

# Tensor networks and quantum field theory

Tobias J. Osborne<sup>1</sup>

<sup>1</sup>*Institut für Theoretische Physik, Leibniz Universität Hannover, Appelstr. 2, 30167 Hannover, Germany*

Here we describe how tensor network states — a quantum information theory inspired tool originally arising in the study of quantum spin systems — can be used to understand the physics of quantum fields. In particular, we review the tensor network formalism and outline its application to the study of ground state physics and scattering amplitudes. We describe two approaches, one based on the Lie-Trotter expansion within the hamiltonian formalism, and a second based on the lagrangian formalism which represents the path integral as a tensor network. Our aim in this paper is to present recent developments in condensed matter theory with a quantum field theory audience in mind.

*Introduction.* — Quantum field theory is now understood as an effective theory describing the large scale low energy physics [1, 2]. By exploiting the correspondence between a statistical physics system at criticality and a quantum field Wilson in one fell swoop resolved all the troubling “infinities” plaguing quantum field theory. And so it is that quantum field theory has become a tool par excellence in physics; to understand the large-scale effective physics of any system we need only identify the fixed point — often uniquely defined by the symmetries of the system — to which it belongs and, often enough, the resulting QFT describing that situation is amenable to perturbation theory allowed us to deploy the apparatus of Feynman diagrams.

Tensor network papers: [3][4–6][7] [8][9][10]

Perturbative QFT is useful and powerful but nonperturbative field theory still contains many mysteries: tantalising hints, via holography, of the power of QFT. Complex network of relationships via holographic duality.

The lattice is a good regulator allowing computers to be brought to bear on deep problems. Amazing insights [11] from confinement to the hadron spectrum [12].

The lattice

The path integral is the way to work with QFT

[13][14][15][16][17][18][19][20][21][22][23]

Here we must cite Swingle, Vidal, Verstraete, Maldacena, Susskind, Takayanagi etc.....

Important points:

1. Path integral formalism to calculate  $n$ -point functions.
2. Path integral on the lattice = statistical mechanics.
3. Method 1: trotter
4. Method 2: Stat. mech. = lifshitz PEPS.
5. Scattering amplitudes as  $n$ -point functions.
6. Removing the cutoff: 2nd (higher) order phase transition; scaling limit.
7. Variational calculation of  $n$ -point functions
8. Path integrals+perturbation theory: good for weak coupling bad for strong coupling; TNS good for strong coupling bad for weak.

Since their emergence in 1987 in the condensed matter theory literature there has been an amazing amount of progress in the development of tools, both analytic and numeric, for the study of tensor network states. We cannot do proper justice to the literature here, and we only summarise a couple of the highlights.

In this paper we argue that tensor network methods are reaching a maturity so that they may be profitably employed as an alternative to the path integral in the study of correlated quantum fields. Here

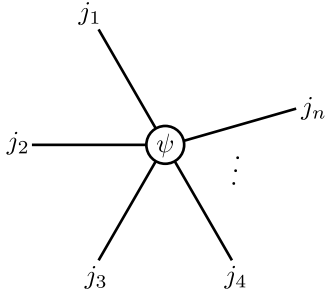
In this sense tensor-network methods have become a natural alternative to the path integral appropriate for

The purpose of this letter is to interpret QFT in the language of tensor networks with a view to formalising possible approaches to studying holographic dualities and the calculation of To this end we describe two natural approaches to studying the statics and dynamics of QFT via TNS, namely, in the hamiltonian setting familiar to condensed matter theory and a second approach by directly representing the full path integral as a TNS. We then argue how to obtain continuum limits and briefly

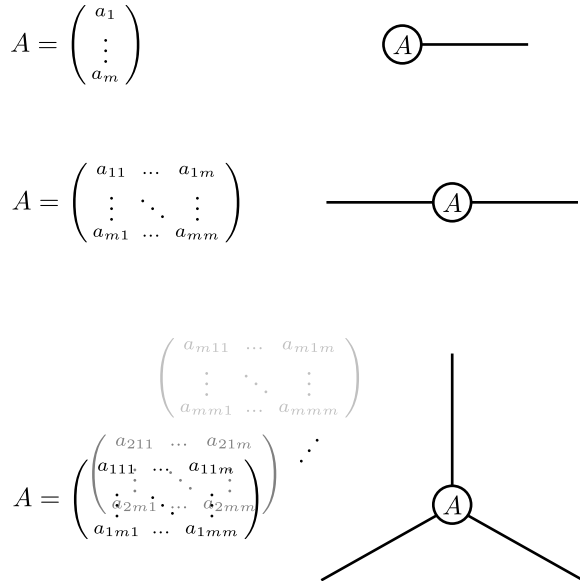
*What is a tensor network?*—Tensor networks provide an economic framework to parametrise tensors with many indices. Here we provide a brief overview of the tensor network formalism and sketch some calculational techniques. For some recent reviews of this material see, e.g., [4–6].

We begin our discussion by considering  $n$ -index tensors: these are, for us, nothing more than collections  $\psi$  of  $d_1 d_2 \cdots d_k$  complex numbers  $\psi_{j_1 j_2 \cdots j_n}$ , where  $j_k = 1, 2, \dots, d_k$  and  $d_k$  is the dimension of the index. While we sometimes use the “upstairs” and “downstairs” notation for indices — basically in order to unclutter our expressions — this should not be interpreted as implying that the tensor transforms in any particular way under Lorentz transformations (whose action is not even defined here).

A key device we exploit in the following is to visualise an  $n$ -index tensor  $\psi$  as a vertex (really a circle) with  $n$  legs, one for each index  $j_k$ ,  $k = 1, 2, \dots, n$ :



According to such a visualisation tensors such as vectors are represented as a circle with one leg, matrices as a circle with two legs, and so on:



The contraction of the indices of two tensors is then represented by joining the legs of the corresponding indices. For example:

$$\sum_{kl} a_{jkl} b_{kl} \quad \text{---} \quad \text{Diagram of two tensors A and B with two legs each, connected by two lines representing the contraction of indices k and l.}$$

A *tensor network* is then the tensor  $\psi$  associated with a diagram, actually a *graph*  $G$ , consisting of a collection of tensors, represented as *vertices*, whose indices are contracted according to the *contraction pattern* described by the edges of the graph. The result is a tensor with a number of indices given by the external or “dangling legs” of the diagram.

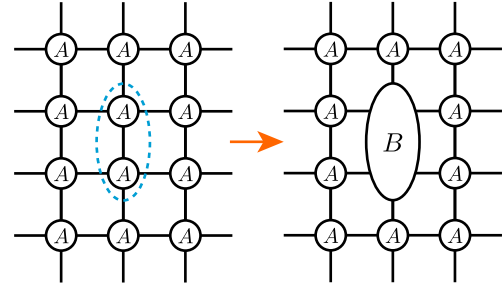
The canonical example of a tensor network is the *matrix product state* tensor:

$$A_{\alpha_1}^{j_1} A_{\alpha_1 \alpha_2}^{j_2} \cdots A_{\alpha_{n-1}}^{j_n} \equiv \text{Diagram of a chain of n tensors A connected by lines representing the contraction of indices } \alpha_i \text{ and } \alpha_{i+1}.$$

Calculating the value of a tensor network for a given configuration  $j_1 j_2 \cdots j_n$  of the external legs is called *contracting* a tensor network. The problem of contracting, or

even approximating the contraction of, an arbitrary tensor network is, in general, *hard*: even *planar* tensor networks can be arbitrarily difficult [24]. Tensor networks are, in general, useful only when they have some additional structure, such as short-ranged correlations (more on this later), topological simplicity, i.e., containing only a few loops, or causal structure, e.g., the tensors are isometries.

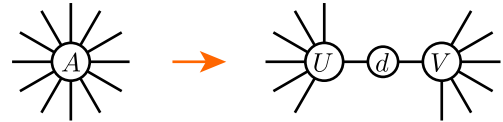
*Operations on tensor networks.*—There are a couple of basic operations one can perform on a tensor network to obtain (sometimes simpler) new networks. The first operation is *partial contraction*: here a collection of vertices and their corresponding tensors are identified and the indices corresponding to all the edges connecting these vertices are contracted. The original tensors are then replaced by the resulting amalgamation:



The second key primitive is that of *tensor splitting*: here we first collect the external legs of a tensor into two groups which are then regarded as two multi-indices. For example, suppose we have the  $n$ -index tensor  $A$  inside our network: we could collect the first  $k$  indices together as the multi-index  $\alpha \equiv [j_1 j_2 \cdots j_k]$  — running, say, in lexicographic order through the  $d_1 d_2 \cdots d_k$  assignments of the indices — and the remaining indices together as  $\beta \equiv [j_{k+1} \cdots j_n]$ . This bundling of indices gives us a 2-index tensor  $A_{\alpha\beta}$ , where  $\alpha$  can be taken to run from 1 to  $d_1 \cdots d_k$  and  $\beta$  from 1 to  $d_{k+1} \cdots d_n$ . We may now regard  $A$  as a *matrix*. The tensor splitting operation is then carried out by exploiting the *singular value decomposition* [25] to write  $A$  as a product of an isometry  $U$ , a diagonal matrix  $d$  with positive entries, and another isometry  $V$ :

$$A = U d V. \quad (1)$$

Graphically this is represented as:



Note that the original tensor is replaced with a contraction of three tensors and, in the process, the *degree* of the new vertices in the network is reduced. This procedure may be recursively applied to all the tensors in an arbitrary network to produce a new tensor network with maximal degree bounded above by 3. (Note: it is impossible, in general, to reduce the degree of all the vertices to 2 or below via the process.)

The third primitive is *truncation*. Focus on one edge  $e$  in a tensor network: this represents a contraction of two specific indices of two tensors. That is, we perform a sum over all the values of the index  $\alpha$  indicated by the edge  $e$  and shared by the two tensors at the ends of the edge:

$$\sum_{\alpha=1}^d \cdots A_{\dots\alpha} B_{\alpha\dots} \cdots \quad (2)$$

The truncation procedure is an approximation technique where we simply sum over fewer of the possible values of  $\alpha$ :

$$\sum_{\alpha=1}^d \cdots A_{\dots\alpha} B_{\alpha\dots} \cdots \mapsto \sum_{\alpha=1}^{d'} \cdots A_{\dots\alpha} B_{\alpha\dots} \cdots, \quad (3)$$

with  $d' < d$ . Truncation is particularly effective in combination with tensor splitting: when the diagonal matrix  $d$  in Eq. (1) has many small entries we can simply replace the sum over  $\beta$  with a sum over a smaller range, including only the indices corresponding to large entries.

A final approximation operation frequently useful in studying quantum states generated by local dynamics is the *Lie-Trotter* expansion:

$$e^{A+B} \approx (e^{A/m} e^{B/m})^m, \quad (4)$$

where  $n$  is a positive integer and  $A$  and  $B$  are operators on some hilbert space  $\mathcal{H}$ . This approximation improves as  $n$  is increased and becomes an identity in the limit  $m \rightarrow \infty$ . This result allows us to approximate the propagator  $U(t) = e^{-itH}$  — an  $n$ -index tensor — for a quantum spin system with a *local* hamiltonian  $H$  by a tensor network involving small-index tensors. To see how this works suppose that  $H$  is a hamiltonian for a chain of  $n$  quantum spins (similar results hold in higher dimensions):

$$H = \sum_{j=1}^{n-1} h_j, \quad (5)$$

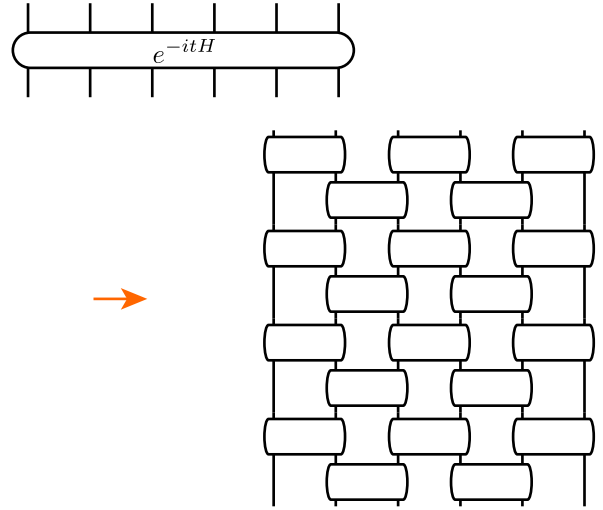
where  $h_j$  is a local interaction term which acts nontrivially only on spins  $j$  and  $j+1$ . Next collect together the even-numbered (respectively, odd numbered) interactions:

$$A = \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} h_{2k} \quad \text{and} \quad B = \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} h_{2k-1}. \quad (6)$$

Now  $A$  and  $B$  only contain terms which mutually commute with each other, so that  $e^A = \prod_k e^{h_{2k}}$  and  $e^B = \prod_k e^{h_{2k-1}}$ . Applying the Lie-Trotter expansion allows us to replace  $e^{zH}$  with the product

$$e^{zH} \approx \left( \prod_k e^{\frac{z}{m} h_{2k}} \prod_{k'} e^{\frac{z}{m} h_{2k-1}} \right)^m, \quad (7)$$

which is a tensor network of  $mn$  degree-4 tensors. Graphically this is represented by:



We've thus reduced the contraction of a single degree  $2n$  tensor to contracting a network of smaller tensors. In general the difficulty of these two different tasks is roughly equivalent, however, there are cases where the Lie-Trotter network can be exploited to develop useful approximation methods

*Tensor network states.*—Tensor network states (TNS) are quantum states of a many body quantum system associated with a tensor network. The systems we typically think about are quantum spin systems comprised of  $n$  distinguishable quantum spins with local hilbert spaces  $\mathcal{H}_j$ ,  $j = 1, 2, \dots, n$ ; the total hilbert space is then  $\mathcal{H} \cong \bigotimes_{j=1}^n \mathcal{H}_j$ . It is common to assume that  $\mathcal{H}_j$  is finite dimensional, with dimension  $d_j$ , although there is no fundamental reason preventing everything from straightforwardly applying to infinite dimensional systems. An arbitrary state  $|\psi\rangle$  of  $\mathcal{H}$  can be written as

$$|\psi\rangle = \sum_{j_1=1}^{d_1} \cdots \sum_{j_n=1}^{d_n} \psi_{j_1 j_2 \dots j_n} |j_1\rangle \cdots |j_n\rangle, \quad (8)$$

where  $|j_k\rangle$  is an arbitrary orthonormal basis for  $\mathcal{H}_j$ . The state  $|\psi\rangle$  is encoded in the  $d_1 d_2 \cdots d_n$  components of the  $n$ -index tensor  $\psi_{j_1 j_2 \dots j_n}$ . When the dangling legs of a tensor  $\psi$  are associated with an orthonormal basis of a quantum spin system we obtain a *tensor network state*.

*Basic properties of tensor network states.*—The tensor network formalism exposes some very powerful structural information about quantum states.

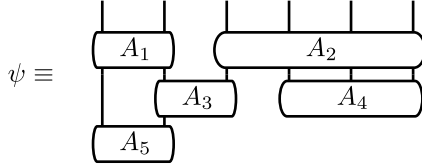
The first property is that a tensor network state is always manifestly a *state*. This is an extremely useful property not shared by other representations (e.g., a truncated perturbation series around some mixed state does not usually determine a positive state).

The second property is that expectation values of (local) operators may be computed by contracting a related tensor network: let  $|\psi\rangle$  be a TNS and  $M$  some (local) operator on the  $l$ th spin. Then the expectation value  $\langle M \rangle$  is determined

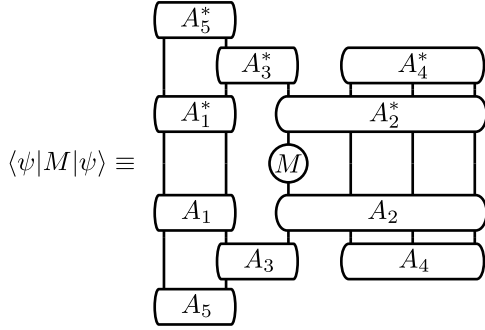
by

$$\sum_{j_1=1}^{d_1} \cdots \sum_{j_l, k_l=1}^{d_l} \cdots \sum_{j_n=1}^{d_n} \psi_{j_1 j_2 \dots j_n}^* \psi_{j_1 \dots j_{l-1} k_l j_{l+1} \dots j_n} M_{j_l k_l}, \quad (9)$$

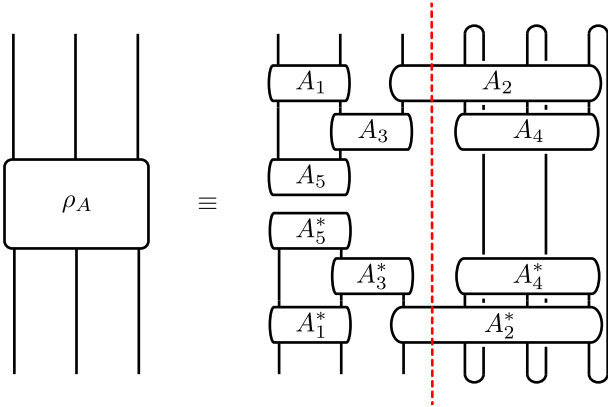
which is the contraction of three tensor networks: one for  $\psi^*$ , one for  $M$ , and one for  $\psi$ . For example, suppose we have a TNS whose coefficients are determined by the network



Then the expectation value of a hermitian operator  $M$  on the 3rd spin is given by the contraction



Closely related to the calculation of expectation values is the construction of the *reduced density operator*. Let  $|\psi\rangle$  be a tensor network state. Form the corresponding density operator  $\rho \equiv |\psi\rangle\langle\psi|$ : this is a tensor network with twice the number of external legs and every tensor doubled. Take the partial trace over all spins outside of region  $A$ ; this operation is found by connecting the indices corresponding to the legs in the complement in  $A$  as follows:



The third remarkable property is that we obtain an upper bound on the entropy of entanglement associated a region  $\mathcal{R}$  in the network. This result may be argued as follows.

*Tensor networks from hamiltonians.*—

*Tensor networks from path integrals.*—

Here we must cite Lifschitz theories; Ardonne; Horava

The basic input to any calculation in QFT is the Lagrangian  $\mathcal{L}$ . To keep a concrete example in mind think of  $\phi^4$  theory in  $D+1$  dimensions:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2 \phi^2 - \frac{\lambda}{4!} \phi^4, \quad (10)$$

however, everything we discuss here is valid in generality (in particular, everything we say here will apply equally to relativistic and nonrelativistic settings). Using the Lagrangian we can calculate any  $n$ -point correlation function using the path integral according to the standard formula

$$\langle T[\phi(x_1) \cdots \phi(x_n)] \rangle \equiv \lim_{T \rightarrow \infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi \phi(x_1) \cdots \phi(x_n) e^{i \int_{-T}^T d^4x \mathcal{L}}}{\int \mathcal{D}\phi e^{i \int_{-T}^T d^4x \mathcal{L}}}. \quad (11)$$

Before such a path integral can be evaluated — or even approximated — it is generally necessary to *regulate* the theory by applying a cutoff  $\Lambda$ . Now, in the context of tensor networks and quantum information, the most natural regulator is the lattice, with lattice spacing  $a = \frac{1}{\Lambda}$ , and we don't hesitate to use it throughout. In this way we discretise our continuous field(s)  $\phi(x)$  onto the integer lattice  $a\mathbb{Z}^{D+1}$ ; we restrict our coordinates to elements of the form  $x_\mu = a n_\mu$ ,  $n_\mu \in \mathbb{Z}^{D+1}$ . Write the field evaluated at lattice points as  $\phi_{n_\mu} = \phi(x_\mu)$ ; thus we discretise derivatives as  $\partial_\mu \phi(x_\nu) \mapsto (\phi_{n_\nu + e_\mu} - \phi_{n_\nu})/a$ , where  $e_\mu$  is the unit lattice vector in the direction  $\mu$ . Thus our potentially ill-defined path integral becomes the better-behaved iterated integral

$$\int \mathcal{D}\phi \mapsto \int [d\phi] \equiv \int \left( \prod_{n \in \mathbb{Z}^{D+1}} d\phi_n \right). \quad (12)$$

A further important device that we also initially exploit is to make a Wick rotation of time  $t \mapsto -i\beta$ : this turns our path integral expression for the correlation functions into a statistical mechanical problem:

$$\langle \phi(x_1^E) \cdots \phi(x_n^E) \rangle \equiv \frac{1}{\mathcal{Z}} \int [d\phi] \phi_{m_1} \cdots \phi_{m_n} e^{-S}, \quad (13)$$

where  $\mathcal{Z} = \int [d\phi] e^{-S}$  and, e.g.,  $S = a^{D+1} \sum_{\langle m, n \rangle \in \mathbb{Z}^{D+1}} (\phi_m - \phi_n)^2 / (2a^2) + \sum_n m^2 \phi_n^2 / 2 + \lambda \phi_n^4 / 4!$ .

To practitioners in tensor networks an expression such as Eq. (12) is strongly suggestive because we can introduce a special pure quantum state  $|\Phi\rangle$  which encodes all the information we can extract from a partition function by taking the square root of the probabilities [17, 18]. The way this works is as follows: if we have a partition function  $\mathcal{Z} = \sum_{j=1}^N e^{-E_j}$  for a system with  $N$  configurations we introduce the quantum state  $|\Phi\rangle \equiv \sum_{j=1}^N e^{-\frac{1}{2}E_j} |j\rangle / \sqrt{\mathcal{Z}}$ , where  $|j\rangle$  is an orthonormal basis corresponding to the  $N$  possible configurations of the classical system.

Remarkably, we can also realise  $|\Phi\rangle$  as the *ground state* of a natural hamiltonian [13, 14, 26, 27]. The key to this construction is to introduce a reversible markov chain  $M$  which has  $p_j \equiv e^{-E_j}/\mathcal{Z}$  as its stationary distribution [28], i.e.,  $\sum_{j=1}^N p_j M_{jk} = p_k, \forall k$ . Reversibility is the condition that  $p_j M_{jk} = p_k M_{kj}, \forall j, k$ . This motivates us to introduce the matrix with matrix elements:

$$K_{jk} \equiv \sqrt{p_j} M_{jk} \frac{1}{\sqrt{p_k}}. \quad (14)$$

Because of the reversibility condition we find that  $K$  is real and symmetric, i.e., hermitian. Further, we have that operator  $\hat{K} \equiv \sum_{j,k=1}^N K_{jk} |k\rangle\langle j|$  has  $|\Phi\rangle$  as an eigenvector corresponding to its largest eigenvalue. We obtain our desired hamiltonian by writing  $\hat{H} \equiv \mathbb{I} - \hat{K}$ , this operator has  $|\Phi\rangle$  as a ground state with eigenvalue 0.

Given a probability distribution of exponential form, i.e.,  $p_j \equiv e^{-E_j}/\mathcal{Z}$ , there is a special reversible Markov chain having  $p_j$  as its stationary distribution, namely, the Metropolis algorithm [29]. To define this Markov chain we introduce the quantity

$$c_{jk} \equiv \min\{1, e^{-E_j+E_k}\}. \quad (15)$$

Let  $F$  be an operator which flips classical configurations, a *flip operator*. In our cases it is convenient to let  $F$  be a permutation of  $j = 1, 2, \dots, N$ . We then use the flip operator to construct

$$\hat{M} = \sum_{j \neq k} c_{jk} |j\rangle\langle k| \quad (16)$$

When we apply the procedure described here to our Wick rotated path integral we obtain the following state of a lattice of a  $(D+1)$ -dimensional spatial lattice of harmonic oscillators

$$|\Phi\rangle \equiv \frac{1}{\mathcal{Z}} \int [d\phi] e^{-\frac{\mathcal{S}}{2}} |\phi\rangle, \quad (17)$$

where

$$|\phi\rangle \quad (18)$$

*Reflection positivity and analytic continuation.—*

*Conclusions.—The*

- 
- [1] K. G. Wilson, Rev. Modern Phys. **47**, 773 (1975).
  - [2] K. G. Wilson and J. B. Kogut, Phys. Rept. **12**, 75 (1974).
  - [3] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Inf. Comput. **7**, 401 (2007), quant-ph/0608197.
  - [4] R. Orus, “A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States,” (2013), arXiv:1306.2164.
  - [5] G. Evenbly and G. Vidal, J. Stat. Phys. **145**, 891 (2011).
  - [6] G. Evenbly and G. Vidal, Phys. Rev. B **79**, 144108 (2009).
  - [7] F. Verstraete and J. I. Cirac, “Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions,” (2004), cond-mat/0407066.
  - [8] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007), cond-mat/0512165.
  - [9] M. Levin and C. P. Nave, Phys. Rev. Lett. **99**, 120601 (2007).
  - [10] Z.-C. Gu and X.-G. Wen, Phys. Rev. B **80**, 155131 (2009).
  - [11] K. G. Wilson, Phys. Rev. D **10**, 2445 (1974).
  - [12] S. Dürr, Z. Fodor, J. Frison, C. Hoelbling, R. Hoffmann, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, K. K. Szabo, and G. Vulvert, Science **322**, 1224 (2008).
  - [13] P. Hořava, Phys. Lett. B **694**, 172 (2010).
  - [14] E. Ardonne, P. Fendley, and E. Fradkin, Annals of Physics **310**, 493 (2004).
  - [15] P. Hořava, Phys. Rev. D **79**, 084008 (2009).
  - [16] B. Swingle, Phys. Rev. D **86**, 065007 (2012).
  - [17] D. Aharonov and A. Ta-Shma, in *Proceedings of the 35th Annual ACM Symposium on Theory of Computing held in San Diego, CA, June 9–11, 2003* (Association for Computing Machinery (ACM), New York, 2003) pp. 20–29, quant-ph/0301023.
  - [18] L. Grover and T. Rudolph, “Creating superpositions that correspond to efficiently integrable probability distributions,” (2002), quant-ph/0208112.
  - [19] S. Ryu and T. Takayanagi, Phys. Rev. Lett. **96**, 181602 (2006).
  - [20] S. Ryu and T. Takayanagi, JHEP **2006**, 045 (2006).
  - [21] T. Nishioka, S. Ryu, and T. Takayanagi, J. Phys. A **42**, 504008 (2009).
  - [22] J. Maldacena and L. Susskind, Fortschr. Phys. **61**, 781 (2013).
  - [23] T. Hartman and J. Maldacena, JHEP **2013**, 014 (2013).
  - [24] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **98**, 140506 (2007).
  - [25] R. A. Horn and C. R. Johnson, *Matrix analysis* (Cambridge University Press, Cambridge, 1990) pp. xiv+561.
  - [26] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. **96**, 220601 (2006).
  - [27] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Inf. Comput. **8**, 650 (2008).
  - [28] J. R. Norris, *Markov Chains* (Cambridge University Press, Cambridge, 1997).
  - [29] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).