Taming sign problems using tensor renormalization

Yannick Meurice

The University of Iowa

yannick-meurice@uiowa.edu

Work done with Alexei Bazavov, Shailesh Chandrasekharan, Alan Denbleyker, Yuzhi "Louis" Liu, Chen-Yen Lai, Shan-Wen Tsai, Judah Unmuth-Yockey, Tao Xiang, Zhiyuan Xie, Ji-Feng Yu, and Haiyuan Zou

SIGN 2014, February 19, 2014

Can we build a Bose-Hubbard quantum simulator for the classical $O(2)$ model with real μ ?

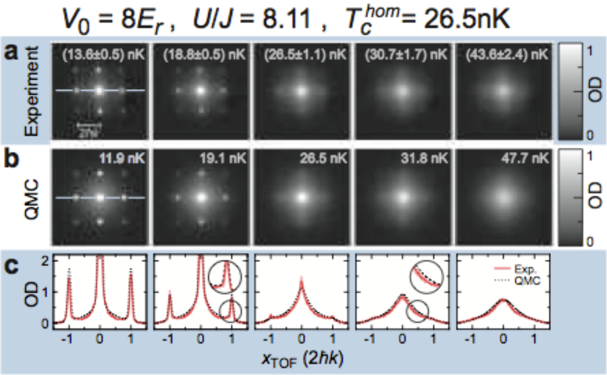


Figure: Comparison of experimental and simulated TOF distributions from optical lattices: From S. Trotzky, L. Pollet, F. Gerbier, U. Schnorrberger, I. Bloch, N.V. Prokof'ev, B. Svistunov, M. Troyer Nature Phys. 6, 998 (2010)

Fermi-Hubbard quantum simulators for LGT?

Baryons and Mesons (Fradkin et al.):after a particle-hole transformation in the spin down operator

$$f_{\mathbf{i},\uparrow}, f^\dagger_{\mathbf{i},\uparrow} \to \Psi_{x,1}, \Psi^\dagger_{x,1}; \qquad f_{\mathbf{i},\downarrow}, f^\dagger_{\mathbf{i},\downarrow} \to \Psi^\dagger_{x,2}, \Psi_{x,2}$$

The Heisenberg Hamiltonian can be written as follows

$$H = \frac{J}{8}\sum_{x,\hat{\mathbf{i}}}[M_x M_{x+\hat{\mathbf{i}}} + 2(B^\dagger_x B_{x+\hat{\mathbf{i}}} + B^\dagger_{x+\hat{\mathbf{i}}} B_x)] - \frac{Jd}{4}\sum_x (M_x - \frac{1}{2}) \; ,$$

where the “meson” and “baryon” operators are

$$M_x = \sum_{a=1,2} \Psi^\dagger_{x,a}\Psi_{x,a}$$

and

$$B_x = \sum_{a=1,2} \frac{\epsilon_{ab}}{2}\Psi_{x,a}\Psi_{x,a} = \Psi_{x,1}\Psi_{x,2}$$

Strong coupling correspondence with LGT

The meson+ baryon Hamiltonian can also be obtained at lowest order in the strong coupling expansion of the Kogut-Susskind Hamiltonian of $SU(2)$ lattice gauge theory with $J = 16/3g^2 \propto \beta$.

$$H = \frac{8}{3J} \sum_{a,\mathbf{x},\mathbf{i}} E^a_{\mathbf{xi}} E^a_{\mathbf{xi}} + \frac{i}{2} \sum_{a,b,\mathbf{x},\mathbf{i}} (\psi^\dagger_{\mathbf{x},a} \hat{U}^{ab}_{\mathbf{x},\mathbf{i}} \psi_{\mathbf{x}+\mathbf{i},b} - h.c.)$$

The approximate (strong coupling) correspondence between the Hubbard model and the above $SU(2)$ theory can be seen as a starting point to implement lattice gauge theory on optical lattice. There are no plaquette interactions for the above lattice gauge Hamiltonian and there are no spin indices for the gauge fields.

Tensor Renormalization Group (TRG): can we do better than the worm algorithm?

- Exact blocking with controllable approximations
- Deals well with sign problems, reliable at larger Imβ than reweighting MC
- Ising case: checked with the complex Onsager-Kaufman exact solution
- Finite Size Scaling of Fisher zeros of $O(2)$ agrees with KT
- Robust estimations of the eigenvalues of the transfer matrix
- Agreement with the worm algorithm for $O(2)$ with a real chemical potential μ. Allows calculations with imaginary μ
- Connects smoothly phase diagrams in the $\beta - \mu$ plane in the time continuum limit
- Connection with Bose-Hubbard and real time evolution?
- References: PRD 88 056005, PRD 89 016008, PRE 89 013308, arXiv 1402-xxxx

TRG blocking for Ising: it's simple and exact!

For each link:

$$\exp(\beta\sigma_1\sigma_2) = \cosh(\beta)(1 + \sqrt{\tanh(\beta)}\sigma_1\sqrt{\tanh(\beta)}\sigma_2) =$$
$$cosh(\beta)\sum_{n_{12}=0,1}(\sqrt{\tanh(\beta)}\sigma_1\sqrt{\tanh(\beta)}\sigma_2)^{n_{12}}.$$

Regroup the four terms involving a given spin σ_i and sum over its two values ± 1. The results can be expressed in terms of a tensor: $T^{(i)}_{xx'yy'}$ which can be visualized as a cross attached to the site i with the four legs covering half of the four links attached to i.
The horizontal indices $x,\ x'$ and vertical indices $y,\ y'$ take the values 0 and 1 as the link variables n_{12}.

$$T^{(i)}_{xx'yy'} = f_x f_{x'} f_y f_{y'} \delta\left(\text{mod}[x + x' + y + y', 2]\right)\ ,$$

where $f_0 = 1$ and $f_1 = \sqrt{\tanh(\beta)}$. The delta symbol is 1 if $x + x' + y + y'$ is zero modulo 2 and zero otherwise.

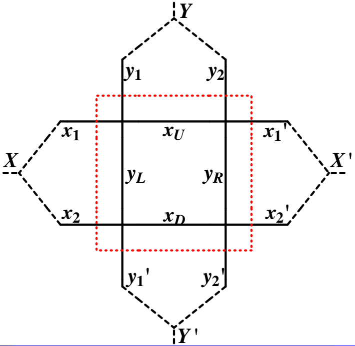
Isotropic TRG blocking (graphically)

Exact form:

$$Z = (\cosh(\beta))^{2V}\,\mathrm{Tr}\prod_i T^{(i)}_{xx'yy'}$$

Tr means contractions (sums over 0 and 1) over the links

TRG blocking separates the degrees of freedom inside the block which are integrated over, from those kept to communicate with the neighboring blocks. Graphically :



TRG Blocking (formulas)

Blocking defines a new rank-4 tensor $T'_{XX'YY'}$ where each index now takes four values.

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

where $X(x_2, x_2)$ is a notation for the product states e. g. , $X(0,0) = 1,\ X(1,1) = 2,\ X(1,0) = 3,\ X(0,1) = 4$. The partition function can be written exactly as

$$Z = (\cosh(\beta))^{2V} \operatorname{Tr} \prod_{2i} T'^{(2i)}_{XX'YY'} \ ,$$

where $2i$ denotes the sites of the coarser lattice with twice the lattice spacing of the original lattice.

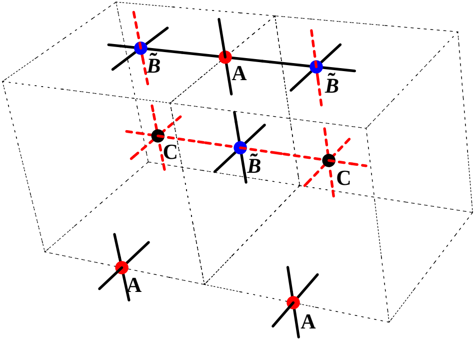
TRG formulations for other lattice models

- $O(2)$ and $O(3)$ models
- Principal chiral models
- Abelian gauge theories
- $SU(2)$ gauge theories

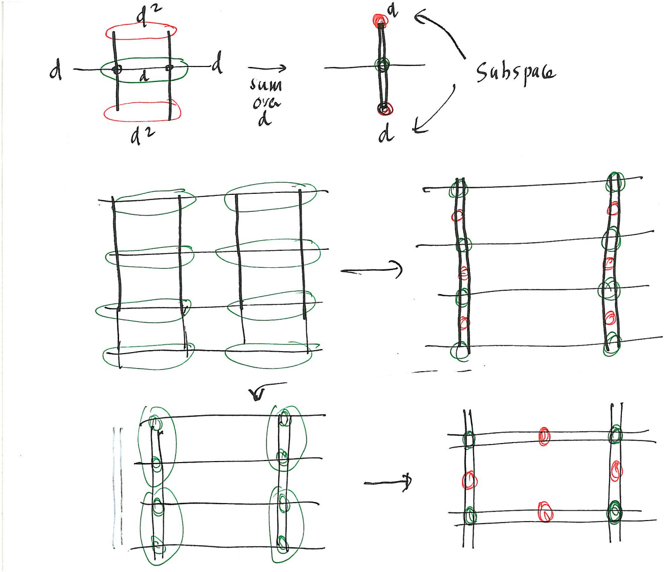
PRD 88 056005
arXiv:1307.6543 [hep-lat]

Blocking for gauge theories (graphically)

A blocking procedure can be constructed by sequentially combining two cubes into one in each of the directions.



Anisotropic blocking (memory economy)



Tensor Renormalization Group (TRG) iterations

Blocking:

$$M^{(n)}_{xx'yy'} = \sum_i T^{(n-1)}_{x_1 x'_1 yi} T^{(n-1)}_{x_2 x'_2 iy'}$$

where $x = x_1 \otimes x_2$ and $x' = x'_1 \otimes x'_2$.

Unitary transformation $U^{(n)}$ followed by a truncation

$$T^{(n)}_{xx'yy'} = \sum_{ij} U^{(n)}_{ix} M^{(n)}_{ijyy'} U^{*(n)}_{jx'}$$

The unitary matrix U is determined by taking the singular value decomposition of a specific matrix denoted as Q. By truncating the number of states to D_s, we mean keeping the eigenvectors corresponding to the first D_s largest singular values of Q. In Xie et al. PRB86 (HOTRG), for real β, Q was chosen as

$$Q \equiv M' M'^{\dagger} = U \Lambda U^{\dagger} = ||T||^2$$

where the matrix $M'_{x,x'yy'}$ is converted from the tensor $M_{xx'yy'}$ by grouping its indices.

Remarks

- The linear algebra seems insensitive to the fact that the values of the initial tensor become complex
- This allows us to deal with complex β, chemical potential (no apparent sign problem)
- However, when one approaches a zero of the partition function, larger truncations are necessary
- The TRG allows us to study the analyticity in complex β and μ planes
- There are subtleties with parity at complex β (H. Zou); need for CP or PT considerations?

2D Ising model on $L \times L$ lattice

$$\beta H = -\beta \sum_{<ij>} \sigma_i \sigma_j .$$

The exact solution for the partition function at finite volume was written by Kaufman in 1949.
When β is complex, the choices of signs for the square roots require care. For even L, the Hamiltonian is always a multiple of four and

$$Z(\beta) = Z(\beta + in\pi/2)$$

In the infinite volume limit, the partition function zeros lies on two circles in the complex $\tanh\beta$ plane given by $\pm 1 + \sqrt{2}\exp(i\theta)$ $(0 \leq \theta \leq 2\pi)$ Fisher 65 (these are the "Fisher's zeros")

The first quadrant part of the zeros with $0 \leq \mathrm{Im}\beta \leq \pi/2$ are shown below. $Z(-\beta) = Z(\beta)$ and $Z(\beta^*) = Z(\beta)^*$.

(Original) Fisher zeros

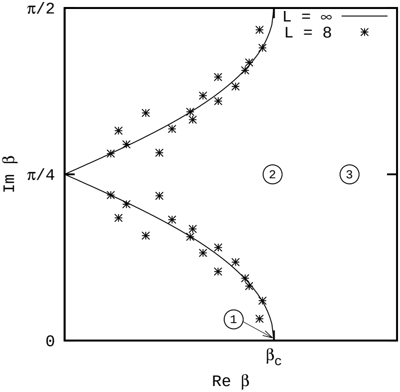


Figure: The zeros form periodic curves in the infinite volume limit. The zeros at finite volume are near the curves. Points are the zeros for a $L = 8$ system. Complex β regions 1, 2, and 3.

Near a Fisher 0, we need larger D_s (# of TRG states)

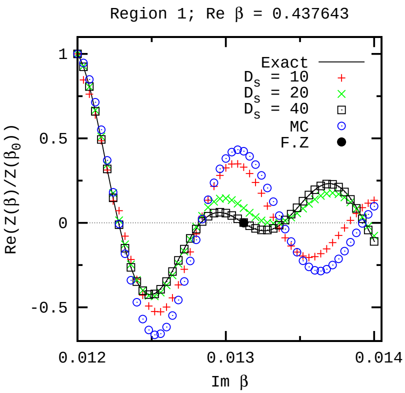


Figure: The real part of the normalized partition function $\mathrm{Re}[Z(\beta)/Z(\beta_0)]$ for β near the Fisher zero $0.437643 + i0.01312$ (the big filled circle on the horizontal axis): result from the HOTRG with $D_s = 10, 20, 40$ ($D_s = 30$ result is not shown as it is close to the $D_s = 40$ case), MC, and exact solution. The MC reweighting results are worse than the HOTRG results with $D_s = 10$.

MC reweighting

The partition function of a spin or gauge model with a complex inverse coupling $\beta = \mathrm{Re}\beta + i\mathrm{Im}\beta$ can be expressed as (Falcioni et al. 82)

$$Z(\mathrm{Re}\beta + i\mathrm{Im}\beta)/Z(\mathrm{Re}\beta) = < \mathrm{e}^{-i\mathrm{Im}\beta E} >_{\mathrm{Re}\beta}$$

Fluctuations become of the same size as the average for too large $\mathrm{Im}\beta$

For a Gaussian distribution, the region of confidence is

$$\mathrm{Im}\beta < C\sqrt{\ln(N_{\mathrm{conf.}}/\tau)}V^{-1/2}$$

where $N_{\mathrm{conf.}}$ is the number of configurations, τ is the integrated correlation time and C a constant depending on β_0 (Alves et al. 1992)

For the 2D Ising model, the region of confidence shrinks at the same rate as the imaginary part of the zeros of the partition function (Fisher zeros) namely L^{-1} since $\nu = 1$.

Error of the free energy for different D_s

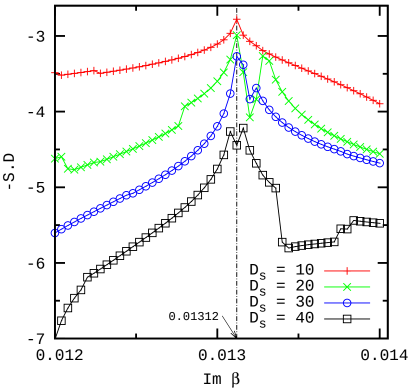


Figure: The relative error (- the number of Significant Digits) of the free energy for HOTRG calculation with $D_s = 10, 20, 30, 40$. The vertical line corresponds to the lowest zero.

Distributions of the normalized eigenvalues of Q

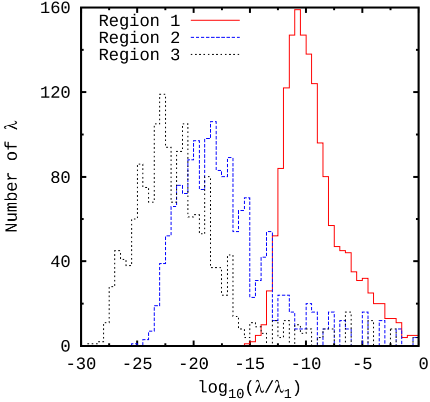


Figure: The distributions of normalized eigenvalues λ/λ_1 for β_0 in regions 1, 2, and 3 (moving away from the Fisher zero) with $D_s = 40$. There are 40^2 eigenvalues for each case. $TrQ = ||T||^2$ with $||T||$ the norm of the tensor .

Comparing with Onsager-Kaufman

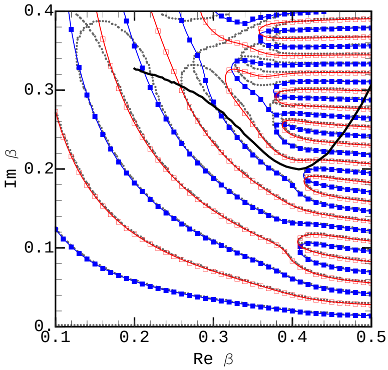


Figure: Zeros of Real (■) and Imaginary (□) part of the partition function of the Ising model at volume 8×8 from the HOTRG calculation with $D_s = 40$ are on the exact lines. Gray lines: MC reweighting solution. Thick Black curve: the "radius of confidence" for MC reweighting result, above this line, the error is large.

2D 0(2) model at complex β

$$\beta H = -\beta \sum_{<ij>} \cos(\theta_i - \theta_j).$$

Finite size scaling for the zeros calculated from the HOTRG with $D_s = 40$ and 50. For one step of the RG transformation with scaling factor $\tilde{b}$ at large L ($L \to L/\tilde{b}$), the correlation lengh scales like $\xi \to \xi/\tilde{b}$. Assuming the singular part of the partition function is a function $f(\xi/L)$. Then at the zeros,

$$Z(\beta_z) = f(z_0) = 0,$$

the values of zeros for different volumes map to the same z_0 at large L.

KT FSS

Correlation length for a K-T transition (with $\nu = 1/2$)

$$\xi = A \exp(bt^{-\nu}),$$

where $t = \beta_c - \mathrm{Re}\beta - i\mathrm{Im}\beta$ for $\mathrm{Re}\beta < \beta_c$. For small $Im\beta$:

$$|\mathrm{Re}\beta_z - \beta_c| = \frac{b^{1/\nu}}{(\ln L + a_1)^{1/\nu}}$$

$$|\mathrm{Im}\beta_z| = \frac{a_2 b^{1/\nu}}{\nu(\ln L + a_1)^{1/\nu+1}}$$

In which $a_1 = \mathrm{Re}(\ln(z_0/A))$, $a_2 = \mathrm{Im}(\ln(z_0/A))$. This implies:

$$\mathrm{Im}\beta_z = \frac{a_2}{b\nu}(\beta_c - \mathrm{Re}\beta_z)^{1+\nu}$$

Calculated zeros confirms KT FSS ($1 + \nu = 1.5$)

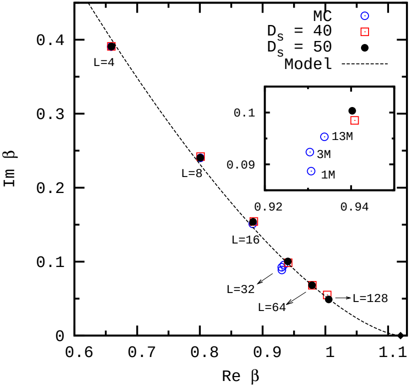


Figure: Zeros of XY model with linear size $L = 4, 8, 16, 32, 64, 128$ (from up-left to down-right) calculated from HOTRG with $D_s = 40, 50$ and zeros with $L = 4, 8, 16, 32$ from MC. The curve is a model for trajectory of the lowest zeros. Fit: $\mathrm{Im}\beta_z = 1.27986 \times (1.1199 - \mathrm{Re}\beta_z)^{1.49944}$.

The simplest example of quantum rotors

$O(2)$ model with one space and one Euclidean time direction.
The $N_x \times N_t$.sites of the lattice are labelled (x, t). We assume periodic boundary conditions in space an time.

$$Z = \int \prod_{(x,t)} \frac{d\theta_{(x,t)}}{2\pi} \mathrm{e}^{-S} \tag{1}$$

$$\begin{aligned} S = & - \beta_t \sum_{(x,t)} \cos(\theta_{(x,t+1)} - \theta_{(x,t)} + i\mu) \\ & - \beta_s \sum_{(x,t)} \cos(\theta_{(x+1,t)} - \theta_{(x,t)}). \end{aligned} \tag{2}$$

In the isotropic case, we have $\beta_s = \beta_t = \beta$.
In the limit $\beta_t >> \beta_s$ we reach the time continuum limit.
For large μ, there is a correspondence with the Bose-Hubbard model (Sachdev, Fisher, ..)

TRG formulation of $O(2)$ with a chemical potential

Using the fact that Fourier coefficients of $e^{\beta\cos\theta}$ are $I_n(\beta)$, the modified Bessel functions of the first kind, we can write

$$Z = \mathrm{Tr}\prod_{(x,t)} T^{(x,t)}_{n_x n'_x n_t, n'_t} \ ,$$

with

$$\begin{aligned} T^{(x,t)}_{n_x,n_{x'},n_t,n_{t'}} &= \sqrt{I_{n_t}(\beta_t)I_{n_{t'}}(\beta_t)\exp(\mu(n_t+n'_t)} \\ & \sqrt{I_{n_x}(\beta_s)I_{n_{x'}}(\beta_s)}\delta_{n_x+n_t,n_{x'}+n_{t'}} \ . \end{aligned} \tag{3}$$

Th indices $n_x,\ n'_x,\ n_t$ and n'_t label with some abuse of notation the four links coming out of (x,t) in the x and t direction and the trace Tr refers to the sum over all these link indices.

Transfer Matrix

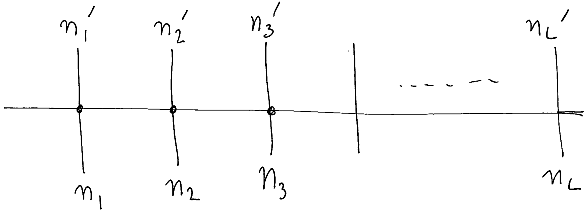
We can now consider a time slice at a given t and perform the traces over the space links (for definiteness we assume periodic boundary conditions in space). This defines a transfer matrix

$$\mathbb{T}_{(n_1,n_2,\dots n_{N_s})(n'_1,n'_2\dots n'_{N_s})} = \sum_{n_{x1}n_{x2}\dots n_{N_s}} T^{(1,t)}_{n_{xN_s}n_{x1}n_1,n'_1} T^{(2,t)}_{n_{x1}n_{x2}n_2n'_2\dots} \dots T^{(N_s,t)}_{n_{x(N_s-1)}n_{xN_s}n_{N_s}n'_{N_s}} \tag{4}$$

The indices $(n_1, n_2, \dots n_{N_s})$ represent the past and $(n'_1, n'_2 \dots n'_{N_s})$ the future and can be interpreted as the two indices of the transfer matrix.

Graphical representation of the transfer matrix

$\mathbb{T}_{(n_1,n_2,\ldots n_L)(n'_1,n'_2\ldots n'_L)}$



The partition function in terms of the transfer matrix

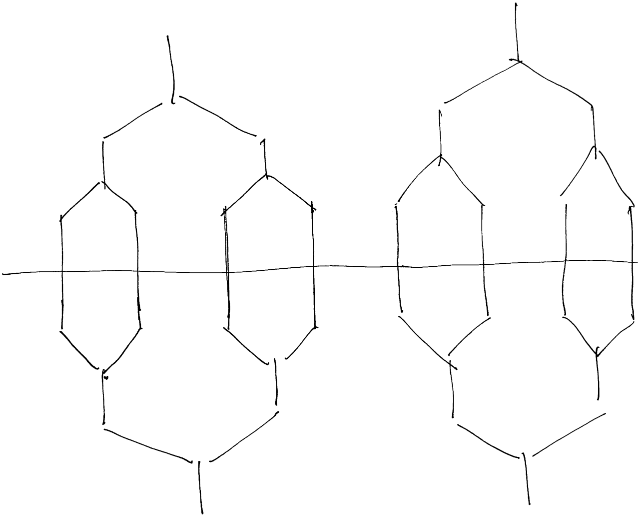
We can sum over the time links between two consecutive time slices, this amounts to squaring the transfer matrix:

$$\mathbb{T}^2_{(n_1,n_2,\ldots n_{N_s})(n''_1,n''_2\ldots n''_{N_s})} = \sum_{(n'_1,n'_2\ldots n'_{N_s})} \mathbb{T}_{(n_1,n_2,\ldots n_{N_s})(n'_1,n'_2\ldots n'_{N_s})}\mathbb{T}_{(n'_1,n'_2,\ldots n'_{N_s})(n''_1,n''_2\ldots n''_{N_s})} \ . \quad (5)$$

If the lattice size in the time direction is N_t, we obtain for periodic boundary conditions in time that

$$Z = Tr\mathbb{T}^{N_t}$$

Blocking in the spatial direction



Comparing TRG with the worm algorithm

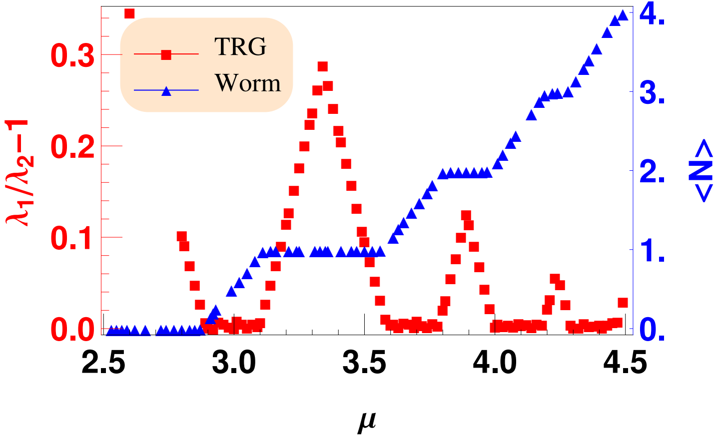


Figure: Consistency check between the worm algorithm and the TRG method at $L_x = 16$.

Comparing TRG with the worm algorithm (Yuzhi Liu)

Lx=16, beta=0.3, mu=1.57 (close to transition); <N>:
TRG
11 0.0177626
13 0.0196462
15 0.0198056
17 0.0203519
19 0.020528
21 0.0205411
23 0.0205439
25 0.0208463
27 0.0208519
Worm
0.0201141 ErrorBar[0.00060065]

Rough idea of CPU times (not apples to apples!)

CPU time scales like D_s^6
Lx=16, beta=0.1, mu=2.85.
Dbond Time(in seconds)
7 6.44
9 17
11 50
13 74
15 122
17 219
19 370
21 955
23 1435
25 2559
For the Worm, it takes about 360 seconds to get SWEEP =50000 MEAS= 500; and 3600 seconds to get SWEEP =500000 MEAS =400. The CPU time just scales with SWEEP*MEAS. The relative error are around 4 percent for 360 seconds run already.

Average occupation and eigenvalues crossing

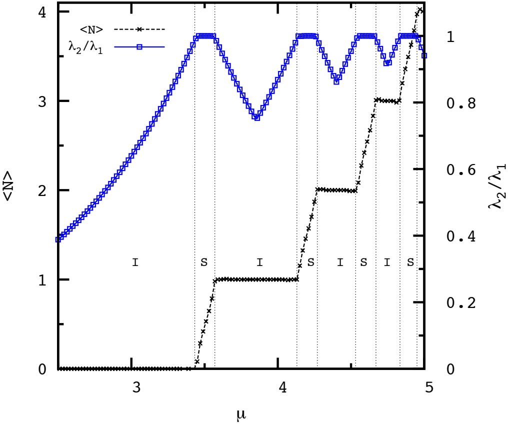


Figure: Second normalized eigenvalues (λ_2/λ_1) of the transfer matrix and particle number density at $\beta = 0.06$ from HOTRG calculation with the number of state $D_s = 15$ are shown.

Mapping the phase diagrams of the isotropic and anisotropic models

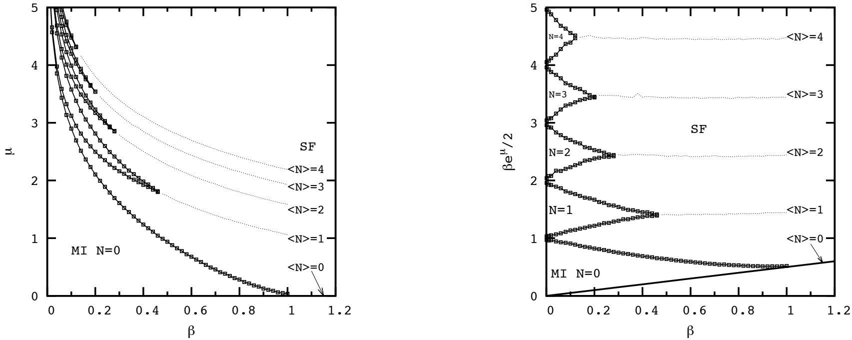


Figure: Phase diagram for 2D $O(2)$ isotropic model in β-μ plane (left) and in the β-$\beta e^{\mu}/2$ plane (right) which resembles the anisotropic case.

Anisotropic phase diagram (time continuum limit)

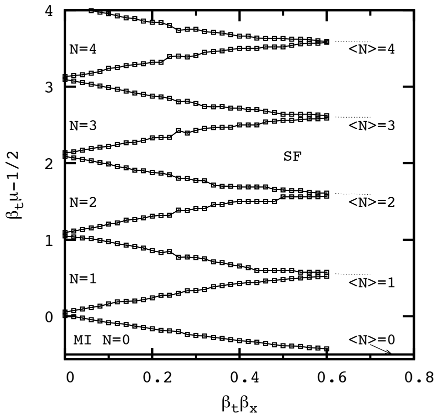


Figure: In the continuum time limit ($\beta_t >> \beta_x$), we define the effective chemical potential $\mu_e = \mu\beta_t - 1/2$ and the effective coupling $\beta_e = \beta_x\beta_t$, we find that the same insulator-superfluid phase transition behavior appears in β_e-μ_e plane. Phase diagram at 1+1D $O(2)$ model at $\beta_t = 10$ in β_e-μ_e plane.

Bose-Hubbard correspondance?

In the limit $\beta_t >> \beta_s$ (Fradkin-Susskind, Kogut, Polyakov), we obtain a continuous time Hamiltonian

$$\hat{H} = \frac{U}{2}\sum_x \hat{L}_x^2 - \tilde{\mu}\sum_x \hat{L}_x - J\sum_{<xy>} \cos(\hat{\theta}_x - \hat{\theta}_y) \ ,$$

with the commutation relations $[\hat{L}, \mathrm{e}^{\pm i\hat{\theta}}] = \pm \mathrm{e}^{\pm i\hat{\theta}}$.

When μ is large enough, one can restrict the occupation numbers to say 1 and 2 and we have the approximate correspondence $\hat{L}_x \rightarrow n_x$ and $\mathrm{e}^{i\hat{\theta}_x} \rightarrow a_x^\dagger$ (Fisher, Sachdev,..)

Another option is
$\hat{L}_x \rightarrow \hat{L}_x^{(z)}$ and $\mathrm{e}^{\pm i\hat{\theta}_x} \rightarrow \hat{L}_x^{(\pm)}$ (like gauge links)

Two species Bose-Hubbard?

In the $O(2)$ rotor Hamiltonian, the eigenvalues of L are allowed to run from positive to negative values, while in the Bose-Hubbard model they are strictly positive. The negative n states in the $O(2)$ model could correspond to antiparticles (but the Fock space generated by $a^\dagger$ and $b^\dagger$ is larger). We considered a two-species Bose-Hubbard model with commensurate filling fraction $\sum_\alpha n^\alpha = n$ on each lattice site. For $n = 2$, this model can be mapped to angular momentum $L = 1$ operators in the strong coupling limit, obtaining $L_z = \pm 1, 0$ eigenvalue states.

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_I$$

$$\mathcal{H}_I = -\sum_{\langle ij\rangle}(t_a a_i^\dagger a_j + t_b b_i^\dagger b_j + h.c.)$$

$$\mathcal{H}_0 = \frac{U}{2}\sum_{i,\alpha} n_i^\alpha(n_i^\alpha - 1) + (U - W)\sum_i n_i^a n_i^b$$

$$+\sum_{i,\alpha}(\mu + \Delta_\alpha)n_i^\alpha$$

Multi species Bose-Hubbard models have been discussed by Kuklov and Svistunov in 2003.

Effective XXZ model

In the strong coupling limit, $|U| \gg t_\alpha$, we choose the basis $|n_i^a, n_i^b\rangle$ with $n_i^a + n_i^b = n$, and treat the hopping terms as perturbation. The second order contribution $\mathcal{H}' = \mathcal{H}_I(E_0 - \mathcal{H}_0)^{-1}\mathcal{H}_I$, yields the effective Hamiltonian

$$\begin{aligned}\mathcal{H}^{eff} &= -J_z \sum_{\langle ij\rangle} L_i^z L_j^z - J_{xy} \sum_{\langle ij\rangle}(L_i^x L_j^x + L_i^y L_j^y) \\ &\quad + h\sum_i L_i^z + W\sum_i (L_i^z)^2\end{aligned}$$

$J_z = \frac{|t_a|^2+|t_b|^2}{U}, \quad J_{xy} = \frac{2t_a t_b}{U}, h = (\Delta_a - \frac{2(n+1)|t_a|^2}{U}) - (\Delta_b - \frac{2(n+1)|t_b|^2}{U}).$

Except for the first term, this looks like a restriction of our model to $n = -1, 0, 1$

TRG with restriction $n = -1, 0, 1$ for the initial tensor

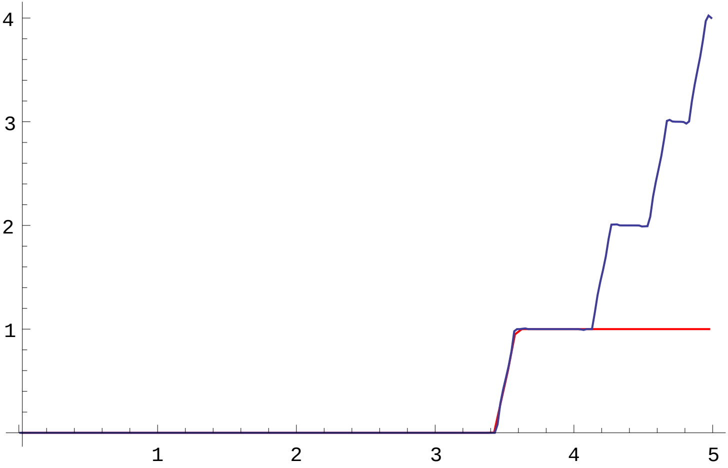


Figure: <N> versus μ for $\beta = 0.06$ with restriction $n = -1, 0, 1$ at each link for the initial tensor (last minute work by Haiyuan Zou).

Real time evolution?

- Eigenvenvalues of the transfer matrix :$A\mathrm{e}^{-aE_n} \rightarrow A\mathrm{e}^{-iaE_n}$
- Eigenvectors $|E_n\rangle$are the same
- $\hat{U}(t) \simeq \sum_{n=1}^{D_s} \mathrm{e}^{-itE_n}|E_n\rangle\langle E_n|$
- Easy tests for short time evolution?

Conclusions

The Tensor RG allows us to

- formulate most lattice models
- interpolate between discrete and continuous time
- deal with the sign problems associated with complex β and complex chemical potential
- understand the merging of the phase diagrams in the time continuum limit
- get a better insight on the classical-quantum connection

Thanks!!