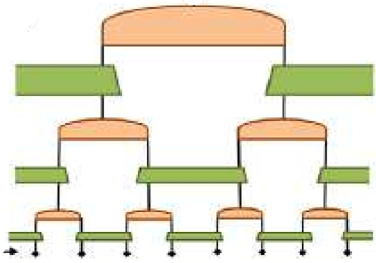
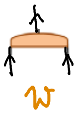
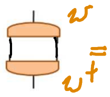
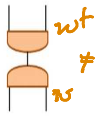
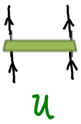
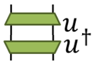
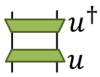
TNR: Tensor network renormalization
MERA: Multi-scale Entanglement Renorm. Ansatz



1. Motivation

- coarse-grain / truncate using isometries
- disentangle neighboring sites using unitary 'disentanglers'

Isometry  : $w^\dagger w = \mathbb{1}$:  $= \,|$, $w w^\dagger \neq \mathbb{1}$:  $\neq \,|\;|$

Unitary disentangler  : $u^\dagger u = \mathbb{1}$:  $= \,|\;|$, $u u^\dagger = \mathbb{1}$  $= \,|\;|$

[Vidal2007] (Dec. 1, 2006) Original idea for entanglement renormalization; application to transverse field quantum Ising model.

[Vidal2008] (Dec. 3, 2006) Detailed proposal for multi-scale entanglement renormalization ansatz (MERA); shows that contractions in MERA can be computed efficiently, $\mathcal{O}(N)$

[Evenbly2009] Evenbly & Vidal: Describe MERA algorithm in detail. Optimization of disentanglers: Sec. IV.

[Aguado2008] Aguado & Vidal: Argue that MERA provides natural description for topological states of matter.
[Koenig2009] Koenig, Reichardt, Vidal: Show Kitaev toric code ground state can be written as a MERA.

[Rizzi2009] Rizzi, Montangero, Vidal: time-dependent MERA (tMERA); optimization via time evolution.

[Cincio2008] Cincio, Dziarmaga, Rams: MERA for 2D system: quantum Ising model.

[Pfeifer2009] Pfeifer, Evenbly, Vidal: MERA for scale-invariant systems.

[Evenbly2009a] Evenbly, Vidal: MERA in 2D: bring down cost from $\mathcal{O}(\chi^{28})$ to $\mathcal{O}(\chi^{16})$

[Evenbly2010] Corboz, Evenbly, Verstraete, Vidal: Fermionic MERA

today: → [Evenbly2015] Evenbly, Vidal: propose tensor network renormalization (TNR): improve TRG via disentanglers.

→ [Evenbly2015a] Evenbly, Vidal: show that TNR, applied to $e^{-\beta H}|\psi\rangle$, yields MERA.

→ [Evenbly2017] Describe TNR algorithm in detail. Optimization of via 'projective truncations': Sec. III.C.

[Evenbly2017a] Implicitly disentangled renormalization (IDR): cheaper way to implement disentangling.

[Evenbly2016] Evenbly, White: Entanglement renormalization and wavelets. Construct the first known analytic MERA for a critical system (critical Ising model).

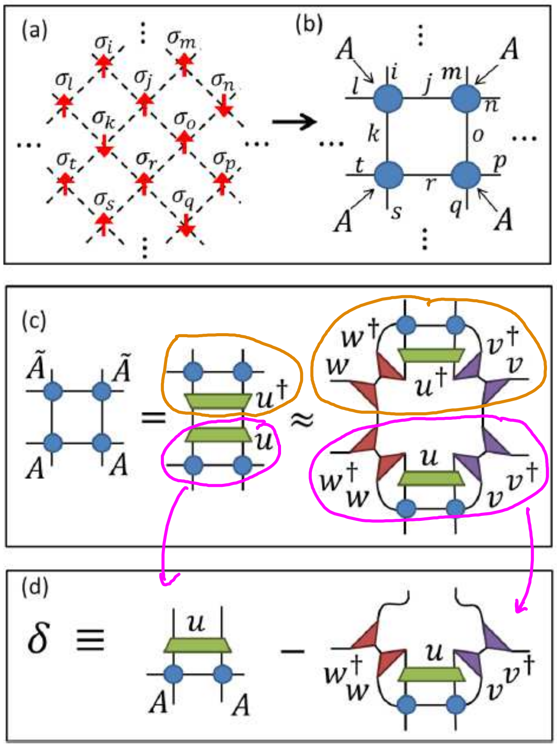
[Haegeman2018] Haegeman, Swingle, Walter, Cotler, Evenbly, Schulz: Rigorous free-fermion renormalization from wavelet theory.

[Evenbly2018] Gauge fixing, canonical forms, and optimal trunctations in tensor networks with closed loops.

2. Tensor network renormalization (TNR) [Evenbly2015]

Goal: improve TRG by fully removing local correlations, including those in local loops.

Strategy: devise truncation scheme involving not only isometries, but also unitary disentanglers.

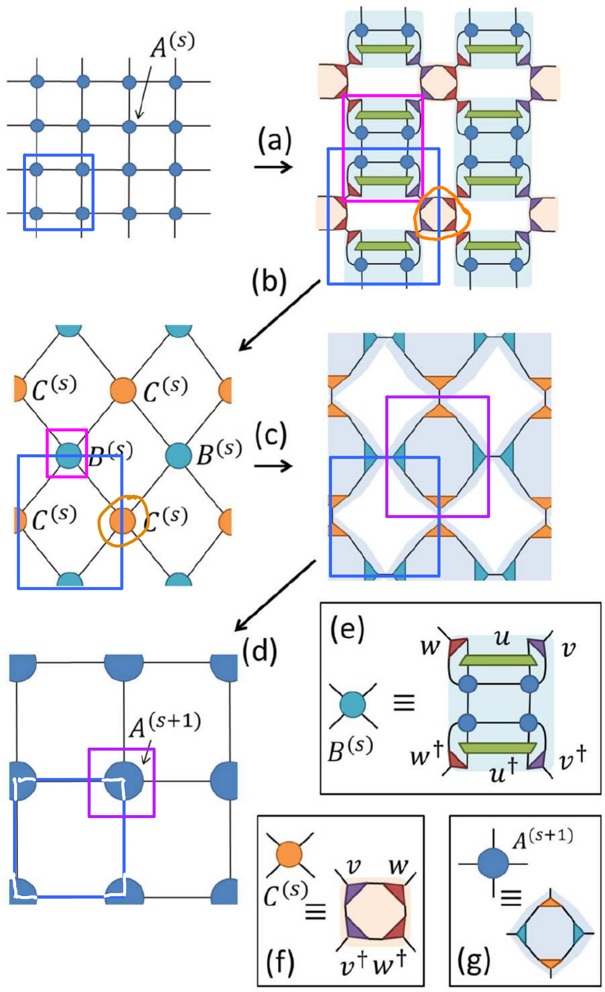


(a) Original lattice for classical spin system on square lattice.

(b) Tensor network representation of partition function.

(c) Disentangle internal loop correlations along a plaquette using a combination of disentanglers u (to disentangle neighboring sites) and isometries $v, w^\dagger$ (to truncate).

(d) Optimize isometries v, w and unitaries u by minimizing the truncation error introduced by isometries. (For details, see 'projective truncations' in Section MERA.4 below.)



One TNR iteration step:

(a) Disentangle internal loop correlations along every other plaquette.

(b) Contract plaquettes into 4-leg tensors $B^{(s)}$, and their links into 4-leg tensors $C^{(s)}$, as defined in (e), (f).

(c) Standard diagonal TRG decomposition.

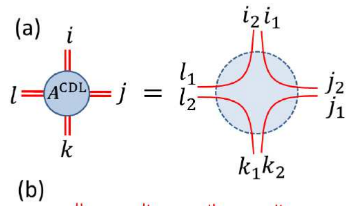
(d) Standard TRG contraction to yield new 4-leg tensor $A^{(s+1)}$, as defined in (g).

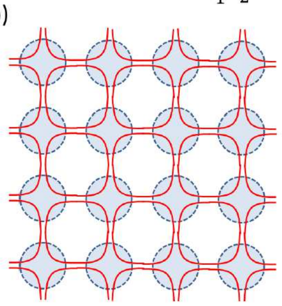
Remark: the details of this algorithm are presented in [Evenbly2017, Figs. 5-7]. There, the 4-leg tensor $C^{(s)}$ is not constructed explicitly, only implicitly. This is explained in Sec. MERA.4 below.

Example: Corner double line (CDL) tensors: how TNR encodes internal loop correlations

CDL model:

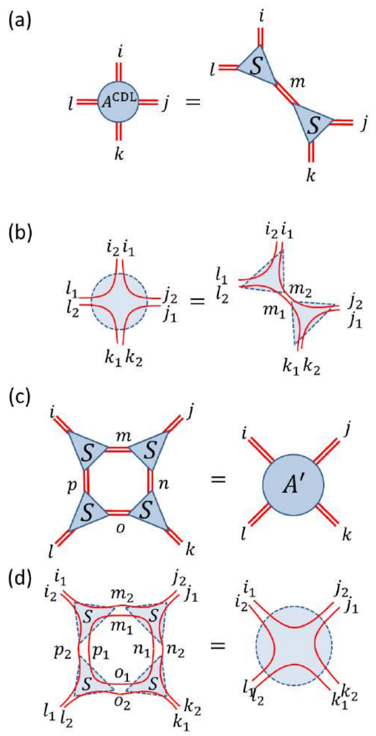
Red lines symbolize local correlations.
Initial A-tensor features 4 red lines.





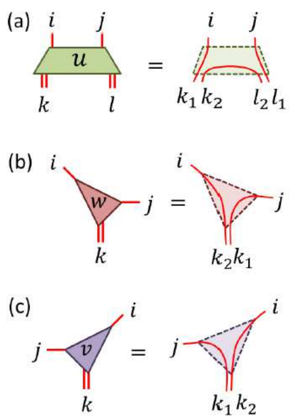
CDL tensors are fixed point of TRG, illustrating that TRG fails to fully encode local loop correlations.

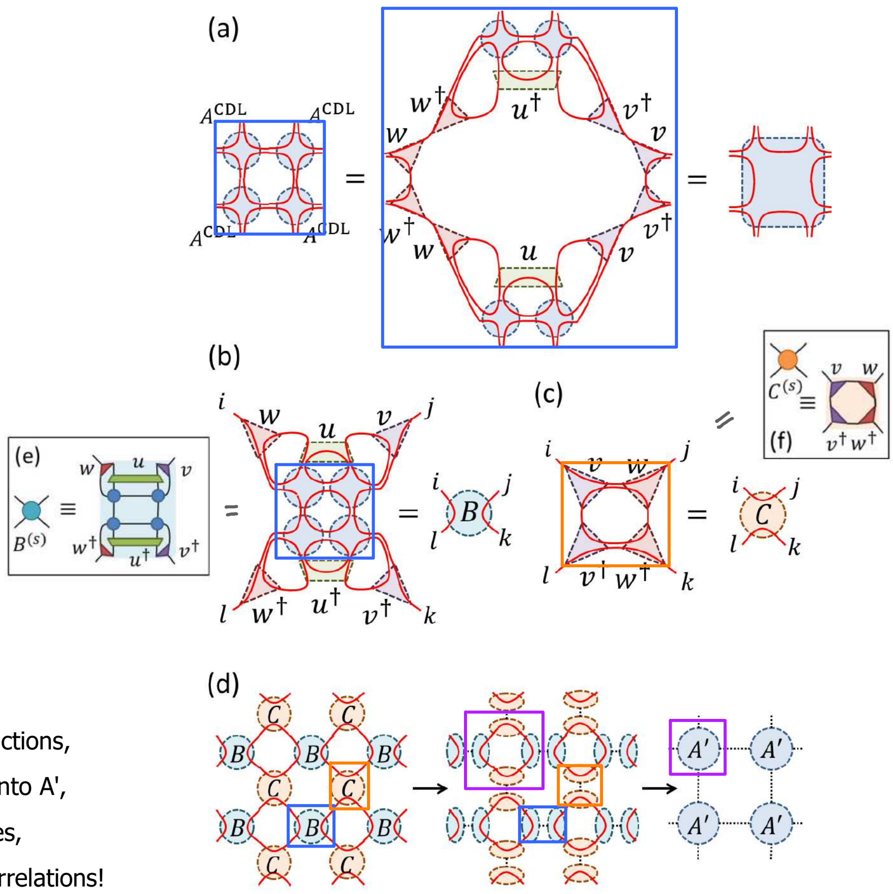
TRG treatment of CDL model:



A' again features 4 red lines, hence renormalized tensor is again of CDL type

TNR treatment of CDL model:





B involves only <u>two</u> red lines after all internal contractions in its definition have been performed. Ditto for C.

After SVD of B and C and further contractions, combining two halves each of B and C into A', the A' are not connected by any red lines, hence each A' fully encodes all local correlations!

Results: TNR for square lattice Ising model with 2^{39} spins

at criticality

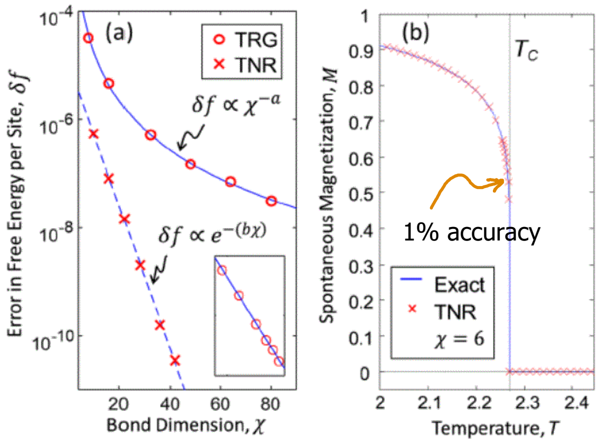
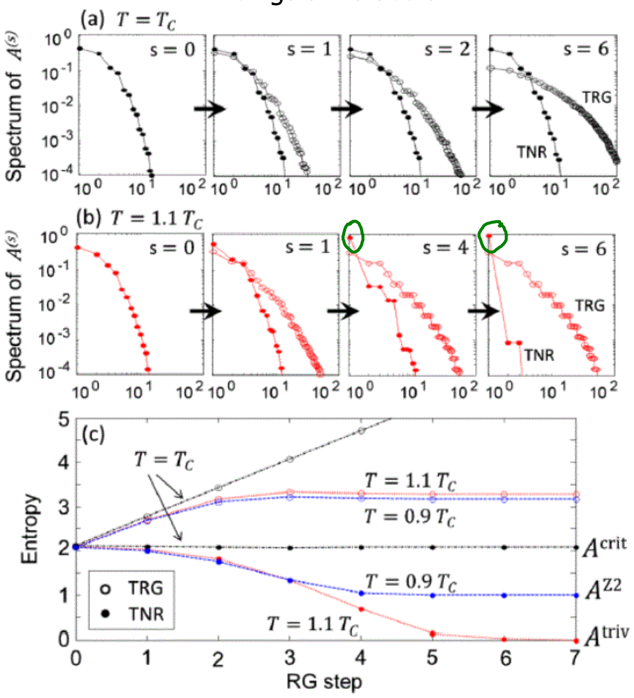


FIG. 3 (color online). Benchmark results for the square lattice Ising model on a lattice with 2^{39} spins. (a) Relative error in the free energy per site δf at the critical temperature T_c, comparing the TRG and TNR over a range of bond dimensions χ. The TRG errors fit $\delta f \propto \chi^{-3.02}$ (the inset displays them using log-log axes), while TNR errors fit $\delta f \propto \exp(-0.305\chi)$. Extrapolation suggests that the TRG would require bond dimension $\chi \approx 750$ to match the accuracy of the $\chi = 42$ TNR result. (b) Spontaneous magnetization $M(T)$ obtained with TNR with $\chi = 6$ [30]. Even very close to the critical temperature, $T = 0.9994T_c$, the magnetization $M \approx 0.48$ is reproduced to within 1% accuracy.

singular values of $A^{(s)}$ TRG iteration number



scale invariant!

Singular value spectrum reaches fixed point for TNR, but not for TRG.

Away from criticality, a single singular value dominates

Entanglement entropy is scale invariant at criticality (black dots).

3. Projective truncations

Truncation for original formulation of MERA [Vidal2007], [Vidal2008] is described in [Evenbly2009, Sec. IV].

We here discuss truncation strategy for TNR, described in [Evenbly2017, Sec. III]. When using TNR to generate MERA, this strategy is simpler and more efficient than previous one from [Evenbly2009, Sec. IV].

Local approximation for 4-index tensor

General idea: replacement of $\mathcal{F}$ by $\tilde{\mathcal{F}}$ (differing only locally), is allowed if $\varepsilon \equiv \| \mathcal{F} - \tilde{\mathcal{F}} \|^2$ is small. (1)

Hilbert-Schmidt norm: $\| A \|^2 = \mathrm{tTr}(A \otimes A^\dagger)$ tTr = tensor trace = contract all matching indices (2)

This is generalization of $|\,|\psi\rangle|^2 = \langle\psi|\psi\rangle = \mathrm{Tr}[\,|\psi\rangle\langle\psi|\,] = \mathrm{Tr}\,\rho$

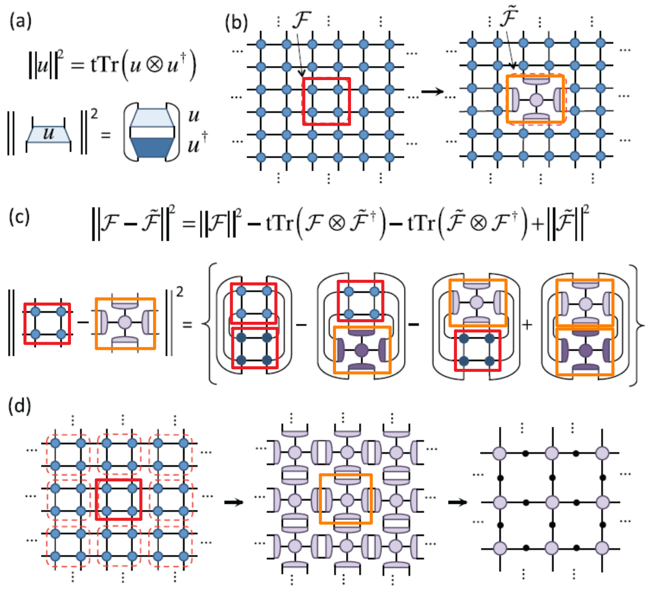


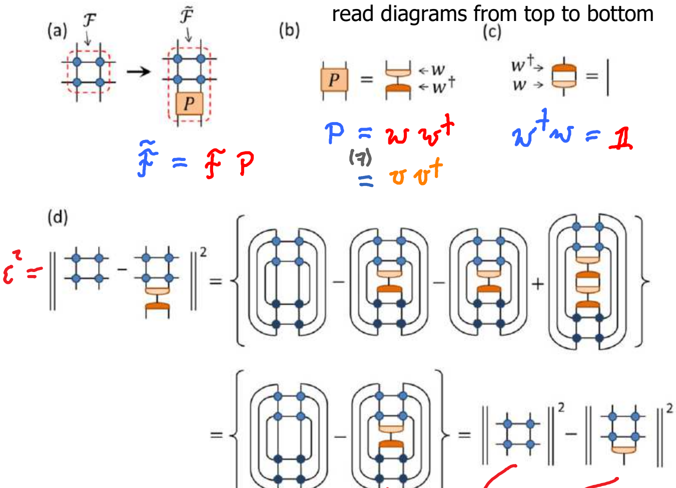
FIG. 3. (a) Depiction of (the square of) the Hilbert-Schmidt norm of a four-index tensor u. Note that a darker shade is used to represent the conjugate tensor, which is also drawn with opposite vertical orientation. (b) Given a square lattice tensor network, we wish to replace a 2×2 block of tensors $\mathcal{F}$ from the network with a different subnetwork of tensors $\tilde{\mathcal{F}}$. (c) The square of the difference between $\mathcal{F}$ and $\tilde{\mathcal{F}}$ under the Hilbert-Schmidt norm is depicted, where darker shades are used to depict conjugate tensors, which are drawn with opposite vertical orientation to regular tensors. The replacement in (b) is valid if the difference $\|\mathcal{F} - \tilde{\mathcal{F}}\|$ is sufficiently small. (d) Assuming that the local square-lattice network is homogeneous, one can replace $\mathcal{F}$ with $\tilde{\mathcal{F}}$ in all 2×2 blocks. A coarser square-lattice network is obtained after contraction between pairs of three-index tensors.

If $\tilde{\mathcal{F}}$ is obtained from $\mathcal{F}$ via truncated SVD, then $\varepsilon = \sum_i{}' s_i^2$ (3)

sum over discarded singular values

TNR introduces more general class of truncations: 'projective truncations'

but more complicated choices are possible and will be used in TNR.4



$\tilde{\mathcal{F}} = \mathcal{F} P$

$P = w w^\dagger$ (7) $= v v^\dagger$

$w^\dagger w = \mathbb{1}$

$\varepsilon^2 =$

$$\varepsilon = \sqrt{\|\mathcal{F}\|^2 - \|\mathcal{F}P\|^2} = \sqrt{\|\mathcal{F}\|^2 - \|\mathcal{F}w\|^2}$$

FIG. 4. (a) In a *projective truncation* a subnetwork $\mathcal{F}$ is replaced by a new subnetwork $\tilde{\mathcal{F}}$, which consists of a projector P applied to the original subnetwork, i.e., $\tilde{\mathcal{F}} = \mathcal{F}P$. (b) Here we assume that P is decomposed as a product of an isometric tensor w and its conjugate, $P = ww^\dagger$. (c) By definition, isometry w contracts to identity with its conjugate, $w^\dagger w = \mathbb{I}$. (d) The square of the error in a projective truncation is expanded as a sum of four terms; however given that $P^2 = P$, two of the terms cancel; see also Eq. (18). (e) The environment Γ_w of isometry w is defined as the network that results by removing a single instance of w from $\|\mathcal{F}w\|^2$; see also Eq. (20). (f) By construction, the contraction of w and its environment Γ_w is equal to $\|\mathcal{F}w\|^2$. (g) Environment Γ_w is decomposed, via singular value decomposition (SVD), into a product of isometric tensors u, v, and diagonal matrix s.

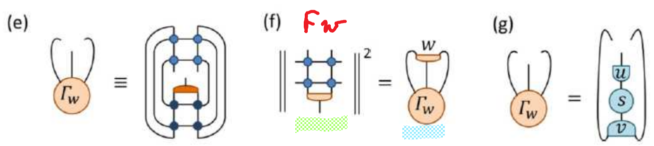
according to the same partitioning of indices for which w is isometric

To minimize error, maximize

$$\|\mathcal{F}w\|^2 = \mathrm{tTr}(\Gamma_w \otimes w) \qquad (4)$$

'environment of w'

with constraint $w^\dagger w = \mathbb{1}$. 'Linearize': hold $w^\dagger$ fixed, optimize w



SVD of environment: $\Gamma_w = u s v^\dagger$ (5) 'align basis vectors'

Optimal choice for updated w: $w = v u^\dagger$ (6)

see [Evenbly2009, Eq. (67)]

Check: $w^\dagger w = u v^\dagger v u^\dagger = \mathbb{1}$ ✓

Then $\|\mathcal{F}w\|^2 = \mathrm{tTr}(\Gamma_w \otimes w) = \mathrm{tTr}(u s v^\dagger w) \xrightarrow{\text{update}} \mathrm{tTr}(u s v^\dagger v u^\dagger) = \mathrm{tTr}(s)$ (7)

Iterate this until convergence of singular value spectrum of Γ_w (typically, hundreds of iterations are needed)

insert new $w = v u^\dagger$ into (e), recompute Γ_w, do SVD, etc.

[Note: in contrast to Gilt [Hauru2018], local truncation does not 'know' about environment of $\mathcal{F}$ itself. Hence, Gilt truncation is 'smarter' than projective truncation, and needs much fewer iterations.]

4. Details of TNR scheme for square lattice. [Evenbly2017, Sec. IV]

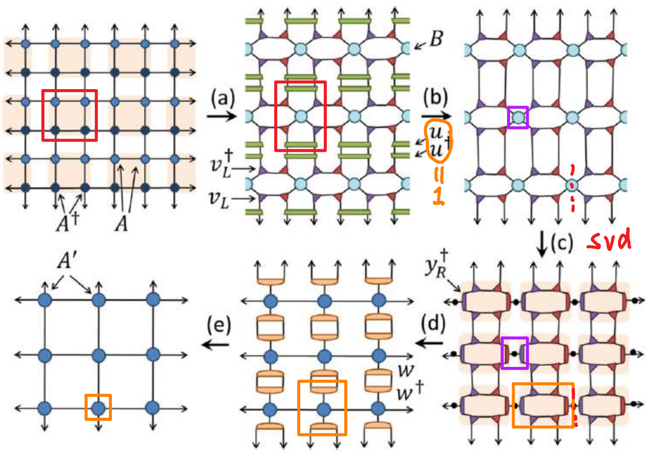


FIG. 5. The sequence of coarse-graining steps used in the binary TNR scheme in order to map an initial square lattice of tensors A, where every second row of tensors has been conjugated as described in Appendix D, to a coarser square lattice composed of tensors A'. (a) A projective truncation is made on all 2×2 blocks of tensors; see Figs. 6(a)–6(c). (b) Conjugate pairs of disentanglers u are contracted to identity. (c) A projective truncation is made on all B tensors; see Figs. 6(d)–6(e). (d) A final projective truncation is made; see Figs. 6(f) and 6(g) for details. (e) Conjugate pairs of isometries w are contracted to identity.

[A gauge transformation was made on every second row, equivalent to flipping tensor indices and taking complex conjugation, to replace A by $A^\dagger$.]

Here several different projectors are used, all aiming to disentangle pairs of legs:

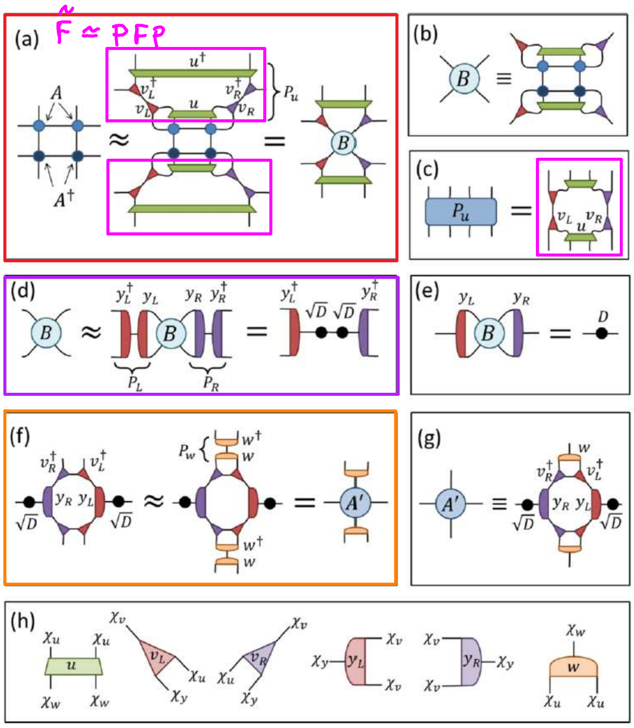


FIG. 6. (a) Details of the projective truncation made at the first step of the TNR iteration; here two copies of a projector P_u, which is composed of a product isometric and unitary tensors, are applied to a 2×2 block of A tensors. (b) Definition of four-index tensor B. (c) Projector P_u is formed from isometries v_L, v_R and disentangler u (and their conjugates). (d) Details of the projective truncation made at the second step of the TNR iteration. (e) Definition of matrix D. (f) Details of the projective truncation made at the third step of the TNR iteration. (g) Definition of new four-index tensor A', copies of which comprise the coarse-grained square-lattice tensor network. (h) Delineation of the different dimensions $\{\chi_u, \chi_v, \chi_w, \chi_y\}$ of indices on tensors $\{u, v_L, v_R, y_L, y_R, w\}$.

(a-c): Projector P is used twice. It is built from a unitary, u , and two isometries, v_L , v_R

(d-e) Two projectors are used, built from two isometries, y_L, y_R

(f-g) One projector is used twice, built from the isometry w

Each of these objects has a corresponding environment:

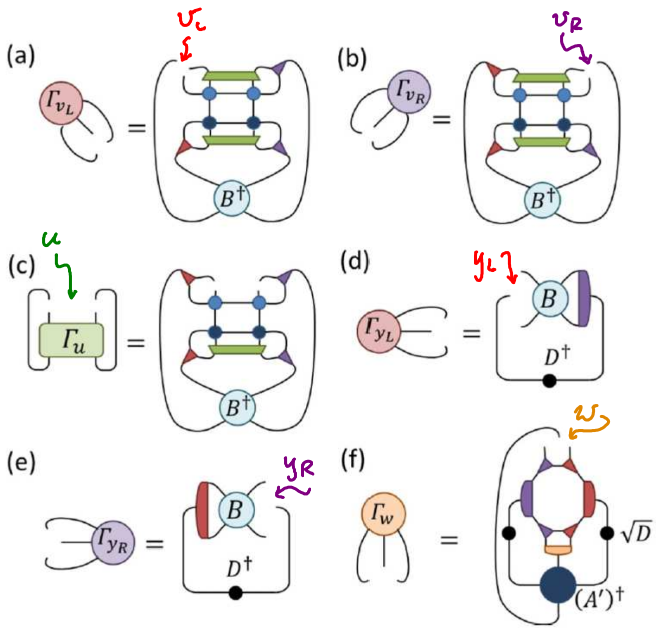


FIG. 7. The linearized environments of tensors $\{v_L, v_R, u, y_L, y_R, w\}$ involved in an iteration of the binary TNR scheme. (a)–(c) Environments Γ_{v_L}, Γ_{v_R}, and Γ_u of the isometries v_L, v_R and disentangler u involved in the first projective truncation of the TNR iteration, as detailed in Fig. 6(a). (d) and (e) Environments Γ_{y_L} and Γ_{y_R} of isometries y_L and y_R from the second projective truncation of the TNR iteration, as detailed in Fig. 6(d). (f) Environment Γ_w of isometry w from the third projective truncation of the TNR iteration, as detailed in Fig. 6(f).

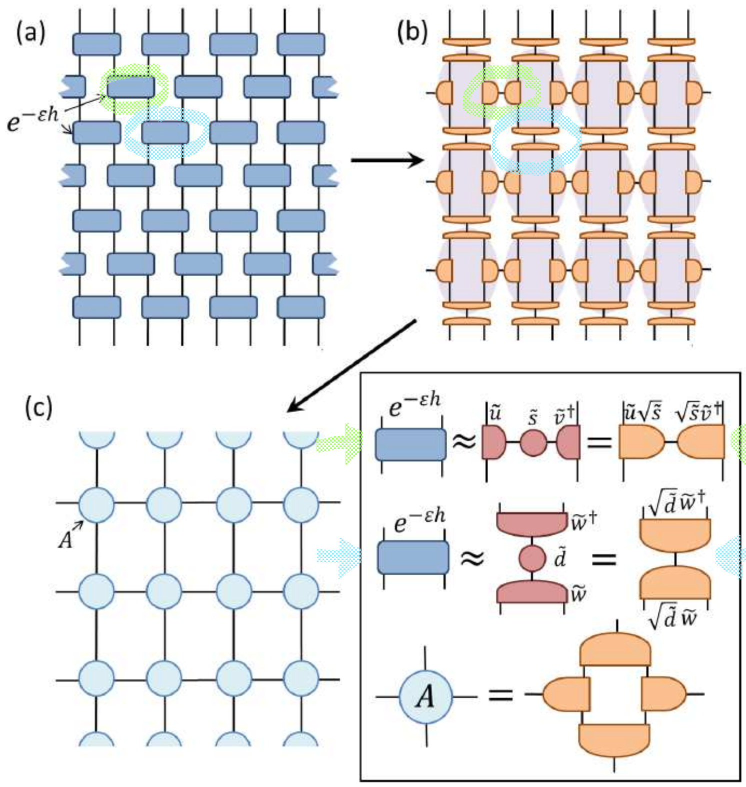
u, v_L, v_R, y_L, y_R, w must be iteratively optimized one after the other, hundreds of times, until convergence is achieved.

5. TNR results in MERA   [Evenbly2015a]   TNR.5

Original formulation of MERA: [Vidal2007], [Vidal2008]; truncation used there: [Evenbly2009, Sec. IV].

We here discuss alternative scheme for obtaining MERA via TNR, proposed in [Evenbly2015a], with truncation strategy described in [Evenbly2017, Sec. III]. This strategy is simpler and more efficient than previous one from [Evenbly2009, Sec. IV].

Express imaginary time evolution of 1D quantum chain as tensor network:



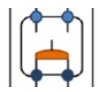
[Evenbly2015a, Fig. C2] = [Evenbly2017, Fig. 2]

$|\psi\rangle = e^{-\beta H}|\psi_0\rangle$

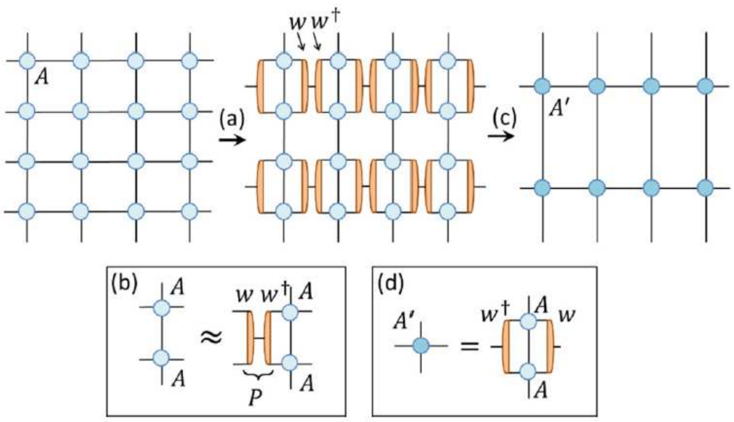
(a) Trotterize imaginary time evolution.

(b) SVD all Trotter gates, horizontally in one layer, vertically in the next.

(c) Contract four half-isometries to obtain effective square lattice with 4-leg tensor A.



Coarse-grain a few times in the Euclidean time direction, to make network more 'isotropic':

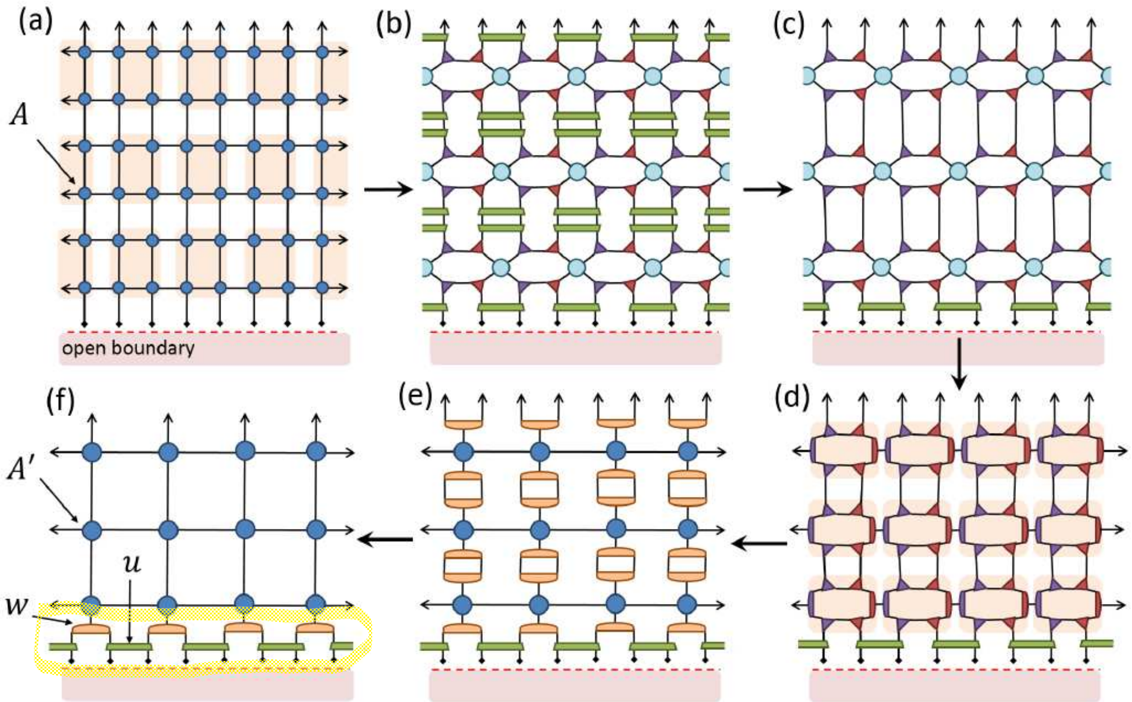


[Evenbly2017, end of Sec. II, and Fig. 17]

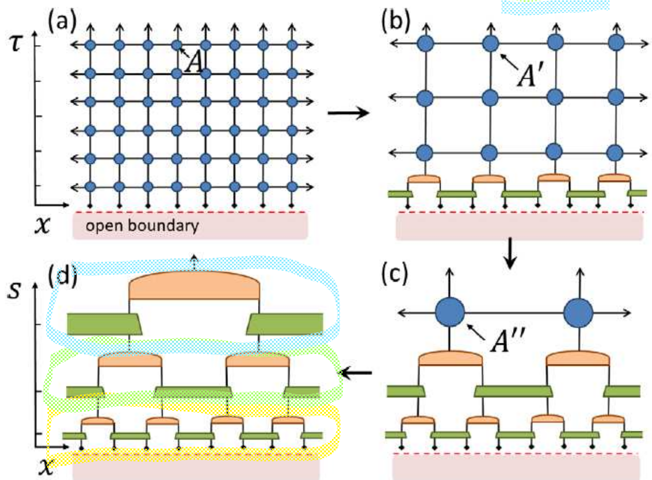
Goal of this step: to make 'bonds' in horizontal/vertical directions become comparable in strength. To 'measure bond strength', compute spectra of transfer matrices

Their decay should match as closely as possible.

Apply one TNR step (see MERA.3); this yields boundary layer of MERA: [Evenbly2015a, Fig. C3]



Iteratively apply RNT; each step yields new layer of MERA:



[Evenbly2015a, Fig. 1]

FIG. 1 (color online). (a) Tensor network, the ground state $|\Psi\rangle$ of H on an infinite lattice. It is made of copies of tensor A and restricted to the upper half plane (x, τ^+), with a row of open indices at $\tau = 0$. (b) By coarse graining the tensor network while leaving the open indices untouched, we obtain a new tensor network with tensors A' together with one row of disentanglers and isometries. (c) Further coarse graining of the tensor network produces new coarse-grained tensors A'' and a second layer of disentanglers and isometries. (d) By iteration we obtain a full MERA approximation for state $|\Psi\rangle$.

Similar constructions are possible for thermal Boltzmann factor, [Evenbly2015a, Fig. 2]

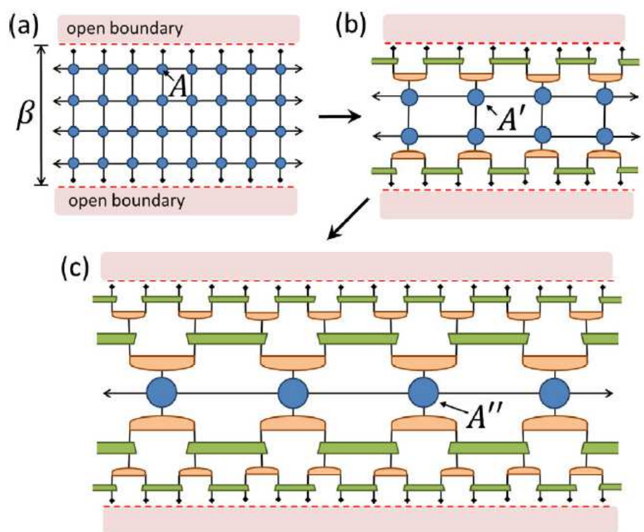


FIG. 2 (color online). (a) Tensor network on an infinite strip of finite width β, with two rows of open indices. It is proportional to the thermal state, $e^{-\beta H}/Z$. (b) By coarse graining the tensor network while leaving the open indices untouched, we obtain a new tensor network with tensors A' together with an upper and lower row of disentanglers and isometries. (c) Further coarse graining produces a thermal MERA.

and for ground state of a periodic strip of finite length L: [Evenbly2015a, Fig. 3]

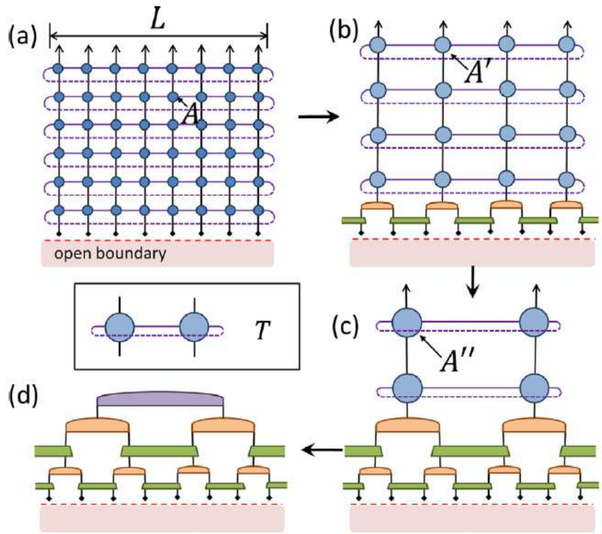


FIG. 3 (color online). (a) Tensor network on a semi-infinite vertical cylinder of finite width L and with a row of open indices, proportional to the ground state of H on a periodic chain made of L sites. (b) Result of coarse graining the initial tensor network while not touching its open indices. (c) MERA connected to a semi-infinite vertical cylinder of $O(1)$ width. Inset: Transfer matrix T of this cylinder. The eigenvectors of T with the largest eigenvalues correspond to the low energy eigenstates of H. (d) MERA for the ground state or low energy excited states of H, where the top tensor is an eigenvector of the transfer matrix T.

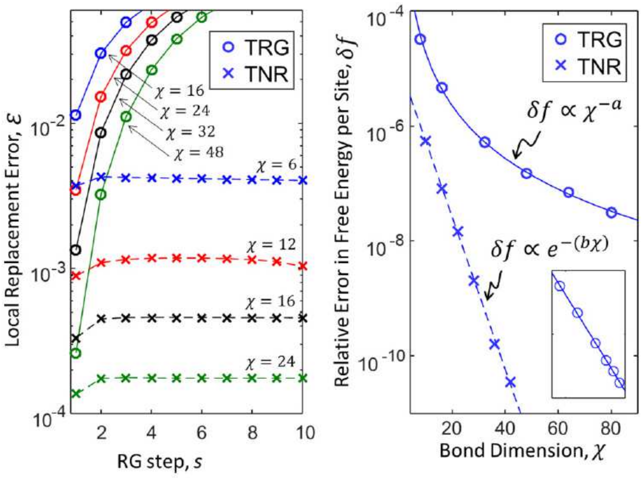


FIG. 11. (a) Comparison between TRG and TNR of the truncation error ε, as defined in Eq. (11), as a function of RG step s in the $2D$ classical Ising model at critical temperature T_c. While increasing the bond dimension χ gives smaller truncation errors, the truncation errors still grow quickly as a function of RG step s under TRG. Conversely, truncation errors remain stable under coarse-graining with TNR. (b) Relative error in the free energy per site δf at the critical temperature T_c, comparing TRG and TNR over a range of bond dimensions χ. The error from TRG is seen to diminish polynomially with bond dimension, with fit $\delta f \propto \chi^{-3.02}$ (where the inset displays the same TRG data with logarithmic scales on both axes), while the error from TNR diminishes exponentially with bond dimension, with fit $\delta f \propto \exp(-0.305\chi)$. Extrapolation suggests that TRG would need bond dimension $\chi \approx 750$ to match the accuracy of the $\chi = 42$ TNR result.

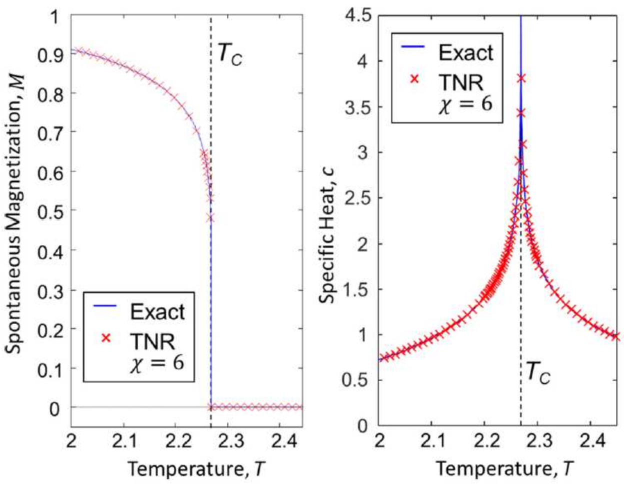


FIG. 12. (a) Spontaneous magnetization $M(T)$ of the $2D$ classical Ising model near critical temperature T_c, both exact and obtained with TNR with $\chi = 6$. Even very close to the critical temperature, $T = 0.9994\ T_c$, the magnetization $M \approx 0.48$ is reproduced to within 1% accuracy. (b) Specific heat, $c(T) = -T\frac{\partial^2 f}{\partial T^2}$, both exact and obtained using TNR with $\chi = 6$.

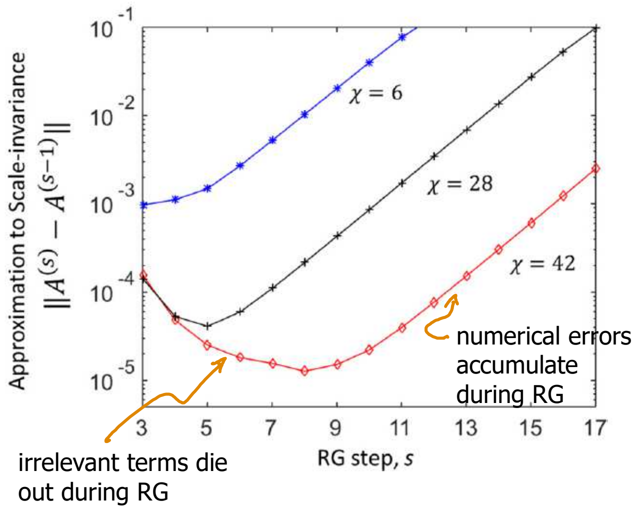


FIG. 13. The precision with which TNR approximates a scale-invariant fixed-point tensor for the $2D$ classical Ising model at critical temperature T_c is examined by comparing the difference between tensors produced by successive TNR iterations $\delta^{(s)} \equiv \|A^{(s)} - A^{(s-1)}\|$, where tensors have been normalized such that $\|A^{(s)}\| = 1$. The precision with which scale invariance is approximated in the initial RG steps (small s) is limited by the presence of RG-irrelevant terms in the lattice Hamiltonian that break scale invariance at short-distance scales, while numerical truncation errors, which can be thought of as introducing RG-relevant terms, shift the system from criticality (and thus scale invariance) in the limit of many RG steps s.

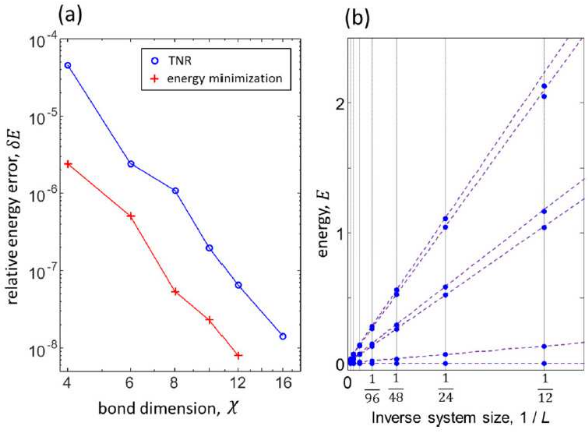


FIG. 15. (a) Relative error in the energy of scale-invariant MERAs optimized for the ground state of the $1D$ quantum Ising model at criticality in terms of bond dimension χ, comparing MERAs optimized using TNR to those optimized using variational energy minimization. Energy minimization produces MERAs with a more accurate approximation to the ground state energy, but is significantly more computationally expensive [with a computational cost that scales as $O(\chi^9)$ versus as $O(\chi^6)$ for TNR]. (b) Low-energy eigenvalues of the $1D$ quantum Ising model at criticality as a function of $1/L$, computed with $\chi = 12$ TNR. Discontinuous lines correspond to the finite-size CFT prediction, which ignores corrections of order L^{-2}.