
- [33] J. I. Cirac and G. Sierra, *Phys. Rev. B* **81**, 104431 (2010).
- [34] G. M. Crosswhite, A. C. Doherty, and G. Vidal, *Phys. Rev. B* **78**, 035116 (2008).
- [35] I. P. McCulloch, “Infinite size density matrix renormalization group, revisited,” (2008), arXiv:0804.2509.
- [36] J. A. Kjäll, M. P. Zaletel, R. S. K. Mong, J. H. Bardarson, and F. Pollmann, *Phys. Rev. B* **87**, 235106 (2013).
- [37] M. Ganahl, M. Haque, and H. G. Evertz, (2013), unpublished, arXiv:1302.2667.
- [38] L. Vidmar, S. Langer, I. P. McCulloch, U. Schneider, U. Schollwöck, and F. Heidrich-Meisner, *Phys. Rev. B* **88**, 235117 (2013).
- [39] A. Reinhard, J.-F. Riou, L. A. Zundel, D. S. Weiss, S. Li, A. M. Rey, and R. Hipolito, *Phys. Rev. Lett.* **110**, 033001 (2013).
- [40] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, (2014), arXiv:1408.5056.

## Appendix A: MPO examples

In this section, we provide explicit examples of MPOs for pedagogical purposes.

To reiterate from the main text, an MPO describes an operators written as a product of  $\hat{W}$ 's

$$\cdots \hat{W}_{(1)} \hat{W}_{(2)} \hat{W}_{(3)} \cdots, \quad (\text{A1})$$

where each  $\hat{W}_{(i)}$  is a matrix of operators acting on site  $i$ . An MPO for a Hamiltonian can always be casted in the form

$$\hat{W}_{(i)} = \begin{matrix} & & 1 & N_i & 1 \\ & 1 & \hat{A} & \hat{C} & \hat{D} \\ & N_{i-1} & \begin{pmatrix} \hat{A} & \hat{C} & \hat{D} \\ 0 & \hat{A} & \hat{B} \\ 0 & 0 & \hat{A} \end{pmatrix} & & \\ & 1 & & & \end{matrix} \Big|_{(i)}. \quad (\text{A2})$$

$\hat{D}$  is simply an operator,  $\hat{C}$  and  $\hat{B}$  are, respectively, a row and column vector, an  $\hat{A}$  is an  $N_{i-1} \times N_i$  matrix of operators.

Consider the transverse field Ising model with Hamiltonian

$$H_{\text{TFI}} = -J \sum_i \hat{Z}_i \hat{Z}_{i+1} - h \sum_i \hat{X}_i, \quad (\text{A3})$$

where  $\hat{X}$  and  $\hat{Z}$  are Pauli operators. This Hamiltonian may be constructed as an MPO with

$$\hat{W}_{(i)} = \begin{pmatrix} \hat{1} & \hat{Z} & -h\hat{X} \\ 0 & 0 & -J\hat{Z} \\ 0 & 0 & \hat{1} \end{pmatrix} \Big|_{(i)}. \quad (\text{A4})$$

Hence  $N_i = 1$  for all bonds, and the MPO has bond dimension  $\chi_i = 3$ . We can also read off the  $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$  operators as  $(0, -J\hat{Z}, \hat{Z}, -h\hat{X})$ . We note that this MPO is not unique for Hamiltonian Eq. (A3) (cf. App. D). Due to the absence of  $\hat{A}$ , the Hamiltonian consists of only onsite and nearest-neighbor terms. Here  $\hat{D}$  always denote the onsite term, and the pair terms are given by  $\hat{C}_i \hat{B}_{i+1}$ .

Our second example is a long-ranged XY-chain, with exponentially decaying couplings.

$$H = J \sum_{i < j} e^{-\alpha|i-j|} (\hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j). \quad (\text{A5})$$

Now using the MPO recursion of Eq. (5), we can peel off one site:

$$H_{R_i} + \phi_i \cdot h_{R_i} = \phi_i \cdot \hat{A}_{i+1} \cdot h_{R_{i+1}} + \phi_i \cdot \hat{B}_{i+1} + \hat{C}_{i+1} \cdot h_{R_{i+1}} + \hat{D}_{i+1} + H_{R_{i+1}}. \quad (\text{B3})$$

Thus if we introduce a new vector of fields  $\phi_{i+1, a_{i+1}}$  which runs over  $a_{i+1} = 1, \dots, N_{i+1}$ , we can write

$$e^{\phi_i \cdot h_{R_i} + H_{R_i}} = \int \mathcal{D}[\phi_{i+1}, \bar{\phi}_{i+1}] \hat{U}_{\phi_i, \bar{\phi}_{i+1}} e^{-\bar{\phi}_{i+1} \phi_{i+1}} e^{\phi_{i+1} \cdot h_{R_{i+1}} + H_{R_{i+1}}}, \quad (\text{B4})$$

where  $\hat{U}_{\phi_i, \bar{\phi}_{i+1}} \equiv e^{\phi_i \cdot \hat{A}_{i+1} \cdot \bar{\phi}_{i+1} + \phi_i \cdot \hat{B}_{i+1} + \hat{C}_{i+1} \cdot \bar{\phi}_{i+1} + \hat{D}_{i+1}}$ .

A corresponding MPO with  $N_i = 2$  is as follows,

$$\hat{W}_{(i)} = \begin{pmatrix} \hat{1} & e^{-\alpha} \hat{X} & e^{-\alpha} \hat{Y} & 0 \\ 0 & e^{-\alpha} \hat{1} & 0 & J\hat{X} \\ 0 & 0 & e^{-\alpha} \hat{1} & J\hat{Y} \\ 0 & 0 & 0 & \hat{1} \end{pmatrix} \Big|_{(i)}. \quad (\text{A6})$$

Here  $\hat{A}$  is a non-trivial  $2 \times 2$  matrix of operators, which allows terms to reach beyond two neighboring sites. Each insertion of the  $\hat{A}$  matrix increases the separation of the bookends  $\hat{X}/\hat{Y}$  by 1 site, and also reduces its amplitude by  $e^{-\alpha}$  factor.

## Appendix B: Exact MPO exponentiation for commuting Hamiltonians

Here we obtain the exact MPO description for  $e^H$  when  $H$  is a sum of commuting terms such as  $\sum_{i,j} \hat{X}_i \hat{X}_j t_{ij}$ . This result generalizes the nearest-neighbor case investigated in Ref. 17. to long range interactions. First, we address the stricter case in which  $\hat{A}, \hat{B}, \hat{C}, \hat{D}$  all commute, then comment briefly on the more general case.

Suppose the data  $(A, B, C, D)_{(i)}$  of the MPO representation for  $H$  is given, with bond dimensions  $\chi_i = 2 + N_i$ . On each bond  $(i, i+1)$ , introduce a vector of complex fields  $\phi_i = (\phi_{i,1}, \dots, \phi_{i,N_i})$ , with complex conjugate  $\bar{\phi}_i$  and indices  $a_i = 1, \dots, N_i$  in correspondence with the non-trivial MPO indices in  $H$ . (That is, any MPO indices that is not L or R.) Using the fundamental rule of complex Gaussian integrals,

$$\frac{1}{\pi} \int d^2 \phi e^{-\bar{\phi} \phi + J \bar{\phi} + \phi \bar{J}} = e^{J \bar{J}}, \quad (\text{B1})$$

the exponential factors as

$$e^H = \int \mathcal{D}[\phi_i, \bar{\phi}_i] e^{H_{L_i} + h_{L_i} \cdot \bar{\phi}_i} e^{-\bar{\phi}_i \cdot \phi_i} e^{\phi_i \cdot h_{R_i} + H_{R_i}} \quad (\text{B2})$$

where the dot-product is the sum  $\sum_{a_i=1}^{N_i}$ , and  $\mathcal{D}[\phi_i, \bar{\phi}_i]$  is shorthand for  $\prod_{a_i} (d^2 \phi_{i,a_i} / \pi)$ . This identity requires that all terms commute; otherwise discrepancies arise at second-order in  $H$ .



By repeating this step on all the bonds, we find

$$e^H = \int \mathcal{D}[\phi, \bar{\phi}] \left[ \cdots e^{-\bar{\phi}_i \phi_i} \hat{U}_{\phi_i, \bar{\phi}_{i+1}} e^{-\bar{\phi}_{i+1} \phi_{i+1}} \hat{U}_{\phi_{i+1}, \bar{\phi}_{i+2}} \cdots \right]. \quad (\text{B5})$$

This is a matrix product operator in which the auxiliary bonds are labeled by a set of continuous numbers  $\phi_i$ , rather than discrete indices; it is a ‘‘coherent state MPO.’’ To bring the result to a discrete form, we note that an integral of the form Eq. (B5) is a discretized coherent state path integral for  $N_i$  bosons, so the integrals can be converted to discrete sums over the *many-body* Hilbert space of  $N_i$  bosons. The basic manipulation is the Taylor expansion:

$$Y_\phi \equiv \sum_{n=0}^{\infty} Y_n \frac{\phi^n}{\sqrt{n!}} \quad (\text{and likewise for any tensor}) \quad (\text{B6})$$

$$\begin{aligned} \frac{1}{\pi} \int d^2\phi X_{\bar{\phi}} e^{-\bar{\phi}\phi} Y_\phi &= \frac{1}{\pi} \sum_{\bar{n}, n} X_{\bar{n}} Y_n \int d^2\phi \frac{\bar{\phi}^{\bar{n}} \phi^n}{\sqrt{\bar{n}! n!}} e^{-\bar{\phi}\phi} \\ &= \sum_n X_n Y_n \end{aligned} \quad (\text{B7})$$

The integer  $n$  is the ‘occupation.’ Note that if a tensor depends on multiple variables (such as the vector  $\phi_{i, a_i}$ ), then the above rule extends via a simple product. So if we define a vector of occupations  $n_i = (n_{i,1}, \dots, n_{i, N_i})$ , whose values index the Hilbert space of  $N_i$  bosons, we can Taylor expand  $U$  as

$$\hat{U}_{\phi_i, \bar{\phi}_{i+1}} \equiv \sum_{\{n_i\}, \{\bar{n}_i\}} \hat{U}_{n_i, \bar{n}_i} \frac{\phi_i^{n_i} \bar{\phi}_{i+1}^{\bar{n}_i}}{\sqrt{|n_i!| |\bar{n}_i+1!|}} \quad (\text{B8})$$

with  $|n_i| = \prod_{a_i} (n_{i, a_i})!$ . The MPO for the exponential is

$$e^H = \sum_{\{n_i\}} \left[ \cdots \hat{U}_{n_i, n_{i+1}} \hat{U}_{n_{i+1}, n_{i+2}} \cdots \right] \quad (\text{B9})$$

Now in principle each sum on the bonds is over the many-body Hilbert space of  $N_i$  bosons, which is infinite. But there will be ‘‘Boltzmann factors’’ associated to these states which allows for a sensible truncation.

Furthermore, in certain situations, such as for a nearest-neighbor interaction of Pauli-matrices,  $H = \sum_i \hat{X}_i \hat{X}_{i+1}$ ,  $\hat{U}_{n_{i+1}, n_{i+2}}$  only has rank 2, resulting in the  $\chi = 2$  MPO reported previously [17].

We must slightly modify the procedure if the Hamiltonian is a sum of commuting terms but  $\hat{A}, \hat{B}, \hat{C}, \hat{D}$  do not commute (for instance, in the Toric code). Then on each bond we can arbitrarily order the  $\phi_a$ , and when expanding the exponential for  $\hat{U}_{\phi_i, \bar{\phi}_{i+1}}$ , order the terms accordingly.

### Appendix C: Characterizing errors

Here we make a brief remark on the correct way to characterize the errors in an approximation  $\tilde{U}(t)$  to the exact evolution  $U(t)$ . We say an approximation has an error *per site* of

order  $t^p$  if

$$U(-t)\tilde{U}(t) = e^{\sum_{j=p}^{\infty} t^j O_j} \sim e^{L(a_p t^p + a_{p+1} t^{p+1} + \cdots)} \quad (\text{C1})$$

where each operator  $O_j$  is a sum of *local* terms, so its spectral radius goes as  $|O_j| \sim L$ . One can see that Suzuki-Trotter and  $W^I$  follow this form, while global Runge-Kutta does not. The locality of the operator generating the error implies that (after normalizing the state) the error in local observables should scale as  $t^p$ .

Now consider the expansion of  $U(t) - \tilde{U}(t)$ , which is more natural to calculate:

$$U(t) - \tilde{U}(t) \sim \sum_{nm} E_{nm} L^n t^m \quad (\text{C2})$$

where  $L^n t^m$  denotes a term of spectral radius  $L^n$  and order  $t^m$ . For an approximation with constant error  $t^p$  per site, we see that  $t_{nm} \neq 0$  only if  $m - n \geq p - 1$ . Order  $p - 1$  Runge-Kutta has errors of the form  $L^m t^m$  for  $m \geq p$ , which violates this constraint.

The proposed form for  $W^II$  was the expansion

$$\begin{aligned} e^{\phi \cdot \hat{A} \cdot \bar{\phi} + \phi \cdot \hat{B} \cdot \sqrt{t} + \sqrt{t} \hat{C} \cdot \bar{\phi} + t \hat{D}} \\ = \hat{W}_D^II + \hat{W}_C^II \cdot \bar{\phi} + \phi \cdot \hat{W}_B^II + \phi \cdot \hat{W}_A^II \cdot \bar{\phi} + \cdots \end{aligned} \quad (\text{C3})$$

Comparing Eq. (C3) with Eq. (B4), we see that  $\hat{W}^II$  is precisely a truncation of  $\hat{U}_{n_i, n_{i+1}}$  to an occupation of at most a single boson on each bond. The occupation number of bosons across a bond encodes the number of terms in the Hamiltonian which cross the bond in the Taylor expansion of  $e^{\sum_x H_x}$ . Hence by truncating  $\hat{U}$  to a maximum occupation of 1, we keep all non bond-overlapping terms. However, in the derivation of the exact MPO  $e^{tH}$  we required all terms to commute. Careful inspection shows that the non-commutivity only shows up at 3<sup>rd</sup>-order in  $H$ . Hence in general  $\hat{W}^II$  is only an approximation to the sum of all non bond-overlapping terms, with errors at  $\mathcal{O}(t^3)$ . But these errors are subleading in comparison to the terms dropped (by the truncation) at  $\mathcal{O}(t^2)$ , so are unimportant.

### Appendix D: Taking advantage of different MPO decompositions

There are numerous ways to decompose a Hamiltonian as  $H = \sum_x H_x$ , and hence many decompositions into an MPO. For instance, a ferromagnetic interaction can be written as

$$\begin{aligned} H_F &= - \sum_i \hat{Z}_i \hat{Z}_{i+1} \\ &= - \sum_i \left[ (\hat{Z}_i - h)(\hat{Z}_{i+1} - h) + 2h\hat{Z}_i - h^2 \right] \end{aligned} \quad (\text{D1})$$

with MPO

$$\hat{W}^{H_F} = \begin{pmatrix} 1 & -\hat{Z} & 0 \\ 0 & 0 & \hat{Z} \\ 0 & 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 1 & -(\hat{Z} - h) & h^2 - 2h\hat{Z} \\ 0 & 0 & (\hat{Z} - h) \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{D2})$$

The MPO  $\hat{W}^H$  is *not* invariant under such shifts (at 2<sup>nd</sup>-order). This can be used to improve the effective accuracy of  $\hat{W}^H$ .

In principle one could try to optimize over all the MPO representations of  $H$  in order to minimize the error in  $\hat{W}^H$ . It is an open question whether there is a practical method to do this. As a toy model we compute the error  $|(U(dt) - W^H(dt))|\psi\rangle|$  for the ferromagnet  $H_F$  as a function of the shift  $h$  given in Eq. (D2). To leading order,

$$\begin{aligned} & \left| (U(dt) - W^H(dt))|\psi\rangle \right|^2 \\ & \propto dt^2 \sum_i \langle \psi | (\hat{Z}_i - h)^2 (\hat{Z}_{i+1} - h)^2 | \psi \rangle, \end{aligned} \quad (\text{D3})$$

since  $W^H$  drops these two-site terms at 2<sup>nd</sup>-order. So, in principle, the optimal  $h$  minimizes this expression.

One possible *heuristic* is to make a mean field approximation and instead minimize  $\langle (\hat{Z}_i - h)^2 \rangle \langle (\hat{Z}_{i+1} - h)^2 \rangle$  by setting  $h = \langle \hat{Z}_i \rangle$ . With this choice the onsite term of Eq. (D2) is  $\hat{D} = h^2 - 2h\hat{Z}$ , the mean field Hamiltonian. Since  $W^H$  treats  $\hat{D}$  exactly, it's not surprising this can reduce the error.

To generalize this heuristic mean field criteria, we can always choose the MPO for  $H$  such that the Hamiltonian cut across any bond (cf. Eq. 4) satisfies  $\langle h_{R_i, a_i} \rangle = \langle h_{L_i, a_i} \rangle = 0$  by shuffling the mean field component into  $H_{L_i}, H_{R_i}$ . Then the errors in  $\hat{W}^H$  at 2<sup>nd</sup>-order will depend only on the connected part of  $\sum_{a_i} h_{L_i, a_i} h_{R_i, a_i}$ . For many relevant models, such as a Heisenberg spin model, this heuristic does not help since  $\langle h_{R_i, a_i} \rangle = 0$  due to the SU(2) symmetry of  $\mathbf{S}$ . But for a model with a long-ranged density-density interaction like

$\frac{1}{2} \sum_{x,y} n_x V(x-y) n_y$ , the mean field approach will treat the 'direct' part of the evolution,  $\sum_{x,y} n_x V(x-y) \langle n_y \rangle$ , exactly.

### Appendix E: Analytical expressions for dynamical correlation functions of Haldane-Shastry spin chain

We provide here the expression found by Haldane and Zirn-bauer [28] for the ground-state dynamical correlations

$$G_{mn}^{ab}(t, t') \equiv \langle 0 | S_m^a(t) S_n^b(t') | 0 \rangle \quad (\text{E1})$$

of the Haldane-Shastry spin chain [26, 27] with Hamiltonian

$$H_{\text{HS}} = J \sum_{m < n, a} \frac{S_m^a S_n^a}{|m - n|^2}. \quad (\text{E2})$$

(The superscript of  $S$  operators denote the spin direction and the subscript denote the lattice site.) The arguments leading to the forms below are somewhat involved and we refer the reader to the original paper for details. In the following  $\hbar = 1$ .  $G_{mn}^{ab}$  is diagonal in spin indices, and translation invariance allows us to define

$$G_{mn}^{ab}(t, t') = \frac{1}{4} \delta^{ab} (-1)^{m-n} C(m - n, t - t'). \quad (\text{E3})$$

The function  $C(x, t)$  is related to the spinon spectrum in the solution for the ground-state wavefunction and can be simplified to two integrals:

$$C(x, t) = \frac{1}{4} \int_{-1}^1 d\lambda_1 \int_{-1}^1 d\lambda_2 e^{i\pi\lambda_1\lambda_2x - \frac{\pi vt}{2}(\lambda_1^2 + \lambda_2^2 - 2\lambda_1^2\lambda_2^2)}. \quad (\text{E4})$$

Here  $v$  is the spinon velocity,  $v = \pi J/2$ , and the prefactor of 1/4 can be understood by noting that  $C(0, 0) = 1$  as  $(S^a)^2 = 1/4$  for each spin direction  $a$ . The numerical integrations used to obtain the comparison curves in Fig. 4 are straightforward and were carried out using commercial software.