

# Efficient representation of long-range interactions in tensor network algorithms

Matthew J. O'Rourke, Zhendong Li, and Garnet Kin-Lic Chan  
*Division of Chemistry and Chemical Engineering,  
 California Institute of Technology, Pasadena, California 91125, USA*

We describe a practical and efficient approach to represent physically realistic long-range interactions in two-dimensional tensor network algorithms via projected entangled-pair operators (PEPOs). We express the long-range interaction as a linear combination of correlation functions of an auxiliary system with only nearest-neighbor interactions. To obtain a smooth and radially isotropic interaction across all length scales, we map the physical lattice to an auxiliary lattice of expanded size. Our construction yields a long-range PEPO as a sum of ancillary PEPOs, each of small, constant bond dimension. This representation enables efficient numerical simulations with long-range interactions using projected entangled pair states.

The accurate description of strongly correlated quantum many-body systems is a major challenge in contemporary physics. Nonetheless, some of the most intriguing macroscopic quantum phenomena, such as high-temperature superconductivity and the fractional quantum Hall effect, arise from strong quantum correlations. In recent years, tensor network states (TNS) [1–6], including matrix product states (MPS) [7–10] and projected entangled-pair states (PEPS) [11–14], have emerged as promising classes of variational states to numerically approximate the low energy physics of correlated quantum systems. Their power stems from systematically improvable accuracy through increasing the tensor bond dimension  $D$ , and the  $O(A)$  linear complexity of the associated algorithms with respect to the system size  $A$ .

One promising application of TNS is to accurate calculations of electronic structure. While the electronic structure Hamiltonian can be represented in multiple ways [15–18], the simplest – and the one of interest in this work – is a real-space grid formulation [19–23],

$$\hat{H} = -t \sum_{\langle i,j \rangle} (a_{i\sigma}^\dagger a_{j\sigma} + h.c.) + \sum_i v_i^{ne} n_i + \hat{V}^{ee},$$

$$\hat{V}^{ee} = \sum_i v_{ii}^{ee} n_{i\alpha} n_{i\beta} + \sum_{i<j} v_{ij}^{ee} n_i n_j, \quad (1)$$

where  $i, j$  label lattice sites,  $\sigma \in \{\alpha, \beta\}$  labels spin, and  $a^{(\dagger)}$  and  $n$  are fermion creation, annihilation, and number operators, respectively. This representation becomes exact as the spacing between the sites goes to zero, and  $v_{ij}^{ee}$  becomes the continuum Coulomb potential  $1/r_{ij}$  with  $r_{ij} \triangleq |\mathbf{r}_i - \mathbf{r}_j|$ . This form of the electronic structure Hamiltonian is especially suited to TNS algorithms as the Coulomb interaction is a pairwise operator as opposed to a general quartic operator when using a non-local basis, and Eq. (1) thus takes the form of an extended Hubbard model with long-range terms. Ground states of such grid Hamiltonians have been computed in 1D using MPS and the density matrix renormalization group (DMRG), yielding near exact electronic structure benchmarks for small lattice spacings [1, 2, 19]. In principle, this success

in 1D should be extensible to 2D and 3D by using PEPS instead of MPS, and would then provide a route to simulate arbitrarily complex electronic structure problems with arbitrarily improvable accuracy.

However, current state-of-the-art PEPS calculations have not yet advanced beyond local lattice models in 2D [24–29]. One reason for this is the difficulty of efficiently including long-range interactions, such as the Coulomb interaction  $\hat{V}^{ee}$  appearing in (1). To see the basic challenge, consider the energy expectation value: for a Hamiltonian with localized interactions, the cost of a term-by-term calculation scales linearly with the size of the system,  $O(A)$ . However, for a Hamiltonian with long-range interactions, the term-by-term evaluation scales like  $O(A^2)$ , which is prohibitively expensive in two (or higher) dimensions, as we take the continuum limit. Alternatively, one might try to use an exact tensor network operator, or projected entangled pair operator (PEPO), to represent the long-range interaction [30], avoiding the explicit term-by-term evaluation. However, the exact PEPO representation for arbitrary long-range interactions in 2D has a bond dimension that scales as  $O(A^{1/4})$ , causing the corresponding cost to compute expectation values to again scale as  $O(A^2)$  [31].

In 1D, the increased computational cost of long-range interactions can be eliminated if they are smooth and decaying. In this case one can approximate the exact matrix product operator (MPO) by a compressed MPO of constant bond dimension  $D$  that generates a sum of exponential interactions, and smoothly decaying interactions can be approximated well by such sums [30, 32, 33]. Exponential interactions in MPOs arise naturally from the matrix product structure, which also gives rise to the exponential decay of two-point correlation functions in MPS. Extending the correlation function analogy to 2D leads to an efficient representation of long range interactions in 2D when their form exactly coincides with the correlation function of a 2D lattice model. This was demonstrated in Ref. [30], which constructed a compact pair interaction PEPO whose interaction potential was given by the critical 2D Ising correlation function.

Building on these ideas, in this work we describe how

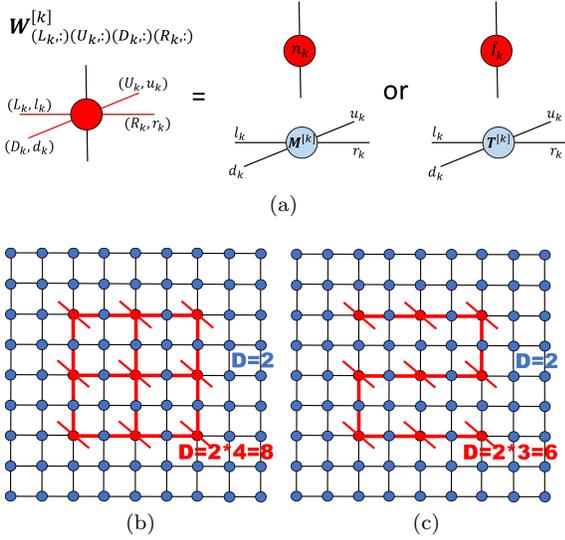


FIG. 1. (a) The construction of the nonzero parts of the CF-PEPO tensor  $\mathbf{W}^{[k]}$  via the coupling of the finite state machine (FSM) tensor (red) with the Ising correlation function tensors (blue). Note that here the physical indices of  $\mathbf{W}^{[k]}$  are explicitly shown, whereas they are suppressed in Eq. (3). (b)-(c) Two possible constructions of the long-range PEPO for a 3x3 physical system with 1 fictitious Ising site (blue) in between adjacent physical sites (red) and a 2 site buffer to help mitigate boundary effects in the encoding of the potential. Black bonds are  $D = 2$  and red bonds are  $D = 8$  (b) and 6 (c).

general long-range interactions in two dimensions, including the Coulomb interaction, can be efficiently encoded as a sum of low rank correlation function valued PEPOs. Although superficially similar to the problem of approximating a smooth interaction in 1D by a sum of exponentials, additional complications arise in two dimensions because physical interactions possess different analytic properties from two-point correlation functions on the same lattice. For example, the Coulomb interaction is radially isotropic at all distances, while the two-point lattice correlation functions are isotropic only at large distances due to the lattice discretization. We show how to overcome these and other difficulties by introducing an expanded auxiliary lattice, and demonstrate the efficiency of the representation in a ground-state finite PEPS simulation of a 2D spin model with Coulombic Heisenberg interactions. Although we specifically treat only the Coulomb interaction and two dimensions in our numerical examples, our arguments naturally extend to representing smooth and radially isotropic interactions in any dimension.

*Correlation function valued PEPOs.* — We first define correlation function valued PEPOs (CF-PEPOs), which are central to this work. As motivation, we recall the construction of MPOs for smooth interactions approximated by sums of exponentials. This is usually done in the language of finite state machines (FSM), where the incoming and outgoing bonds of each MPO tensor

are interpreted as machine states [32, 33]. An FSM can encode an exponentially decaying strength  $e^{-\lambda r_{ij}}$  via a single non-zero element in each MPO tensor with value  $e^{-\lambda}$ , that gets multiplied along the lattice as long as the FSM stays in a specified state. The pairwise operator  $\sum_{i < j} e^{-\lambda r_{ij}} n_i n_j$  can then be represented by an MPO with bond dimension 3, with the two additional states in the FSM acting to combine the exponential scalar values with the operators  $n_i n_j$ . The construction can be extended to the general 1D interaction  $\sum_{i < j} V(r_{ij}) n_i n_j \approx \sum_{i < j} \sum_{t=1}^{N_t} c_t e^{-\lambda_t r_{ij}} n_i n_j$  by introducing additional states for each of the  $N_t$  exponential decays, for a total MPO bond dimension of  $N_t + 2$  (or alternatively,  $N_t$  MPOs of bond dimension 3). However, while this representation is natural in 1D, its direct extension to 2D is not. This is because multiplying the element  $e^{-\lambda}$  along any single FSM path between two sites  $i$  and  $j$  creates an exponentially decaying strength as a function of the Manhattan distance  $|x| + |y|$ , not the desired Euclidean distance  $(x^2 + y^2)^{1/2}$ , as the elements are multiplied out along the grid lines [31].

A different starting point, that is more natural in higher dimensions, is to consider scalar interaction strengths generated by the two-point correlation function  $\langle o(\mathbf{r}_i) o(\mathbf{r}_j) \rangle_\beta$  of a classical model at inverse temperature  $\beta$ . We term the PEPO for the operator  $\sum_{i < j} \langle o(\mathbf{r}_i) o(\mathbf{r}_j) \rangle_\beta n_i n_j$ , a correlation function valued PEPO (CF-PEPO). Using a classical model with local interactions yields a CF-PEPO with low bond dimension, as noted in Ref. [30]. For example, consider the spin-spin correlation function of the 2D Ising model with  $H = -\sum_{\langle i, j \rangle} \sigma_i \sigma_j$ ,  $\sigma_i \in \{+1, -1\}$ , which can be exactly represented by the Ising PEPS with  $D = 2$  [13, 34], viz.,

$$\langle \sigma_i \sigma_j \rangle_\beta = \frac{1}{Z} \text{Tr} \left( \prod_{k \neq i, j} T_{l_k u_k d_k r_k}^{[k]} M_{l_i u_i d_i r_i}^{[i]} M_{l_j u_j d_j r_j}^{[j]} \right), \quad (2)$$

where the partition function  $Z = \text{Tr} \prod_k \mathbf{T}^{[k]}$ , and the trace is over all bond indices. The  $\mathbf{M}$  and  $\mathbf{T}$  tensors are the tensors of the PEPS on and off the correlation function sites, respectively, and depend on  $\beta$ . Given the Ising CF-PEPS, we obtain the Ising CF-PEPO by combining its tensors with the tensors of an FSM that generates all pairwise interaction terms  $n_i n_j$  on the square lattice,

$$\sum_{i < j} \langle \sigma_i \sigma_j \rangle_\beta n_i n_j = \text{Tr} \left( \prod_k \mathbf{W}_{(L_k, l_k)(U_k, u_k)(D_k, d_k)(R_k, r_k)}^{[k]} \right),$$

$$\mathbf{W}_{(L_k, :)(U_k, :)(D_k, :)(R_k, :)}^{[k]} \triangleq \mathbf{X}^{[k]},$$

$$\mathbf{X}^{[k]} \in \{0, \mathbf{T}^{[k]} \otimes \hat{I}_k, \mathbf{M}^{[k]} \otimes n_k\}. \quad (3)$$

Here  $\mathbf{W}^{[k]}$  (Fig. 1(a)) is the operator valued tensor in the new CF-PEPO,  $(L_k, l_k)$  is a composite index of the bond  $L_k$  for the 2D FSM and the bond  $l_k$  of the Ising PEPS, and  $\mathbf{X}^{[k]}$  takes one of the three values depending

on the machine state  $(L_k, U_k, D_k, R_k)$  [35]. Since the FSM tensors only need to encode the two operators  $n_i n_j$  and contain no information about the distance between them, there is some flexibility in the FSM geometry (see Fig. 1). The snake geometry in (c) has a significantly reduced computational complexity compared to the full 2D FSM in (b) [31], and it also imposes an ordering that allows for a simple way to include fermionic statistics (via Jordan-Wigner strings) at the operator level, eliminating the need for swap gates in fermionic PEPS [36]. Both of these constructions are compatible with existing iPEPS [37] algorithms.

*CF-PEPOs and the auxiliary lattice.* — Formally, we can consider approximating the form of a physical, smooth, and isotropic interaction  $V(r_{ij})$  by a sum of  $N_t$  correlation functions at different temperatures,  $V(r_{ij}) \approx V_{\text{fit}}(r_{ij}) = \sum_{t=1}^{N_t} c_t f_{\beta_t}(r_{ij})$  [ $f_{\beta_t}(r_{ij}) \triangleq \langle o(\mathbf{r}_i) o(\mathbf{r}_j) \rangle_{\beta_t}$ ], giving the interaction operator as a sum of CF-PEPOs. However, taking  $V(r_{ij}) = 1/r_{ij}$  and  $f_{\beta_t}(r_{ij})$  as the Ising correlation function as examples, we see that a direct expansion is not efficient because, although the correlation functions are isotropic asymptotically, the lattice discretization prevents radial isotropy at short lattice distances. In addition, for finite lattices, boundary effects cause additional errors in isotropy and translational invariance. To illustrate this, we show the maximal absolute error in a direct fit of  $1/r_{ij}$  by Ising correlation functions on an  $L \times L$  lattice (with unit spacing) in Fig. 2(a). For small values of  $r_{ij}$ , the radial anisotropies in the basis  $\{f_{\beta_t}\}$  cause the fit to always be poor, even with large numbers of fitting functions, as seen in Fig. 2(b).

At larger  $r_{ij}$ , however, the maximal absolute error can be seen to converge rapidly, with a fitted convergence rate of  $\sim O(r_{ij}^{-2.7})$  (Fig. 2(a)). This suggests that the natural way to approximate isotropic interactions is to use correlation functions generated on an *expanded auxiliary lattice*, where the physical distance  $r_{ij}$  between operators maps to the expanded distance  $(N_f + 1)r_{ij}$  on the auxiliary lattice ( $N_f$  denotes the number of additional “fictitious” sites added to the sides of one unit square on the original lattice). For the Coulomb potential, this more accurate approximation can be found by a simple rescaling of the  $N_f = 0$  direct fit from above,

$$\tilde{V}_{\text{fit}}^{[N_f]}(r_{ij}) \triangleq (N_f + 1) \cdot V_{\text{fit}}\left((N_f + 1)r_{ij}\right), \quad (4)$$

where  $h = 1/(N_f + 1)$  is analogous to the lattice discretization of a continuous model of the correlation functions. This fitting scheme is attractive because it allows for precise extrapolation to the continuous limit as  $N_f \rightarrow \infty$  ( $h \rightarrow 0$ ) and requires only one least squares fitting procedure for  $V_{\text{fit}}(r_{ij})$  to obtain all  $\tilde{V}_{\text{fit}}^{[N_f]}(r_{ij})$ . Combining this lattice expansion with a suitably large side length buffering the physical region, we can also avoid any boundary effects in a finite PEPS simulation. (Note that for an infinite PEPO to be used with iPEPS, no

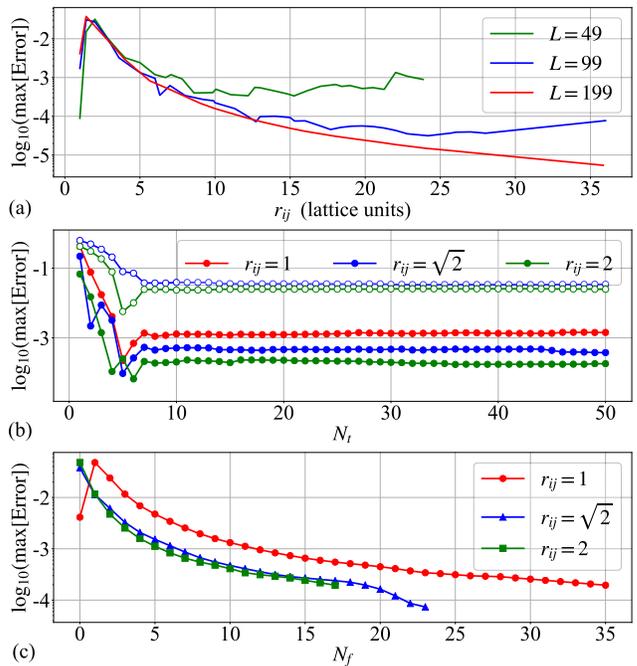


FIG. 2. Convergence properties of Coulomb fitting. For all plots  $r_{ij} = 0$  is the central point in the lattice. (a) The upper envelope of  $|V_{\text{fit}}(r_{ij}) - 1/r_{ij}|$  obtained with  $N_t = 12$  and a least squares weight function of  $r_{ij}^{1.5}$  for Ising model lattices with different side lengths  $L$ . The fits were performed on a disc with radius  $R$  equal to the maximum  $r_{ij}$  displayed for a given curve. (b) and (c): The maximum fitting error  $|\tilde{V}_{\text{fit}}^{[N_f]} - 1/r_{ij}|$  at selected values of  $r_{ij}$  as functions of  $N_t$  (b) and  $N_f$  (c). In (b), the open circles correspond to  $N_f = 0$  and the closed circles to  $N_f = 10$ . In (c),  $N_t = 12$ . The fits in (b) and (c) were performed on discs of radius  $R = 36$  Ising units with  $L = 199$  and a weight function of  $r_{ij}^{1.5}$ .

such boundary is required). The full CF-PEPO is thus obtained by coupling the FSM of the operators (either the snake form, or the full 2D FSM) to the new set of expanded Ising CF-PEPS specified by Eq. (4), as shown in Fig. 1(b)-(c).

As can be seen in Fig. 2(a), for  $N_f > 2$  the total error in the fit will be strongly dominated by the errors at the few shortest lattice spacings. Figs. 2(b)-(c) show the behavior of these dominant errors when using  $\tilde{V}_{\text{fit}}^{[N_f]}(r_{ij})$  to fit  $1/r_{ij}$  on the unit-spacing physical lattice, as a function of both the number of fictitious sites  $N_f$  and fitting terms  $N_t$ . On the interval  $r_{ij} \in [1, 36]$  (Ising units) we see that using  $N_f = 10$  and a modest  $N_t = 8$ , we are able to fit  $1/r_{ij}$  to a maximum absolute error of  $10^{-3}$ .

For system sizes  $r_{ij} \in [1, R]$  that are not unreasonably small, the behavior of these small- $r_{ij}$  errors is unaffected by the extent of the outer boundary  $R$ . Thus, the total error of the fit is strongly dominated by terms for which the saturated value of  $N_t$  is small and fixed (independent of  $R$ ). It can therefore be concluded that, for a given desired level of accuracy in fitting  $1/r_{ij}$ , the

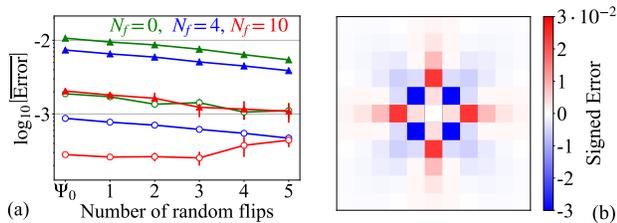


FIG. 3. (a) Average accuracy of energy per site expectation values for 8x8 FM and AFM trial PEPS with  $D_S = 1$ . The solid triangular markers show FM states while the open circles show AFM states.  $\Psi_0$  is a true FM or AFM state, while the “ $x$  flip” regions are  $\Psi_0$  perturbed by  $x$  random spin flips. The average error is taken over 5 PEPS for each  $x$  and each  $N_f$ . (b) The signed error  $1/r_{ij} - \tilde{V}_{\text{fit}}^{[0]}(r_{ij})$ , where  $r_{ij} = 0$  is the white square in the center, each adjacent square is  $r_{ij} = 1$ , etc. Use with Fig. 2(c). For (a)-(b) the fitted potentials are obtained from Eq. (4) with  $N_t = 12$ .

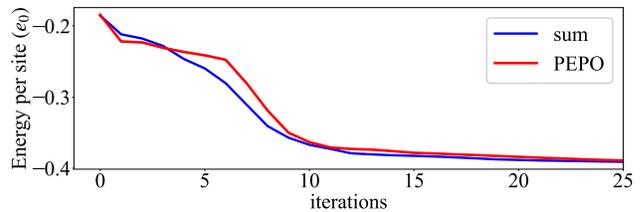
number of terms  $N_t$  in  $\tilde{V}_{\text{fit}}^{[N_f]}(r_{ij})$  is independent of, or at most weakly dependent on, system size. This is similar to what is observed in MPO fits in one dimension [19, 30, 32, 33] as well as analytical work on exponential fits of the Coulomb operator in 2D [38].

*Computational cost.* — Consider the evaluation of a finite PEPS expectation value for a PEPS of bond dimension  $D_S$  and an Ising CF-PEPO using the full 2D FSM (Fig. 1(b)), which has bond dimension  $D = 8$  for the bonds emanating from the physical sites and  $D' = 2$  for bonds that only connect fictitious sites. Following a simple generalization of the optimized contraction scheme proposed in Ref. [39] to include a PEPO, the leading cost can be derived to be  $N_t[O(A\chi^3 D^3) + O(AN_f \chi^3 D'^2 D) + O(AN_f^2 \chi^3 D'^3) + O(A\chi^3 D_S^3)]$ , where  $\chi \sim D_S^2 D$ . For the snake FSM construction (Fig. 1(c))  $D = 6$  instead of 8, and the physical PEPO tensors only have two large bond dimensions instead of four. This reduces the overall scaling to  $N_t[O(A\chi^3 D'^2 D) + O(AN_f \chi^3 D'^2 D) + O(AN_f^2 \chi^3 D'^3) + O(A\chi^3 D_S^3)]$ . The linearity of these expressions in the system size  $A$  is the crucial feature for making long-range interactions computationally viable.

*Long-range  $S=1/2$  Heisenberg model.* — To demonstrate our PEPO’s faithful discretized representation of  $1/r_{ij}$  and potential utility for studying systems with long-range interactions, we now use it to study a long-range  $S=1/2$  Heisenberg Hamiltonian on 4x4 and 8x8 square lattices,

$$H = \sum_{i < j} \frac{\vec{S}_i \cdot \vec{S}_j}{r_{ij}}, \quad (5)$$

in which every pair of spins has an interaction strength of Coulomb form. To represent this operator, we first used the fitting scheme described in Eq. (4) with  $N_t = 12$ . Figure 3(a) shows the accuracy of the energy per site expectation value ( $e_0$ ) for 8x8 trial ferromagnetic (FM) and anti-ferromagnetic (AFM) PEPS with  $D_S = 1$ . The FM state expectation values provide a non-pathological upper bound for the error of a given interaction fit, as all



	$\chi$	sum $e_0$	PEPO $e_0$	$\langle \psi_0^{[P]}   \psi_0^{[S]} \rangle$
4x4, $D = 1$	40	-0.184314	-0.184425	0.999244
4x4, $D = 2$	100	-0.408209	-0.408492	0.999070
8x8, $D = 1$	40	-0.193983	-0.193861	0.994549
8x8, $D = 2$	120	-0.414653	-0.414422	0.989271

FIG. 4. Top: The trajectories over the first 25 iterations of the energy optimization for the 4x4  $D_S = 2$  system using the PEPO and the explicit sum over all  $O(A^2)$  terms in (5). The long tails of the trajectories are excluded for clarity. Bottom: Ground state energies per site  $e_0$  for the Hamiltonian (5) with various system sizes and bond dimensions. The fifth column is the overlap of the ground states obtained with the two different methods. In all cases  $N_f = 4$  and  $N_t = 12$ .

interaction terms are of the same sign, while the AFM state expectation values exhibit better accuracy due to enhanced signed error cancellation, as shown in Figs. 3(b) and 2(c). The clearest evidence for this enhanced cancellation of the AFM states is the negative slope of the FM lines in Fig. 3(a) as the number of random spin flips increases.

We next performed a simple gradient-based variational optimization for the ground state PEPS with  $D_S = 1, 2$  [40, 41]. We employed a more accurate but less general least-squares fitting which obtained separate fitted potentials for each different system size [42]. This allowed us to, in all cases, tune the maximum PEPO fitting error to  $\sim 4.5 \cdot 10^{-4}$  with only  $N_f = 4$ ,  $N_t = 12$ . Fig. 4 shows the initial convergence behavior of the energy optimization using the PEPO compared to the same optimization using the more expensive sum over terms formalism. We observe that the trajectories are similar and the use of the PEPO does not change the stability of the gradient optimization, although it does require a larger value of  $\chi$ . The converged energies and wavefunction overlaps are given in Fig. 4. The maximum fitting error in the PEPO is faithfully reflected in the accuracy of  $e_0$ .

*Conclusions.* — In summary, we have detailed the efficient construction of a PEPO capable of encoding long-range interactions in 2D TNS that maintains the strengths of tensor network algorithms: systematically improvable accuracy and linear computational complexity in the system size. Despite an increased cost prefactor compared to local simulations, this approach allows for the possibility of practically including long-range interactions in numerical studies that use PEPS. Finally, this advance presents a first step toward *ab initio* electronic structure calculations with higher dimensional tensor networks.

*Acknowledgements.* — Primary support for this work was from MURI FA9550-18-1-0095, which supported MJO. Additional support was from the US National Science Foundation via grant CHE-1665333 for ZL. GKC acknowledges support from the Simons Foundation.

- 
- [1] S. R. White, Physical review letters **69**, 2863 (1992).  
 [2] S. R. White, Physical Review B **48**, 10345 (1993).  
 [3] G. Vidal, Physical review letters **99**, 220405 (2007).  
 [4] H. J. Changlani, J. M. Kinder, C. J. Umrigar, and G. K.-L. Chan, Physical Review B **80**, 245116 (2009).  
 [5] F. Mezzacapo, N. Schuch, M. Boninsegni, and J. I. Cirac, New Journal of Physics **11**, 083026 (2009).  
 [6] G. Carleo and M. Troyer, Science **355**, 602 (2017).  
 [7] S. Östlund and S. Rommer, Physical review letters **75**, 3537 (1995).  
 [8] M. Fannes, B. Nachtergaele, and R. F. Werner, Communications in mathematical physics **144**, 443 (1992).  
 [9] M. Fannes, B. Nachtergaele, and R. Werner, Journal of functional analysis **120**, 511 (1994).  
 [10] U. Schollwöck, Annals of Physics **326**, 96 (2011).  
 [11] T. Nishino and K. Okunishi, Journal of the Physical Society of Japan **65**, 891 (1996).  
 [12] F. Verstraete and J. I. Cirac, arXiv preprint cond-mat/0407066 (2004).  
 [13] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Physical review letters **96**, 220601 (2006).  
 [14] R. Orús, Annals of Physics **349**, 117 (2014).  
 [15] R. M. Martin, *Electronic structure: basic theory and practical methods* (Cambridge university press, 2004).  
 [16] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Intro to Advanced Electronic Structure Theory* (Dover publications, 1996).  
 [17] S. R. White and R. L. Martin, The Journal of chemical physics **110**, 4127 (1999).  
 [18] G. K.-L. Chan, A. Keselman, N. Nakatani, Z. Li, and S. R. White, The Journal of chemical physics **145**, 014102 (2016).  
 [19] E. Stoudenmire, L. O. Wagner, S. R. White, and K. Burke, Physical review letters **109**, 056402 (2012).  
 [20] L. O. Wagner, E. Stoudenmire, K. Burke, and S. R. White, Physical Chemistry Chemical Physics **14**, 8581 (2012).  
 [21] E. M. Stoudenmire and S. R. White, Physical review letters **119**, 046401 (2017).  
 [22] M. Dolfi, B. Bauer, M. Troyer, and Z. Ristivojevic, Physical review letters **109**, 020604 (2012).  
 [23] N. Mardirossian, J. D. McClain, and G. K.-L. Chan, The Journal of chemical physics **148**, 044106 (2018).  
 [24] Z. Y. Xie, J. Chen, J. F. Yu, X. Kong, B. Normand, and T. Xiang, Phys. Rev. X **4**, 011025 (2014).  
 [25] P. Corboz, T. M. Rice, and M. Troyer, Phys. Rev. Lett. **113**, 046402 (2014).  
 [26] P. Corboz and F. Mila, Phys. Rev. Lett. **112**, 147203 (2014).  
 [27] T. Picot and D. Poilblanc, Phys. Rev. B **91**, 064415 (2015).  
 [28] T. Picot, M. Ziegler, R. Orús, and D. Poilblanc, Phys. Rev. B **93**, 060407 (2016).  
 [29] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, Science **358**, 1155 (2017).  
 [30] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, New Journal of Physics **12**, 025012 (2010).  
 [31] F. Fröwis, V. Nebendahl, and W. Dür, Physical Review A **81**, 062337 (2010).  
 [32] G. M. Crosswhite and D. Bacon, Physical Review A **78**, 012356 (2008).  
 [33] G. M. Crosswhite, A. C. Doherty, and G. Vidal, Physical Review B **78**, 035116 (2008).  
 [34] H. Zhao, Z. Xie, Q. Chen, Z. Wei, J. Cai, and T. Xiang, Physical Review B **81**, 174411 (2010).  
 [35] See Supplementary Material for a comprehensive discussion on the construction of the finite state machine pairwise interaction tensors.  
 [36] P. Corboz, R. Orús, B. Bauer, and G. Vidal, Physical Review B **81**, 165104 (2010).  
 [37] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Physical review letters **101**, 250602 (2008).  
 [38] D. Braess and W. Hackbusch, IMA journal of numerical analysis **25**, 685 (2005).  
 [39] Z. Xie, H. Liao, R. Huang, H. Xie, J. Chen, Z. Liu, and T. Xiang, Physical Review B **96**, 045128 (2017).  
 [40] L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, Physical Review B **94**, 155123 (2016).  
 [41] P. Corboz, Physical Review B **94**, 035133 (2016).  
 [42] Instead of rescaling the fit done over the entire Ising lattice, a unique fit was performed for each system that accounts for the system size and the number fictitious sites. By doing this, the fit is performed on only the Ising grid points whose correlation functions are explicitly evaluated for the given system. This reduces the fitting error by avoiding both the additional factor of  $N_f$  incurred in (4) and the fitting of extraneous points. However, the penalty for performing the fit in this way is that it sacrifices the extrapolation properties of Eq. (4).

## SUPPLEMENTARY MATERIAL

### Coulomb fitting details

There are many different ways to fit a given long-range potential with the correlation functions of an auxiliary system. In this work, focusing specifically on the Coulomb potential and classical Ising model spin-spin correlation functions, we first computed the correlation functions at 60 different temperatures. To choose these temperatures, we first note that away from  $T_c$ , the correlation functions behave according to  $\sim e^{-r/\xi}$ , where

$$\xi \propto \left( \frac{T - T_c}{T_c} \right)^{-1}, \quad (6)$$

is the correlation length. Thus, a geometric series in  $(T - T_c)/T_c$  was used to select the temperatures, starting from  $T_1 = T_c + \delta$  and ending at  $T_{60} = 50J/k_B$ , where in this case  $\delta = 5 \cdot 10^{-4}$ , but in general it is just a small number that can be tuned.

With all of this data, a rank-revealing QR decomposition was performed on the full “basis matrix”  $\mathbf{A}$  corresponding to the system  $\mathbf{A}\vec{c} = 1/\vec{r}$ , where  $\vec{c}$  contains the linear fitting coefficients. This decomposition gives a best guess at the  $N_t$  most relevant basis functions (temperatures), allowing for a new data matrix  $\hat{\mathbf{A}}$  that is smaller and more well-conditioned to be used in the subsequent weighted least squares fitting of  $V(r_{ij}) = 1/r_{ij}$ . This fitting was described extensively in Figure 2 of the main text. Figures S5-S7 in this Supplementary Material are included to provide additional detail regarding the fitting displayed in Figure 2 of the main text.

Despite its simplicity, weighted least squares fitting appears to be a very natural approximation scheme for this problem. When viewed as a special case of the more general class of quantile regressions, it can be easily seen that least squares (ie. the 50% quantile regression) can maximize the inherent signed error cancellation that comes along with simple magnetic ordering patterns.

As a final note, in practice we have found the fitting procedure and the accuracy it achieves to be very sensitive to most parameters aside from  $N_t$  and the precise temperatures that are used for the basis functions. For this reason, it is possible that the fitting results presented in the current work do not represent the best solution for all possible system sizes, choices of  $N_f$ , and specific Hamiltonians of interest. In future work a thorough analytical treatment will be given to fermionic correlation functions, which may be able to provide more rigorous answers about the best way to perform this fitting.

### Finite state machine rules

The finite state machine picture of a PEPO views each tensor as a node in a graph, and each virtual bond of di-

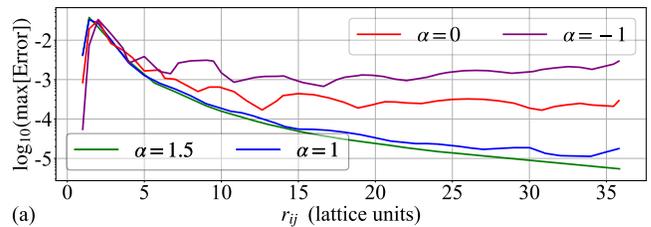


FIG. S5. The upper envelope of the errors of the fitted potential  $|V_{\text{fit}}(r_{ij}) - 1/r_{ij}|$  for the  $N_t = 12$  least square fittings  $L_\alpha = \sum_{ij} r_{ij}^\alpha (V_{\text{fit}}(r_{ij}) - 1/r_{ij})^2$  with different weighting exponent  $\alpha$  on a 2D Ising model lattice of side length  $L = 199$ .

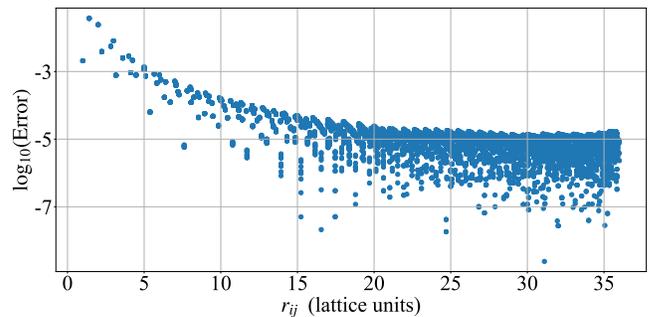


FIG. S6. All the errors  $|V_{\text{fit}}(r_{ij}) - 1/r_{ij}|$  at each  $r_{ij}$  for the  $N_t = 12$ ,  $\alpha = 1.5$ ,  $L = 199$  fit described in Figure 2 of the main text. Note that most of the errors for a given  $r_{ij}$  are significantly smaller than the upper envelope that was shown in the main text.

mension  $D$  as a directed edge in that graph that can pass  $D$  different signals (or has  $D$  different possible states).

*Full 2D FSM.* — By convention we have chosen our directed edges to point up and right so that, for a given tensor at position  $k$ , its  $U$  and  $R$  indices pass outgoing signals while its  $D$  and  $L$  indices receive incoming signals. When there are certain combinations of incoming and outgoing signals for a tensor at position  $k$ , that tensor’s two physical indices encode a local operator  $O_{n_k n'_k}^{[k]}$ , which is either a physical operator or the identity operator. These “certain combinations” of index values are precisely the state machine rules that construct the corresponding desired state machine. When the four virtual index values do not match any of these desired rules, the value of  $O_{n_k n'_k}^{[k]}$  is the zero operator  $\mathbf{0}$ , meaning such a configuration of the state machine (and therefore such a configuration of the local operators) is disallowed. The complete list of these rules that define the full 2D FSM PEPO which generates all pairwise interactions  $\sum_{i < j} \hat{A}_i \hat{B}_j$  with bond dimension  $D = 4$  is given in Table S1.

Each index value corresponds to a different signal, which is used to pass a different message. “0” is the default signal, which generally means that nothing interesting is happening along that signal path. “1” is the

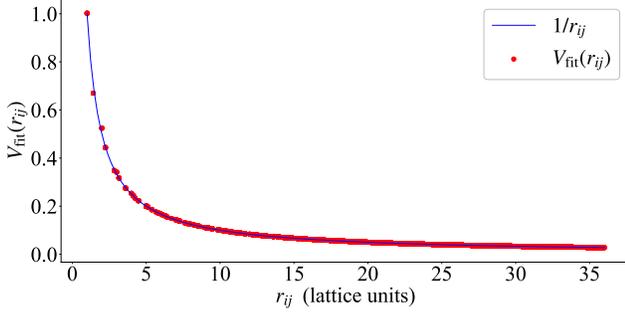


FIG. S7. The lattice discretized  $V_{\text{fit}}(r_{ij})$  compared to the continuous Coulomb potential for the  $N_t = 12$ ,  $\alpha = 1.5$ ,  $L = 199$  fit. Note that at small values of  $r_{ij}$  the values of  $V_{\text{fit}}$  visibly deviate from the exact solution, while as  $r_{ij}$  grows their agreement gets significantly better.

Rule number	Index values ( $L_k, U_k, D_k, R_k$ )	$O_{n_k n'_k}^{[k]}$
1	(0,0,0,0)	$I_k$
2	(0,2,2,0)	$I_k$
3	(2,1,0,2)	$I_k$
4	(0,1,1,0)	$I_k$
5	(1,1,0,1)	$I_k$
6	(0,2,0,0)	$\hat{A}_k$
7	(0,1,0,2)	$\hat{A}_k$
8	(0,1,2,2)	$I_k$
9	(0,1,2,1)	$\hat{B}_k$
10	(2,1,0,1)	$\hat{B}_k$
11	(3,1,0,3)	$I_k$
12	(3,1,2,1)	$I_k$
13	(0,1,0,3)	$\hat{B}_k$
14*	$P_{0,0,0,0}^{\text{top right}}$	$\hat{0}_k$

TABLE S1. The rules for the full 2D FSM PEPO that generates all pairwise interactions  $\sum_{i<j} \hat{A}_i \hat{B}_j$  with  $D = 4$ . All combinations of indices not listed in this table correspond to  $O_{n_k n'_k}^{[k]} = \hat{0}_k$ . Importantly,  $\hat{A}$  and  $\hat{B}$  do not have to be the same, although for the *ab initio* Hamiltonian under consideration in the main text, they are both  $n_k$ .  $I_k$  is simply the identity operator.

signal that tells nearby tensors that they should not turn on their physical operator  $O_{n_k n'_k}^{[k]}$ , but instead should just be the identity operator. This is used when another tensor along a certain signal path has turned on its physical operator and does not want an interaction to be generated along the signal path on which it just sent a “1” message. “2” is the signal that is passed along the “typical” interaction path between the physical operator at site  $i$  and the physical operator at site  $j$ . A typical interaction path is one in which a signal traveling from site  $i$  to site  $j$  must only propagate upward and to the right (along the allowed directions of the directed edges). The signal “3” is reserved for the cases in which the signal traveling from site  $i$  to site  $j$  must travel to the left. In

order to generate all pairs of sites, one must either have signals that travel up and to the left or down and to the right (violating one of the directed edge directions), but the case of down and to the left can be avoided due to the fact that we are generating all pairs of interactions only once (hence  $i < j$  in the summations). By convention, we have chosen this pathological case to be described by a signal that travels up and to the left. Since a signal cannot truly travel against the direction of a directed edge, this case is resolved by having the operator at site  $j$  (the operator at the “end” of the signal) send a “3” signal to the right, which then meets with a “2” signal that was sent upwards from site  $i$ , generating an interaction along a “non-typical” path. These cases are illustrated diagrammatically in Fig. S8.

The rules in Table S1 are broken up into different groups according to what they describe. Rules 1-5 are background rules that account for the propagation of “1” and “2” signals through the FSM. Rules 6-10 give the additional rules necessary for describing a typical interaction. Rules 11-13 add the rules for non-typical interactions. Finally, Rule 14 is a special rule that only applies to the top right tensor in the network, where all signals terminate. This rule is included to disallow the state of the machine where all tensors have virtual index values (0, 0, 0, 0) and a spurious 1 is added so that the final operator is  $1 + \sum_{i<j} \hat{A}_i \hat{B}_j$  instead of  $\sum_{i<j} \hat{A}_i \hat{B}_j$ . *Snake FSM.* — The snake construction for the FSM shown in Fig. 1(c) of the main text is much simpler than the full 2D FSM above because it is precisely just an MPO with a few extra dummy legs at each site so that the direct product with the Ising tensors can be done properly. As discussed briefly in the main text, the operator-valued local matrices for an MPO that encodes the interactions  $\sum_{i<j} \hat{A}_i \hat{B}_j$  are given by,

$$M^{[k]} = \begin{bmatrix} \hat{I}_k & \hat{A}_k & \hat{0}_k \\ \hat{0}_k & \hat{I}_k & \hat{B}_k \\ \hat{0}_k & \hat{0}_k & \hat{I}_k \end{bmatrix}. \quad (7)$$

Since this snake imposes an explicit ordering of all the sites on the 2D square lattice, it very naturally lends itself to the inclusion of fermionic statistics at the operator level via Jordan-Wigner strings. If the operators  $\hat{A}_i$  and  $\hat{B}_j$  are *spinless* fermionic creation or annihilation operators (and  $i < j$ ), then we have,

$$M^{[k]} = \begin{bmatrix} \hat{I}_k & \hat{a}_k(1 - 2\hat{n}_k) & \hat{0}_k \\ \hat{0}_k & 1 - 2\hat{n}_k & \hat{b}_k \\ \hat{0}_k & \hat{0}_k & \hat{I}_k \end{bmatrix}, \quad (8)$$

where  $\hat{a}_k$  and  $\hat{b}_k$  are the bosonic creation/annihilation operators and  $1 - 2\hat{n}_k$  encodes the fermionic statistics. For *spinful* fermionic operators we have to distinguish between spin up and spin down cases. For terms like

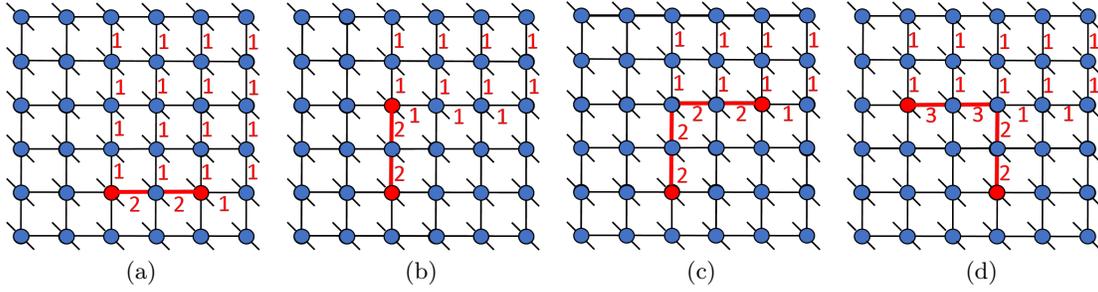


FIG. S8. The four cases of rules needed to build the PEPO that encodes all the pairwise terms in  $\sum_{i<j} \hat{A}_i \hat{B}_j$  for arbitrary operators  $\hat{A}$  and  $\hat{B}$ . All virtual bonds are labeled with their index value, except those that are indexed 0 which are left unlabeled. The red path denotes the path of the signal from  $\hat{A}_i$  to  $\hat{B}_j$ , which are signified by the two red tensors. Note that all the blue sites will be  $\hat{I}$  in these cases.

$\hat{A}_{i\uparrow} \hat{B}_{j\uparrow}$  we have ,

$$M_{\uparrow\uparrow}^{[k]} = \begin{bmatrix} \hat{I}_k & \hat{a}_k (-1)^{\hat{n}_k} & \hat{O}_k \\ \hat{O}_k & (-1)^{\hat{n}_k} & \hat{b}_k \\ \hat{O}_k & \hat{O}_k & \hat{I}_k \end{bmatrix}, \quad (9)$$

and for terms like  $\hat{A}_{i\downarrow} \hat{B}_{j\downarrow}$ ,

$$M_{\downarrow\downarrow}^{[k]} = \begin{bmatrix} \hat{I}_k & \hat{a}_k & \hat{O}_k \\ \hat{O}_k & (-1)^{\hat{n}_k} & (-1)^{\hat{n}_k} \hat{b}_k \\ \hat{O}_k & \hat{O}_k & \hat{I}_k \end{bmatrix}. \quad (10)$$

Here  $1 - 2\hat{n}_k$  changes to  $(-1)^{\hat{n}_k}$  because we need to account for the possibility of double occupancy at a given site  $k$ . This is also why we need to distinguish between spin up and spin down cases.