Accurate exponents from approximate tensor renormalizations

Y. Meurice

Department of Physics and Astronomy, The University of Iowa, Iowa City, Iowa 52242, USA

We explain the recent numerical successes obtained by Tao Xiang's group, who developed and applied tensor renormalization group methods for the Ising model on square and cubic lattices, by the fact that their new truncation method sharply singles out a surprisingly small subspace of dimension two. We show that in the two-state approximation, their transformation can be handled analytically, yielding a value of 0.964 for the critical exponent ν much closer to the exact value 1 than the 1.338 value obtained in the Migdal-Kadanoff approximation. We propose two alternative blocking procedures that preserve the isotropy and improve the accuracy to $\nu = 0.987$ and 0.993, respectively. We discuss applications to other classical lattice models, including models with fermions, and suggest that it could become a competitor for Monte Carlo methods suitable for accurate calculations of critical exponents, taking continuum limits, and the study of near-conformal systems in arbitrarily large volumes.

The renormalization group (RG) ideas have triggered considerable conceptual and numerical progress in many branches of physics.[1,2] However, the basic method to thin down the number of degrees of freedom in configuration space,[3] often called "block spinning", has remained a formidable computational challenge for most classical lattice models [e.g., O(N) spin models and lattice gauge theories]. A few years ago, inspired by the so-called tensor network states[4,5] introduced in the context of the density matrix RG method,[6,7] a tensor RG (TRG) approach of two-dimensional (2D) classical lattice models was proposed.[8] Successful approximations[8–10] were found for the Ising model on honeycomb and triangular lattices.

Very recently, the TRG method was successfully extended to the Ising model on square and cubic lattices by Tao Xiang's group.[11] There are two important ingredients in their calculations. First, their formulation allows an *exact* block-spinning procedure which separates neatly the degrees of freedom inside the block, which are integrated over, from those kept to communicate with the neighboring blocks. As explained below, this provides a more systematic way to implement ideas initiated by Migdal[12] and Kadanoff[13] (abbreviated as MK hereafter). The indices of the tensors run over some finite vector space of "states" associated with finite volume link configurations. Second, they used a new method, based on higher-order singular value decomposition, which selects in a very economical way the most important states that ensure communication among the blocks. Calculations using of the order of 20 states can be carried on a laptop computer. The excellent agreement found with the Onsager solution in two dimensions for arbitrarily large volume suggests that TRG-based methods could become competitors for conventional Monte Carlo methods.

In this paper, we show that the truncation method of Ref. 11 for the two-dimensional (2D) Ising model sharply singles out a 2D subspace of states. Keeping only these two states, we show that we can construct approximate RG transformations with three or four parameters, find the nontrivial fixed point, and obtain precise estimates of the critical exponent ν associated with the correlation length from a linear analysis. The accuracy of the estimates is significantly better than for textbook examples such as the

MK approximation,[12–14] the so-called approximate recursion formula,[15] or other hierarchical approximations.[16,17]

The TRG formulation for the Ising model can be extended to O(N) nonlinear σ models and recent numerical implementations for O(2) (Ref. 18) indicate an optimistic outlook. It also seems possible to formulate models with local invariance and avoid sign problems. In this context, it is important to understand why the method works so well for the Ising model.

The paper is organized as follows. We review the basic TRG formulation for the Ising model on a square lattice, emphasizing the connection with the MK ideas. We then consider the cases of an isotropic blocking (as in the Migdal recursion) and an anisotropic blocking (as in the Kadanoff version and in Ref. 11). We also propose a new type of accurate isotropic projection based on a transfer matrix. We briefly discuss the 3D Ising model and how the method can be applied for lattice fermions. The implications for the study of near-conformal systems and the calculations of critical exponents are discussed in the conclusions.

We consider the nearest-neighbor Ising model on a square lattice with an inverse temperature β. For easy reference, we stay close to the notations of Ref. 11, where it is shown that the partition function can be written as

$$Z = Tr \prod_i T^{(i)}_{xx'yy'}. \qquad (1)$$

The tensor $T^{(i)}_{xx'yy'}$ is attached to each site i, and Tr is a short notation for contractions over the links joining nearest neighbors on the lattice. The horizontal indices x, x' and vertical indices y, y' take the values 0 and 1. The tensor is zero for an odd number of 1. For an even number of 1, a factor $t^{1/2}$ [with $t \equiv \tanh(\beta)$] appears for each 1, irrespective of the direction. The partition function can be interpreted as a sum over intermediate states attached to the links. The reader familiar with the high-temperature expansion of the model will recognize that this expression reproduces exactly the proper closed paths with the proper weights.

We now use this reformulation to blockspin. First, we follow Migdal[12] by using an isotropic procedure. We consider a square block enclosing four sites and sum over the states, inside the
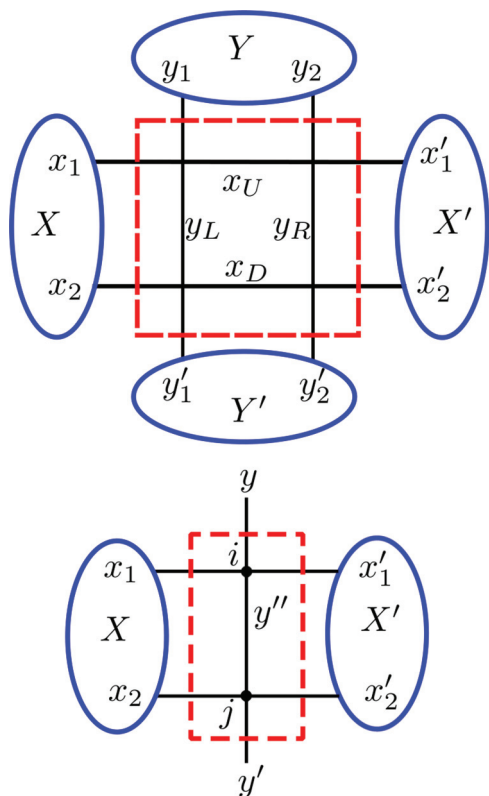
FIG. 1. (Color online) Graphical representation of $T'_{XX'YY'}$ (top) and $M^{(ij)}_{X(x_1,x_2)X'(x'_1x'_2)yy'}$ (bottom). The boundary of the block is represented by the dashed lines and the product states by ovals.

block, associated with the nearest-neighbor links joining these four points. This defines a new rank-4 tensor $T'_{XX'YY'}$, where each index now takes four values:

$$
\begin{aligned}
T'_{X(x_1,x_2)X'(x'_1,x'_2)Y(y_1,y_2)Y'(y'_1,y'_2)} \\
= \sum_{x_U,x_D,x_R,x_L} T_{x_1 x_U y y_L} T_{x_U x'_1 y_2 y_R} T_{x_D x'_2 y_R y'_2} T_{x_2 x_D y_L y'_1},
\end{aligned}
\tag{2}
$$

where $X(x_2,x_2)$ is a notation for the product states. For reasons that will become clear later, we use the convention $X(0,0) = 1$, $X(1,1) = 1$, $X(1,0) = 3$, $X(0,1) = 4$. Later, we also use the ket notation $|00\rangle$ for $X = 1$, etc. This is represented graphically in Fig. 1.

The new tensor can be used to define an *exact* expression of the partition function of the same form as (1); however, the number of states proliferates as 2^{2^n} after n steps and approximations are needed in order to get an expression useful for practical purposes.

The truncation method of Ref. 11 relies on an anisotropic blocking involving two sites as shown at the bottom of Fig. 1. This provides a new rank-4 tensor:

$$
M^{(ij)}_{X(x_1,x_2)X'(x'_1x'_2)yy'} = \sum_{y''} T^{(i)}_{x_1,x'_1,y,y''} T^{(j)}_{x_2 x'_2 y'' y'},
\tag{3}
$$

which can be put in a canonical form by using a higher-order singular value decomposition defined by a unitary transformation on each of the four indices (see Ref. 11 for justifications and refinements). The unitary transformation for each index is the one that diagonalizes the symmetric tensor
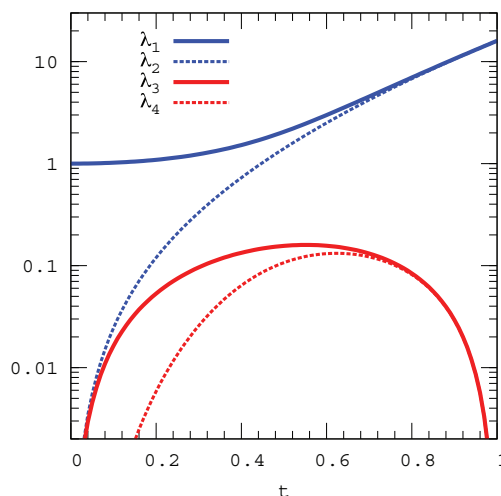

FIG. 2. (Color online) The four eigenvalues of $G_{XX'}$ at the first step, on a logarithmic scale, as a function of $t = th\beta$.

obtained by summing over all the other indices in the following way:

$$
G_{XX'} = \sum_{X''yy'} M_{XX''yy'} M^*_{X'X''yy'}.
\tag{4}
$$

We now consider approximations where for each index we only keep the two states which correspond to the two largest eigenvalues of $G_{XX'}$. With our convention on the product states, this matrix is block diagonal because it does not connect states with an even number of 1's ($X = 1,2$) to states with an odd number of 1's ($X = 3,4$). Numerically, we found that the two smallest eigenvalues are always very small compared to the largest one for any value of β and that the second-largest one is small at small β and almost as large as the largest one at large β. This situation is illustrated for the initial step in Fig. 2. After iterations the gap sharpens as if going to smaller t for $t < t_c$ or larger t for $t > t_c$.

The two new states have the form

$$
|0'\rangle = \cos\phi|00\rangle + \sin\phi|11\rangle, \quad |1'\rangle = (|10\rangle + |01\rangle)/\sqrt{2}.
\tag{5}
$$

The angle ϕ is obtained by diagonalizing the even-even block. The symmetric form of $|1'\rangle$ is a consequence of $G_{33} = G_{44}$, itself due to reflection symmetry.

With this rather drastic projection, we obtain a new rank-4 tensor with indices taking again two values and the same parity selection rule. We divide the new tensor by a constant in such a way that, dropping all the primes hereafter, $T_{0000} = 1$. In the isotropic case, the reflection symmetry imposes that $T_{1010} = T_{0110} = T_{1001} = T_{0101} \equiv t_1$, $T_{1100} = T_{0011} \equiv t_2$, $T_{1111} \equiv t_3$. For the initial tensor, we have what we call later the "Ising condition",

$$
t_1 = t_2 \quad \text{and} \quad t_3 = t_1^2,
\tag{6}
$$

but this property is not preserved by the blocking procedure, which can be expressed as a mapping of the three-dimensional (3D) parameter space (t_1, t_2, t_3) into itself. This can be expressed as rational expressions involving the values of the tensor and trigonometric functions of ϕ. They are analytical

expressions and derivatives can be taken to calculate the linearized map. They are not written explicitly here but can be obtained easily with symbolic manipulation programs.

A fixed point is found for $\beta = 0.411\,594\,547\ldots$ not far from the exact value $0.440\,68.\ldots$. The fixed point is approximately $t_1^\star = 0.443\,492$, $t_2^\star = 0.298\,462$, $t_3^\star = 0.304\,291$, and shows significant departure from the Ising condition (6). The eigenvalues of the linearized RG transformation are $\lambda_1 = 2.017\,7$, $\lambda_2 = 0.202\,2$, and $\lambda_3 = 0.099\,67$. The scaling factor b is 2, and this implies the values for the critical exponents $\nu = \log b / \log \lambda_1 \simeq 0.987$ and $\omega = -\log \lambda_2 / \log b \simeq 2.31$, which can be compared with the exact values 1 and 2, respectively.

For comparison, the also isotropic Migdal recursion for the $D = 2$ Ising model and $b = 2$ can be written as

$$t' = 2t^2/(1 + t^4). \tag{7}$$

The fixed point is at $t^\star = 0.543\,689$ and corresponds to $\beta_c = 0.609\,378$, and $\nu = 1.338$. This approximation can be seen as a two-state approximation, which in addition requires the Ising condition. Improvements of MK discussed in Ref. 14 lead to a value of $\nu = 0.796$ at first order in their expansion parameter and 0.93(1) at second order, the error bar coming from the use of different Padé approximants.

This suggests that the Ising condition is too restrictive. It is nevertheless possible to modify the angle θ in Eq. (5) in such a way that Ising condition remains valid. This is nontrivial because the Ising condition amounts to two equations. However, an explicit calculation shows that if $\tan\phi = 1/\tan\beta$, the *two* conditions are satisfied and the mapping takes the form

$$t' = (1 + t^2)/2t , \tag{8}$$

which unfortunately has only the low-temperature fixed point.

The above results can be compared with the two-state truncation for the anisotropic blocking used in Ref. 11, where a first projection occurs after blocking pairs of vertical sites, as already introduced in Eq. (3). The complete transformation is then obtained by repeating the procedure with a horizontal blocking as in Kadanoff's proposal.[13] We use our previous notations but with h and v subscripts denoting the horizontal and vertical couplings, respectively: $T_{1100} \equiv t_{2h}$ and $T_{0011} \equiv t_{2v}$. We keep the same notations for t_1 and t_3, which are invariant under the interchange of the vertical and horizontal directions. We have now a map with four parameters. In the Ising condition, we need to replace t_2 by $\sqrt{t_{2v}t_{2h}}$. We take initial values that satisfy the Ising condition and are isotropic [$t_{2h} = t_{2v} = t = \tanh(\beta)$]. The critical value is then $t_c = 0.379\,453\,244\,411\,09\ldots$, with a nontrivial fixed point for approximately $t_{2h}^\star = 0.213\,58$, $t_{2v}^\star = 0.379\,24$, $t_1^\star = 0.419\,98$, and $t_3^\star = 0.271\,77$, which is clearly anisotropic and violates the Ising condition. We have an additional, intermediate eigenvalue which is approximately 0.657 and specific to the anisotropic case. This value is not very far from unity, which is why it requires more fine tuning of β to get rid of the irrelevant directions. The first and third eigenvalues are 2.052 and 0.193 4, respectively, which implies $\nu = 0.964$ and $\omega = 2.37$.

This anisotropic version can be compared with the Kadanoff recursion (Ref. 13) for $b = 2$, where first, the horizontal bonds are slid vertically with β_h doubled and t_v squared. This corresponds to doubling the vertical lattice spacing first as we did above. After repeating the procedure with horizontal

moves, we obtain the Migdal recursion of Eq. (7) for t_v while for t_h we obtain

$$t_h' = \left[2t_h/\left(1 + t_h^2\right)\right]^2 , \tag{9}$$

which corresponds to reversing the order of the two operations. The fixed point is $t_v^\star = 0.543\,689$ and $t_h^\star = (t_v^\star)^2 = 0.295\,598$. The eigenvalue is the same in both directions and the value of ν identical to the Migdal case.

We have also considered the isotropic map with a different truncation. Instead of Eq. (4), we use

$$\tilde{G}_{XX'} = \sum_{X''yy'} M_{XX''yy} M^*_{X'X''y'y'} . \tag{10}$$

The trace of this matrix is the partition function for a 2×2 model with periodic boundary conditions. This gives a slightly displaced fixed point at $\beta_c = 0.394\,867\,8$, $t_1^\star = 0.422\,29$, $t_2^\star = 0.286\,37$, and $t_3^\star = 0.274\,66$. The values of the exponents are $\nu = 0.993$ and $\omega = 2.37$. It should be noted that in this case the two small eigenvalues at criticality (0.001 28 and 0.000 069 8) are much smaller than in the first calculation (0.118 and 0.052 5), which may explain the improved accuracy on ν.

Extensions of these methods for more states, more components, and more dimensions are in progress. For practical purposes, the analytical methods discussed above need to be implemented numerically. We have succeeded in reproducing all the results obtained so far with adequate accuracy using numerical procedures which can be implemented using the most common programming languages. The fixed point can be found by monitoring successive bifurcations in tensor values. The high- and low-temperature phases are characterized by the fact that some tensor values go to zero (when β is too small) or one (when β is too large) if we iterate enough times. We can then fine tune β and observe the stabilization of tensors at some nontrivial values. The eigenvalues can be found by taking numerical derivatives of the one-step transformation with respect to the initial values as close as possible to the nontrivial fixed point. This requires variations small enough but not too small, since we have only limited accuracy on β_c; however, accurate results can be obtained using linear extrapolations to zero variation.

The formulation can be extended in D dimension using tensors with with $2D$ indices. The reason the TRG blocking works well in any dimension is that the links are orthogonal (dual) to domain boundaries. We keep the link variables across the boundaries of the block fixed and sum unrestrictedly over all the states inside the block. We have extended the third method described above to the 3D Ising model. The block is then a cube with four external legs coming out of each of the six faces. The blocking and the projection can be built out of the loop made by the four edges of a face with four external legs attached to each of the four corners. The initial transfer matrix can be obtained by tracing the external legs in the plane of the loop with their opposite leg. We then obtain a 16×16 matrix corresponding to the four legs coming out of the plane of the loop in each direction. This matrix splits into two 8×8 blocks. Numerically, we found $\beta_c = 0.199\,659\,777\,323\,9\ldots$ for a two-state projection. The initial eigenvalues of $\tilde{G}_{XX'}$ at β_c are 1.232 5, 0.508 2, and a pair with value 0.168 2 which cannot be considered as small. For this reason, the value of $\nu \simeq 0.74$

for the 2D Ising model in the two-states approximation is not very close to the accurate value 0.630(2),[19] but it is nevertheless more accurate than the MK approximation value 1.055.

There is good empirical evidence[11,18] supporting the idea that as we increase the number of states in TRG calculations, the numerical estimates of energy and entropy get closer to values obtained by exact methods (for the 2D Ising model) or Monte Carlo simulations [for the 2D O(2) model]. However, in order to reach a reasonably large number of states (20–30) on a laptop, one needs to use efficient methods. The anisotropic TRG methods discussed above involve fewer contractions or external legs and the computational cost for N_s states can be limited[11] to a N_s^7 growth in two dimensions and a N_s^{11} growth in three dimensions. Numerical implementations of our method for up to 12 states are now in progress.[20] In all cases, we observe a sharp split between a few large eigenvalues, their number being characteristic of the phase, and the small ones as in the two-states case. The interesting behavior of these new maps will be discussed elsewhere.[20]

Because of the binary nature of Grassmann numbers, the techniques developed for Ising models can also be used for lattice models with fermions. In the case where we have Grassmann variables $\psi_\alpha^{(i)}$ with $\alpha = 1, 2, \dots, q$ at every site i and nearest-neighbor interactions specified by a $q \times q$ matrix $A_{\alpha\beta}^{\langle ij \rangle}$ at every link $\langle ij \rangle$, we can write

$$\exp\left(\psi_\alpha^{(i)} A_{\alpha\beta}^{\langle ij \rangle} \psi_\beta^{(j)}\right) = \sum_{n_\alpha^{\langle ij \rangle} = 0, 1} \prod_{\alpha=1}^{q} \left(\tilde{\psi}_\alpha^{(i)} \lambda_\alpha \tilde{\psi}_\alpha^{(j)}\right)^{n_\alpha^{\langle ij \rangle}}.$$

The $\tilde{\psi}$ are linear combinations of the ψ at the same site obtained from the decomposition $A = U\Lambda V^\dagger$, with U and V unitary and Λ diagonal with elements λ_α. The terms can then be factorized at every site and the local integrations performed. The states are now parametrized by $n_\alpha^{\langle ij \rangle}$ and there are 2^q of them at every link. Translational invariance is essential to perform large-volume calculations. For this reason, possible gauge interactions would need to be averaged inside the blocks.

In summary, we have shown that two-state approximations of the TRG capture the universal behavior of Ising models much better than the MK approximation. Building on the numerical success of Ref. 11, which uses more states, we expect to be able to use the methods presented here to calculate the exponents of the 3D Ising model with unprecedented accuracy and study the analytical picture of the critical behavior provided in Refs. 21 and 22. Recent numerical results for the TRG method applied to the O(2) model[18] suggest that improvements of the MK approximation could be applied to Abelian models and resolve the controversy regarding the confinement in four-dimensional U(1) gauge theory discussed in Ref. 23. We are hoping to be able to extend the TRG method for lattice gauge theories with fermions. Being able to block spin accurately would provide an efficient tool to study the continuum limits of asymptotically free models and the conformal window of models that could provide alternatives to the fundamental Higgs mechanism.[24]

[1] J. L. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, UK, 1996).

[2] Y. Meurice, R. Perry, and S.-W. Tsai, Philos. Trans. R. Soc. A **369**, 2602 (2011), and refrences therein.

[3] K. Wilson and J. Kogut, Phys. Rep. **12**, 75 (1974).

[4] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.

[5] J. I. Cirac and F. Verstraete, J. Phys. A **42**, 4004 (2009).

[6] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[7] U. Schollwoeck, Philos. Trans. R. Soc. A **369**, 2643 (2011).

[8] M. Levin and C. P. Nave, Phys. Rev. Lett. **99**, 120601 (2007).

[9] Z. Y. Xie, H. C. Jiang, Q. N. Chen, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **103**, 160601 (2009).

[10] H. H. Zhao, Z. Y. Xie, Q. N. Chen, Z. C. Wei, J. W. Cai, and T. Xiang, Phys. Rev. B **81**, 174411 (2010).

[11] Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang, and T. Xiang, Phys. Rev. B **86**, 045139 (2012).

[12] A. A. Migdal, Sov. Phys. JETP **42**, 743 (1975).

[13] L. Kadanoff, Ann. Phys. **100**, 359 (1976).

[14] G. Martinelli and G. Parisi, Nucl. Phys. B **180**, 201 (1981).

[15] K. Wilson, Phys. Rev. B **4**, 3184 (1971).

[16] G. Baker, Phys. Rev. B **5**, 2622 (1972).

[17] For a review, see Y. Meurice, J. Phys. A **40**, R39 (2007).

[18] A. Denbleyker, Y. Liu, Y. Meurice, M. Qin, T. Xiang, Z. Xie, and J. Yu (unpublished).

[19] A. Pelissetto and E. Vicari, Phys. Rep. **368**, 549 (2002).

[20] Y. Meurice, J. Unmuth-Yockey, and H. Zou (unpublished).

[21] S. El-Showk, M. F. Paulos, D. Poland, S. Rychkov, D. Simmons-Duffin, and A. Vichi, Phys. Rev. D **86**, 025022 (2012).

[22] S. El-Showk and M. F. Paulos (submitted for publication), arXiv:1211.2810.

[23] E. T. Tomboulis, Mod. Phys. Lett. A **24**, 2717 (2009).

[24] T. DeGrand, Phil. Trans. R. Soc. A **369**, 2701 (2011).