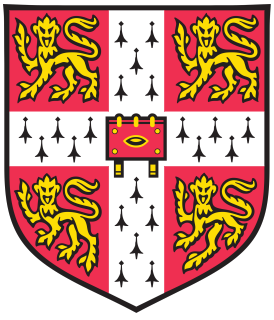
Cambridge University
Department of Applied Mathematics and Theoretical Physics
Part III - Master of Advanced Study

Alexandre Krajenbrink

Essay on Tensor Network Renormalization
Brief review of renormalization, applications and outlook.

Essay proposed by Dr. Wingate



Abstract

Many-body phenomena carry a great interest of research as the range of application is enormous : from superconductivity that has the possibility to provoke a revolution in medicine, to Bose-Einstein condensation that enables us to manipulate qubits to design potential quantum computers. Yet, some models, either classical or quantum are still not solved nowadays. There is therefore a need to develop tools, methods, approximations to capture the properties of these models, such as phase transitions. One path that is thought to lead to a good understanding of numerous models is renormalization. This path has been developed first by physicists like Wilson [1] and Kadanoff [2], it has evolved to propose several schemes used in different contexts, and recently, it has found applications in other fields such as probability theory [3] and deep learning [4].

This essay deals about a very new research subject in the field of renormalization : tensor network renormalization. While in the first part, we aim at introducing a general framework to study statistical physics and renormalization, we then briefly review the ideas of real space renormalization, recalling that renormalization is a fruitful tool for plenty of topics such as probability theory. We then enter the core of this essay, we define a universality class which is tensor network models, derive several of the properties of this class and exhibit different examples. Finally, we introduce renormalization schemes for tensor network models. We will carefully emphasize on the mathematical foundations of tensor network models as this will help us to understand intrinsic links between tensor network renormalization and quantum information.

Contents

1 Problems in classical statistical physics

1.1 Definitions

This section has the purpose of giving a framework to study statistical physics problems. In this essay, we will restrict ourselves to classical problems.

Definition 1.1 (Problem in classical statistical physics)**.** Let $\{\sigma_{\vec{\mathbf{r}}}\}$ be an ensemble of stochastic variables taking values in Ω - it can be a discrete or a continuous space - and represented by a position $\vec{r}$ in $\mathbb{R}^n$. A problem is represented by a Hamiltonian $\mathcal{H}\{\sigma_{\vec{\mathbf{r}}}\}$ and we introduce a topology by defining a measure[1], $\exp(-\beta\mathcal{H}\{\sigma_{\vec{\mathbf{r}}}\})$, where β is an open parameter usually equal to the inverse of the thermodynamic temperature of the system[2]. We define the free energy $\mathcal{F}$ and the partition function $\mathcal{Z}$ as follows :

$$e^{-\beta\mathcal{F}} = \mathcal{Z} = \mathrm{Tr}_{\{\sigma_{\vec{\mathbf{r}}}\}}(e^{-\beta\mathcal{H}\{\sigma_{\vec{\mathbf{r}}}\}}),$$

where the trace is a linear operator that allows us to sum over all different configurations.

Definition 1.2 (Hamiltonian in classical statistical physics)**.** Without any loss of generality, we can define operators $\Theta_i(\{\sigma_{\vec{\mathbf{r}}}\})$ and coupling constants u_i such that the Hamiltonian is written in the following form :

$$\mathcal{H}(u,\Theta) = \sum_i u_i \Theta_i(\{\sigma_{\vec{\mathbf{r}}}\})$$

These operators represents different interactions, for instance nearest neighbor, next to nearest neighbor, three-body interaction, self-interaction.

In the past few years, problems were redefined to obey a tensor network representation. The object of this essay is to understand the construction of this representation and its applications.

1.2 Temperature as a parameter

It is often seen in the literature that temperature in a statistical physics problem is the temperature of an external thermal bath. Yet, the paradigm is to manipulate temperature as a parameter in our system. We shall give in this section an interpretation of the temperature as a parameter and not as a variable of the environment.

We shall consider a stochastic system which has N different microstates, with energy $\mathcal{E}_i$ and weight p_i. The system has an average energy equal to $\mathcal{E}$.

Let $\mathcal{S}$ be the Shanon-Gibbs entropy, namely :

$$\mathcal{S} = -k_B \sum_{i=1}^{N} p_i \ln p_i,$$

[1] As we will see later, this measure is a consequence of the energy conservation and the principle of maximum entropy.

[2] In this definition, there is no need to introduce a thermal bath to define the temperature, we will come back to this later.

where k_B is Boltzmann's constant. The constraints on our system are the following :

$$\begin{cases} \sum_{i=1}^{N} p_i = 1 \quad \text{(normalization of the weights of the system)} \\ \sum_{i=1}^{N} p_i \, \mathcal{E}_i = \mathcal{E} \quad \text{(conservation of the energy)} \end{cases}$$

In order to define equilibrium in our system, we have to maximize the entropy under the constraints. Using Lagrange multipliers λ and μ, we can write the optimization problem as follows :

$$\max_{p_1,\dots,p_N \in [0,1]} \left\{ -k_B \sum_{i=1}^{N} p_i \ln p_i + \lambda [\sum_{i=1}^{N} p_i - 1] + \mu [\sum_{i=1}^{N} p_i \, \mathcal{E}_i - \mathcal{E}] \right\} \tag{1}$$

Since the optimization problem is the sum of N independent optimization problems, we just have to optimize for one p_i. Let f be defined by :

$$f(p) = -k_B p \ln p + \lambda p + \mu p \mathcal{E}$$

The extremum of f obtained for $p = e^{-1+(\lambda+\mu\mathcal{E})/k_B}$, and therefore the solution of equation 1 is :

$$\forall i \in [1, N], \quad p_i = e^{-1+(\lambda+\mu\mathcal{E}_i)/k_B}$$

Hence, the Maxwell weight is the distribution with the greatest entropy for our constraints. Let us look at the constraints to determine the Lagrange multipliers.

$$\sum_{i=1}^{N} p_i = e^{-1+\lambda/k_B} \sum_{i=1}^{N} e^{\mu\mathcal{E}_i/k_B} = 1$$

We deduce that :

$$\forall i \in [1, N], \quad p_i = \frac{e^{\mu\mathcal{E}_i/k_B}}{\sum_{i=1}^{N} e^{\mu\mathcal{E}_i/k_B}}$$

Let $\mathcal{Z}$ be the partition function of our system. Its purpose it to normalize the weights of our system.

$$\mathcal{Z} = \sum_{i=1}^{N} e^{\mu\mathcal{E}_i/k_B}$$

and

$$\lambda = k_B(1 - \ln \mathcal{Z})$$

The other constraint yields :

$$\mathcal{S} = -k_B \sum_{i=1}^{N} p_i \ln p_i = -\sum_{i=1}^{N} p_i \mu \mathcal{E}_i + k_B \ln \mathcal{Z} = k_B \ln \mathcal{Z} - \mu\mathcal{E} \tag{2}$$

The interpretation of equation 2 is that the first term refers to Boltzmann entropy formula and the second term corresponds to a correction to Boltzmann's

formula due to the constraint of energy conservation[3]. Recalling that the thermodynamic temperature is defined by :

$$\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}E} = \frac{1}{T}$$

We deduce that :

$$\mu = -\frac{1}{T}$$

Rewriting our initial optimization problem, replacing Lagrange multipliers and subtracting all constants, we get :

$$\max_{p_1,\dots,p_N \in [0,1]} \left\{ -k_B \sum_{i=1}^{N} p_i \ln p_i + k_B(1 - \ln \mathcal{Z}) \sum_{i=1}^{N} p_i - \frac{1}{T} \sum_{i=1}^{N} p_i \, \mathcal{E}_i \right\}$$

We can now consider the temperature as a parameter that can vary, and we can look at regimes where $T \to +\infty$ and $T \to 0$. In the regime $T \to +\infty$, the probability distribution if uniform, and the disorder is maximum. In the regime $T \to 0$ we have to minimize the quantity $\sum_{i=1}^{N} p_i \, \mathcal{E}_i$, this quantity if lower bounded by the minimum energy $\mathcal{E}_{min}$, the solution is therefore getting all weights equal to zero except p_{min} equal to 1.

We have seen in this section that temperature can be used as a parameter in the sense that it can be seen as a Lagrange multiplier. The behavior of this "Lagrange temperature" has the same interpretation as the usual kinetic temperature and furthermore we do not need to introduce any thermal bath to manipulate it.

[3]If μ is equal to zero, all the states have a uniform probability and equation 2 is exactly the Boltzmann entropy.

2 Real space renormalization group schemes

2.1 What is renormalization ?

All the systems that present the property of phase transition at a critical point - such as water boiling at ∼100 Celcius at standard pressure - fall into a universality class. This universality class[4] described in [6] can be characterized by several features :

Scale invariance : A system set at a critical point of a phase transition will present correlations at all length scales. No matter if we zoom in or out, we expect the behavior to remain the same.

Scale covariance : Near a phase transition, many physical phenomena are fully described by critical exponents. These exponents relate to how the system will react as we come closer of farther from the critical point.

Fixed point : As a result of the scale invariance, the Hamiltonian that describes the system set at the critical point should be a function that does not depend on any system of unit. Therefore the Hamiltonian should be a fixed point of any scale transformation.

The first intuition we can have lies in the fact that a large scale physical phenomenon should not depend on every small detail at a small scale, so that universality is something macroscopic and not microscopic. Another way to understand this is to view our physical variables - spins for instance - as stochastic variables. Statistically, by summing lots of interactions on stochastic variables, we might expect a "concentration of measure" such as the one we have with the Central Limit Theorem. The Central Limit Theorem can also be interpreted in a way useful to introduce renormalization. We know that two systems with different probability laws on the stochastic variables can lead to the same Gaussian measure at the end, this Gaussian measure erased the local degrees of freedom of the initial variables so as to reveal a macroscopic degree of freedom.

Renormalization is a framework that allows us to study the behavior of systems that belong to the universality class. It enables us to reduce recursively the degrees of freedom of a system to reveal critical properties, namely the ones that arise at a phase transition. This recursive procedure will define a flow that will be characteristic of the renormalization. Critical properties will be found at the basin of attraction of this flow.

Several techniques of renormalization exists, like momentum space renormalization, real space renormalization, or tensor network renormalization. We will briefly review what is real space renormalization, provide original examples and then we will enter the core of this essay : tensor network renormalization.

[4] As we will see in this essay, this universality class goes way beyond statistical physics as it is present in probability theory and even in economy and sociology, an example is provided in [5].

2.2 Real space renormalization

We will review in this section one strategy of real space renormalization called blocking. The idea is to group our stochastic variables, spins, into blocks. If for simplicity we consider a d-dimensional regular lattice with a lattice spacing a, we now create blocks of s^d spins so that the new spacing between two blocks is sa for a parameter $s > 1$. At this point we should relate the new variables of the blocked lattice $\{\sigma'_{\vec{\mathbf{r}}}\}$ to the old ones $\{\sigma_{\vec{\mathbf{r}}}\}$. From this point we will omit the space label $\vec{\mathbf{r}}$ to have lighter notations. A way to formalize that is to introduce the blocking kernel $\mathrm{T}(\sigma', \sigma)$ such that the probability measure is transformed as :

$$\exp(-\beta\mathcal{H}'\{\sigma'\}) = \sum_{\sigma} \mathrm{T}(\sigma', \sigma) \exp(-\beta\mathcal{H}\{\sigma\}) \tag{3}$$

A way to interpret equation (3) is to say that the blocking kernel just rearranges the measure inside the system by marginalizing the variables inside each block. Depending on the problem we want to solve, T might take different forms. For simplicity, let us take an Ising-like model where $\{\sigma\}$'s take only a discrete set of value.

- If we want to keep this structure, so we can opt for "the majority rule" where the new variable takes the value taken by the majority of the block, or for "decimation" where we just integrate some variables out of the block.

- If in the contrary we do not care about keeping the structure of σ's values, we can choose σ' to be a linear combination of σ.

In any case, the choice of the blocking kernel is purely arbitrary. The only constraint we will impose on T is to uniquely define the blocks, namely[5] :

$$\forall\sigma, \quad \sum_{\sigma'} \mathrm{T}(\sigma', \sigma) = 1 \tag{4}$$

Equations (3) and (4) suggest that the partition function of the system remains the same :

$$\begin{aligned}
\mathcal{Z}' &= \sum_{\sigma'} \exp(-\beta\mathcal{H}'\{\sigma'\}) \\
&= \sum_{\sigma} \sum_{\sigma'} \mathrm{T}(\sigma', \sigma) \exp(-\beta\mathcal{H}\{\sigma\}) \\
&= \sum_{\sigma} \exp(-\beta\mathcal{H}\{\sigma\}) \\
&= \mathcal{Z}
\end{aligned}$$

This is something good since as all macroscopic properties can be deduced from the partition function and we only want to remove some microscopic details.

[5] In Markov chain theory, T can be interpreted as a transfer matrix from the state σ to the state σ', T(y,x) is the probability that the state x is transformed into the state y, and equation (4) simply means that the state x is transformed into another state, or in terms of our variables, that the variable σ is sent to some block.

Equation (3) defines a flow of measure, or equivalently a flow of Hamiltonian. Following a flow of measure or Hamiltonian is not something very intuitive, therefore generally one follows the flow of coupling constants.

$$\mathcal{H}(u,\Theta)=\sum_i u_i\Theta_i(\{\sigma\})\longmapsto\mathcal{H}'(u',\Theta)=\sum_i u'_i\Theta_i(\{\sigma'\})$$

Calling $\underline{u}$ the vector of all coupling constants, the flow can be written :

$$\underline{u}'=\mathcal{R}_s(\underline{u})$$

At this point, it is important to recall that the Hamiltonians are defined in lattices with different spacings a and $a'=sa$. So as to compare the properties of these Hamiltonians (such as correlation lengths) we must define distances relative to $\mathcal{H}'$, $\vec{\mathbf{r}'}$ in the way that $\vec{\mathbf{r}'}=\frac{\vec{\mathbf{r}}}{s}$ so to have $a'\vec{\mathbf{r}'}=a\vec{\mathbf{r}}$. There is one last degree of freedom that we did not modify during our transformation : the scaling dimension of $\{\sigma'\}$. To get some quantities adimensional, we can proceed to the following transformation $\sigma'\longmapsto\lambda(s)\sigma'$ and we will call λ the renormalization factor.

At this point, let us look at some mathematical properties of the flow of coupling constants $\mathcal{R}_s$.

Property 2.1 (Semi-group property)**.** *If we proceed to the renormalization scheme with a factor s_1 and then a factor s_2, it should physically provide the same answer as we did it with a factor s_2 then s_1 or directly s_1s_2. It provides a relation called a semi-group property :*

$$\mathcal{R}_{s_1}\mathcal{R}_{s_2}=\mathcal{R}_{s_1s_2}$$

We have the same semi-group property for the renormalization factor λ. The term semi-group is actually justified, as our transformations are not invertible. Each time we look at larger scales, we loose some information about the lower scales just like the Central Limit Theorem erases almost all information about the probability law of each random variable.

Property 2.2 (Characterization of critical points)**.** *We characterize a critical point as a fixed point of the renormalization scheme, which is written as :*

$$\underline{u}^*=\mathcal{R}_s(\underline{u}^*)$$

Property 2.3 (Linearization near fixed point)**.** *Let $\delta\underline{u}$ be a small perturbation around the fixed point, so that $\underline{u}=\underline{u}^*+\delta\underline{u}$. Linearizing the flow of coupling constants we find :*

$$\delta\underline{u}'=\underline{\underline{\nabla}}\mathcal{R}\,.\,\delta\underline{u}+\mathcal{O}(\delta\underline{u}^2)$$

Suppose we can diagonalize $\underline{\underline{\nabla}}\mathcal{R}$ - which is not trivial at all -, let $\underline{e}_i(s)$ be its eigenvectors, the semi-group property imposes the eigenvalues of this matrix to be of the form s^{α_i} so that $\underline{\underline{\nabla}}\mathcal{R}\,.\,\underline{e}_i(s)=s^{\alpha_i}\underline{e}_i(s)$. Writing $\delta\underline{u}'=\sum_i\psi'_i\underline{e}_i(s)$ and $\delta\underline{u}=\sum_i\psi_i\underline{e}_i(s)$, we get $\psi'_i=s^{\alpha_i}\psi_i$.

Coming back to our Hamiltonian $\mathcal{H}(u,\Theta)$, near the critical point, its expression in a compact form is $\mathcal{H}(u,\Theta) = \mathcal{H}(u^*,\Theta) + \delta\underline{u}.\underline{\Theta}$. If we choose the basis of eigenvectors, we get $\mathcal{H}(u,\Theta) = \mathcal{H}(u^*,\Theta) + \sum_i \psi_i\Theta_i$. This tells us that there exists three categories of interactions.

- $\alpha_i > 0$ then the intensity of the interaction Θ_i grows at each renormalization iteration.
- $\alpha_i < 0$ then the intensity of the interaction Θ_i decays at each renormalization iteration.
- $\alpha_i = 0$ then the intensity of the interaction Θ_i remains the same at the first order.

We therefore see that only the first category will play a role at the macroscopic scale. The general goal is to find the fixed points, and then the relevant interactions and their related exponents α_i.

Following this section, we give two examples of real space renormalization that lie outside the field of statistical physics showing how universal this framework is and also showing how we can make connections between different fields such as probability theory, random walks and statistical physics.

2.3 Original proof of the central limit theorem

We shall see in this section an original proof of the Central Limit Theorem using a blocking renormalization method. The proof is inspired from ideas from [7] and the whole calculation is derived here.

Let us consider N random variables $(\sigma_i)_{i\in[1,N]}$ independent and identically distributed of moments $(m_i)_{i\in\mathbb{N}}$ [6] and the following Hamiltonian :

$$\mathcal{H} = \sum_{i=1}^{N} \sigma_i$$

We know from the Central Limit Theorem that $\mathcal{H}$ will have a typical $\sqrt{N}$ behavior for N large. We will now demonstrate that by renormalization.

Let us group our N variables into $\frac{N}{s}$ blocks of s variables and let us define new variables[7] :

$$\forall\alpha \in [1, \frac{N}{s}], \quad \sigma'_\alpha = \frac{\lambda(s)}{s}\sum_{i\in\alpha}\sigma_i$$

where $\lambda(s)$ is a renormalization factor which role will be explained below. We may describe the distribution of these new variables in terms of their moments

[6] We may suppose for simplicity that the distribution of the variables is even.

[7] The new variables correspond to the mean of variables inside a block rescaled with the renormalization factor that should accordingly to the semi-group property be of the form s^{d_σ}.

$m'_2(s)$, $m'_4(s)$...[8] The moments of σ' can be expressed in terms of σ's moments.

$$m'_2(s) =< (\sigma')^2 >= \frac{\lambda^2}{s^2} \sum_{i,j \in \alpha} < \sigma_i \sigma_j >= \frac{\lambda^2}{s^2} s < \sigma^2 >= \frac{\lambda^2}{s} m_2$$

$$m'_4(s) =< (\sigma')^4 >= 3 < (\sigma')^2 >^2 + \frac{\lambda^4}{s^3} m_4 = 3\frac{\lambda^4}{s^2} m_2^2 + \frac{\lambda^4}{s^3} m_4$$

Getting the general moment analytically requires much work that is linked to the notion of cumulants in probability, but we do not need all of that for our purpose.

$$m'_{2n}(s) = \frac{\lambda^{2n}}{s^{2n}} \sum_{i_1 ... i_{2n} \in \alpha} < \sigma_{i_1} ... \sigma_{i_{2n}} >$$

Since any odd moment is equal to zero, we have to group σ's into groups of even size. Let us look at the extreme ways to group them. The first extreme way is to form one group of size 2n, namely 2n replicas of the same variable, we have s possibilities for that as each group contains s variables, this gives a contribution to $m_{2n}(s)$ equal to :

$$m'^{size(2n)}_{2n}(s) = \frac{\lambda^{2n}}{s^{2n}} \, s \, m_{2n} \tag{5}$$

The other extreme way consists on forming n groups of size 2. We have to find the number of ways of joining 2n objects into pairs. Given 2n objects, we can pair the first one with $2n-1$ any other. We can pair the next one to any of the $2n-3$ remaining. Therefore the number of ways of joining 2n objects into pairs[9] is :

$$(2n-1)(2n-3)...(1) = \frac{(2n)!}{n!\, 2^n}$$

For each pair of random variables we have to sum the indices to s different values, therefore this way of forming n groups of size 2 gives a contribution equal to :

$$m'^{size(2)}_{2n}(s) = \frac{\lambda^{2n}}{s^{2n}} \frac{(2n)!}{n!\, 2^n} s^n \, m_2^n \tag{6}$$

Now that we have the extreme contributions, lets look at the relative power of λ and s, in equation (5) we have λ^{2n}/s^{2n-1} and in equation (6) λ^{2n}/s^n. Because we have s variables inside any group, if we form a group of size 2k, namely 2k replicas of the same variable, we have the identity :

$$\forall k \in \mathbb{N}, \quad \sum_{i \in \alpha} < \sigma_i^{2k} >= s \, m_{2k}$$

This identity tells us if we form q groups so that the sizes of the groups sums to 2n, we will get a contribution :

$$m'^{q\ groups}_{2n}(s) = a_q \frac{\lambda^{2n}}{s^{2n-q}},$$

[8] Because of the assumption made on the distribution, all odd moments are trivially equal to zero.

[9] Something interesting is that this little graph combinatorics calculation plays a important role in Quantum Field Theory as it is used to derive Feynman rules in a QFT in 0 dimension with a ϕ^4 interaction `http://www.damtp.cam.ac.uk/user/dbs26/AQFT/chaps1+2.pdf`

where a_q plays the role of a symmetry factor, includes moments of different orders and is independent of s. We have written explicitly a_1 in equation (5) and a_n in equation (6). This gives us a general expression for the moment of the coarse grained variables :

$$m'_{2n}(s) = \frac{\lambda^{2n}}{s^n} \frac{(2n)!}{n!\, 2^n} m_2^n + \sum_{q=1}^{n-1} \frac{\lambda^{2n}}{s^{2n-q}} a_q \tag{7}$$

Let us now define a new Hamiltonian :

$$\mathcal{H}'_s = \sum_{\alpha=1}^{N/s} \sigma'_\alpha$$

The problem remains the same as before, we still have to characterize the distribution of $\mathcal{H}'$. The moments of the variables changed during our coarse graining procedure, and we may continue the procedure recursively, which will define a flow in the space of moments. The procedure might stop when we have only one random variable, this happens when $s \to N \to +\infty$.

Looking at equation (7), we see that there is a unique choice for $\lambda(s)$ that leads to a non-trivial fixed point [10] :

$$\lambda(s) = \sqrt{s}$$

For that value, the general moment is equal to :

$$m'_{2n}(s) = \frac{(2n)!}{n!\, 2^n} m_2^n + \sum_{q=1}^{n-1} \frac{1}{s^{n-q}} a_q$$

We can see that m_2 stays unchanged by this renormalization and all moments converge to the moments of a normal distribution as $s \to N \to +\infty$.
As $\mathcal{H} = \sqrt{s}\mathcal{H}'_s$ and $\mathcal{H}'_\infty = \sigma_\infty$ for $s \to N \to +\infty$, we deduce that in the limit $N \to +\infty$, $\mathcal{H} \sim \sqrt{N}\sigma_\infty$ where σ_∞ is a random variable normally distributed.

We have therefore proved the Central Limit Theorem by looking at a fixed point of a renormalization flow for the moments of the variables. Following this approach, what could be interesting is to do the calculations again for correlated variables where for instance the correlation function would describe a Ornstein-Uhlenbeck process :

$$< \sigma_i \sigma_j >= m_2 \, \exp(-\frac{\mid i - j \mid}{N_c}),$$

where $2N_c$ can be seen as the number of variables strongly correlated to any given one.

Another interesting application would be to apply the methods just described to a Hamiltonian (let us say for simplicity an Ising Hamiltonian) with a random external field. We would therefore need to follow the flow of the moments of the spins and the flow of the coupling constants.

[10] Trivial fixed points are either all moments being equal to zero or at least one moment equal to $+\infty$.

2.4 Renormalization and self avoiding walks

Another application of real space renormalization lies in studying an object in probability called self avoiding walk. A self avoiding walk is a random walk that cannot intersect itself, it was first introduced to model the geometry of polymers. The interest of that object are numerous. It was proved by Gennes [8] that self avoiding walks were describing several properties of the O(n) model in the limit $n \to 0$. Besides, the continuum limit of a self avoiding walk is not perfectly known and is currently being investigated in the framework of Schramm-Loewner theory.

To understand why this object is much more complicated than a simple random walks that converges to a Brownian motion let us consider a square lattice in two dimensions. After N steps, there are 4^N different random walks possible since each site has four neighbors. Yet, as we constrain the self avoiding walks not to intersect themselves, their number after N steps is smaller than 4^N. It is proven in [3] that the number of self avoiding walks after N steps, in the limit N large is :

$$\mathcal{N}(\mathrm{N}) \sim \bar{z}^{\mathrm{N}} \mathrm{N}^{\gamma-1}, \tag{8}$$

where $\bar{z}$ has the physical interpretation of the effective number of neighbors on the lattice. $\bar{z}$ depends only on the lattice whereas γ is a universal exponent that depends only on the dimension of the lattice.

We are now going to develop an example in which we can find the $\bar{z}$ of a lattice thanks to a decimation real space renormalization method. The proof is inspired from [7]. Let us consider a particular fractal lattice : Koch snowflake. The construction of this lattice is described in figure 1.

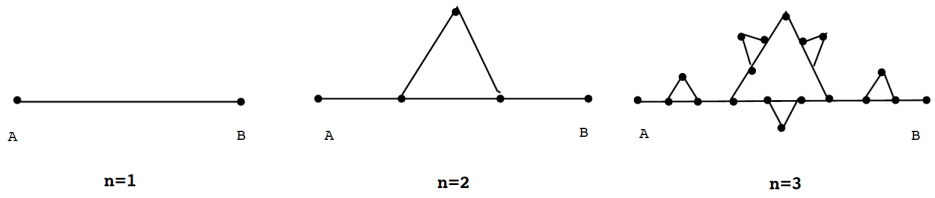


Figure 1: Construction of Koch snowflake, taken from [7]

Let us then define the generating function of the self avoiding walks joining the extreme points A and B of figure 1 at step n of the recursive construction :

$$G_n(\beta) = \sum_N \beta^N \mathcal{N}_N^{(n)},$$

where $\mathcal{N}_N^{(n)}$ is the number of self avoiding walks going from A to B in N steps on the lattice obtained after n iterations. Thanks the result of equation (8), the asymptotic term of this series is roughly $\beta^N \bar{z}^N$ so there might be a critical value of β above which the generating function always diverges.

Before entering into the details of renormalization, let us look at a recursive relation between G_n and G_{n+1}. If we consider a walk of N steps contributing to G_n, it gives birth to one path of length 3N, $\binom{N}{1}$ paths of length 3N+1 and more generally to $\binom{N}{k}$ paths of length 3N+k for $k \in [0, N]$. Therefore $\mathcal{N}_{3N}^{(n+1)} = \mathcal{N}_N^{(n)}$ and more generally $\mathcal{N}_{3N+k}^{(n+1)} = \binom{N}{k}\mathcal{N}_N^{(n)}$. Plugging this recursive relation into the generating function we find :

$$G_{n+1}(\beta) = G_n(\beta^3 + \beta^4) \tag{9}$$

Let us now define our renormalization procedure so to find the critical value of β. We decimate each elementary cell to a single line as in figure 2.

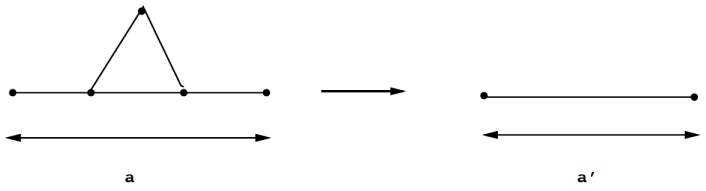


Figure 2: Renormalization procedure, taken from [7]

It defines a flow in the space of β's :

$$\beta' = \beta^3 + \beta^4$$

This renormalization transformation has three real fixed points :

$$\begin{cases} \beta_1 = 0 & \text{(this the infinite temperature fixed point)} \\ \beta_2 = +\infty & \text{(this the zero temperature fixed point)} \\ \beta_3 = \frac{1}{6}(-2 + 2^{\frac{2}{3}}(25 - 3\sqrt{69})^{\frac{1}{3}} + 2^{\frac{2}{3}}(25 + 3\sqrt{69})^{\frac{1}{3}}) \simeq 0.755 \end{cases}$$

Only β_3 is a non trivial fixed point. It is the critical point of the system. What is more, the physical intuition tells us that it is a repulsive fixed point. And as for all n in $\mathbb{N}$, $G_n(+\infty) = +\infty$ we deduce that β_3 is the radius of convergence of the generating function and therefore that $\bar{z} = \frac{1}{\beta_3} \simeq 1.32$.

We have a limited knowledge of $\bar{z}$ for different lattices, for instance only the connective constant of the hexagonal lattice has been rigorously calculated, the proof was stated in [9]. We may therefore think to extend renormalization methods to different lattices to extend the actual state of art.

3 Tensor network states

Several authors discovered in the past few years a new class of universality among classical lattice models with local interactions : all these models can be written as tensor network models. A first vague definition of a tensor network is the following :

Definition 3.1 (Tensor network model)**.** A tensor network model is a high dimensional generalization of the one dimensional matrix product model.

We will first see in paragraph 3.1 the proof of existence of the universality class of tensor network models. We will then see in paragraph 3.2 several examples and answer the question of how we can derive a matrix product model from the 1D Ising model with a magnetic field which is the simplest example of tensor network. Finally we will address in paragraph 3.3 several questions on the notion of lattices and dual lattices that arise during the construction of tensor networks.

3.1 Existence of tensor network states

The results of this section aim at motivating the study of tensor network states by proving their existence within the framework of local interactions. We will state that any classical statistical model in any spatial dimension on a lattice with nearest neighbor interactions can be expressed as a tensor-network model. A sketch of a proof was initiated in [10] and we complete it here with additional remarks.

Theorem 3.2 (Existence of a tensor network state)**.** *Let us consider a model with nearest neighbor interactions on a lattice $\mathcal{L}$ with $\tilde{N}$ stochastic variables taking values in Ω which cardinal is $\mid \Omega \mid$. There are N variables inside the lattice (not at the boundary), each of them has D neighbors. Let $\mathcal{Z}$ be the partition function of the model.*

$$\exists (T^{(1)},...,T^{(N)}) \in (\mathbb{R}^{|\Omega|})^{\otimes D},\ \phi : [(\mathbb{R}^{|\Omega|})^{\otimes D}]^N \to \mathbb{R}^+ \text{ such that } \mathcal{Z} = \phi(T^{(1)},...,T^{(N)})$$

Proof. Let $(\sigma)_{i\in[1,\tilde{N}]}$ be the stochastic variables of the model. Under the assumptions of Theorem 3.2, there exists $K : \Omega^2 \to \mathbb{R}$ such that the Hamiltonian of the model is written :

$$\mathcal{H} = \sum_{<i,j>} K(\sigma_i, \sigma_j),$$

where $<>$ represents nearest neighbor summation. We can separate the variables being at the boundary of $\mathcal{L}$ from the others.

$$\mathcal{H} = \sum_{\substack{<i,j> \\ i\notin\partial\mathcal{L}}} K(\sigma_i, \sigma_j) + \sum_{\substack{<i',j'> \\ i'\in\partial\mathcal{L}}} K(\sigma_{i'}, \sigma_{j'})$$

The partition function of the system is given by :

$$\mathcal{Z} = \sum_{\sigma_i,\sigma_{i'}} \prod_{\substack{<i',j'> \\ i'\in\partial\mathcal{L}}} \prod_{\substack{<i,j> \\ i\notin\partial\mathcal{L}}} W(\sigma_i,\sigma_j)W(\sigma_{i'},\sigma_{j'}),$$

where $W(\sigma_i, \sigma_j) = \exp -\beta K(\sigma_i, \sigma_j) \in \mathcal{M}_{|\Omega|}(\mathbb{R})$ - the set of $\mid \Omega \mid \times \mid \Omega \mid$ real matrices. W defines a matrix whose rows and columns indices are the stochastic variables (σ_i). By proceeding to a Singular Value Decomposition (SVD)[11] on W, we get the existence of S, a diagonal positive matrix, U and V, unitary matrices, in $\mathcal{M}_{|\Omega|}(\mathbb{R})$ such that :

$$W(S_i, S_j) = \sum_{l=1}^{|\Omega|} U(\sigma_i, l) S(l, l) V(\sigma_j, l) \qquad \text{or} \quad W = USV^t,$$

As S is diagonal and positive, we can define its square root so to introduce two matrices Q_a and Q_b such that :

$$Q_a = U\sqrt{S} \qquad Q_b = \sqrt{S}V^t \qquad W = Q_a Q_b$$

By using matrices Q_a and Q_b we have managed to separate our variables σ_i and σ_j at the price of introducing an intermediate variable l. This gives the possibility to group all Q's that connect to a site i. We now introduce a tensor $T^{(i)}$ that is the result of integrating σ_i from all Q's. Denoting D(i) the number of neighbors of σ_i [12], T is given by :

$$T^{(i)}_{l_{1,i},\dots,l_{D(i),i}} = \sum_{\sigma_i} Q_{\alpha_1}(\sigma_i, l_{1,i}) \dots Q_{\alpha_{D(i)}}(\sigma_i, l_{D(i),i})$$

The order of this local tensor is therefore the number of neighbors σ_i has. Coming back to the partition function, we find :

$$\mathcal{Z} = \mathrm{Tr} \prod_{i=i}^{\tilde{N}} T^{(i)}_{l_{1,i},\dots,l_{D(i),i}},$$

where the trace operation is defined by summing over all different l indices. Separating the variables at the boundary and at the interior we get :

$$\mathcal{Z} = \mathrm{Tr} \prod_{i \in \partial\mathcal{L}} T^{(i)}_{l_{1,i},\dots,l_{D(i),i}}, \prod_{j \notin \partial\mathcal{L}} T^{(j)}_{l_{1,j},\dots,l_{D,j}},$$

We have simplified the notation for the tensors that belong to the interior of the lattice as all variables have exactly D neighbors. Now that we have separated the boundary and the interior of the lattice, let us conclude by introducing ϕ, the linear form that depends on the boundary conditions of the problem such that :

$$\phi(T^{(1)}, \dots, T^{(N)}) = \mathrm{Tr} \prod_{i \in \partial\mathcal{L}} T^{(i)}_{l_{1,i},\dots,l_{D(i),i}}, \prod_{j \notin \partial\mathcal{L}} T^{(j)}_{l_{1,j},\dots,l_{D,j}}$$

□

Remark 3.3. *This proof can be easily extended to the case of a multipartite lattice by introducing a tensor for each sublattice.*

[11] We will come back later in section 4.2 to mathematical details of the Singular Value Decomposition (SVD).

[12] The purpose of that is to separate the boundary and the interior of the lattice.

Remark 3.4. *If the matrix W is symmetric and positive, then, by the spectral theorem, U equals V and Q_a equals Q_b. In that case, all the tensors that belong to the interior of the lattice are identical.*

Remark 3.5. *If we consider a periodic boundary condition and a symmetric interaction, the partition function simplifies into :*

$$\mathcal{Z} = \mathrm{Tr}\, T^{\tilde{N}},$$

so that in the thermodynamic limit $\tilde{N} \to +\infty$, the properties of the system are governed by the highest eigenvalues of the tensor. Further work could be done to determine whether the highest eigenvalue of the tensor T is simple and whether the simplicity of the highest eigenvalue influences the physical properties of the model. A way to go to that direction would consist in looking at the generalization of the Perron-Frobenius theorem for tensors. A discussion was initiated in [11].

Remark 3.6. *It seems important to consider models where there are no periodic boundary conditions as particular conditions (such as vacuum or random boundary interaction) could change the behavior of the physical properties by suppressing some eigenvalues of the tensor network.*[13]*. Besides, we could consider a lattice with several vacancies inside and study it easily with a tensor network model.*

Corollary 3.7 (Correlators as tensor network states)**.** *Any correlator of a system described by Theorem 3.2 can also be described by a tensor network model.*

Proof. The idea of this proof is based on operator insertion. For simplicity, we will prove the corollary only for the mean of a variable, but the generalization will be immediate. The mean of a variable is given by :

$$< \sigma_\epsilon >= \frac{1}{\mathcal{Z}} \sum_{\sigma_i, \sigma_{i'}} \prod_{\substack{<i',j'> \\ i' \in \partial\mathcal{L}}} \prod_{\substack{<i,j> \\ i \notin \partial\mathcal{L}}} \sigma_\epsilon \, W(\sigma_i, \sigma_j) W(\sigma_{i'}, \sigma_{j'})$$

Following the proof of Theorem 1, we construct all the tensors in the same way as before except $T^{(\epsilon)}$ which is defined by :

$$T^{(\epsilon)}_{l_{1,\epsilon},...,l_{D(\epsilon),\epsilon}} = \sum_{\sigma_\epsilon} \sigma_\epsilon \, Q_{\alpha_1}(\sigma_\epsilon, l_{1,\epsilon}) ... Q_{\alpha_{D(\epsilon)}}(\sigma_\epsilon, l_{D(\epsilon),\epsilon})$$

The only thing we did in this construction is to insert the variable σ_ϵ inside one of the tensors. The end of the proof is strictly the same as the proof of Theorem 3.2. □

Remark 3.8. *Following Corollary 3.7, if we choose $\tilde{N}$ operators $\Theta_1, ..., \Theta_{\tilde{N}}$, we can easily construct the quantity :*

$$< \Theta_1(\sigma_1) ... \Theta_{\tilde{N}}(\sigma_{\tilde{N}}) >$$

[13] In classical physics, the modes that propagate along a string are determined by boundary conditions for example.

Remark 3.9. *Tensor network models have an underlying gauge invariance. In our construction, an index i is shared by only two tensors, let us call them* $T^{(1)}_{ijkl}$ *and* $T^{(2)}_{ipqr}$*. Recalling that i can take* $\mid \Omega \mid$ *values we can introduce a matrix* O *in* $GL_{|\Omega|}(\mathbb{C})$ *such that making the transformation* $T^{(1)}_{ijkl} \longmapsto \sum_m T^{(1)}_{mjkl} O_{i,m}$ *and* $T^{(2)}_{ipqr} \longmapsto \sum_m T^{(2)}_{mpqr} O^{-1}_{m,i}$ *leaves the partition function and any correlator unchanged as* $\sum_i O^{-1}_{m,i} O_{i,m'} = \delta_{m,m'}$*.*

3.2 Examples

3.2.1 Matrix product model in 1-dimension

It is known in the literature that the 1D Ising model can be solved analytically by representing it as a matrix product model. We derive in this section this simple example of tensor network state. Let $\mathcal{H}$ be the following Hamiltonian :

$$\mathcal{H} = \sum_{i=0}^{N-1} (-J\sigma_i\sigma_{i+1} - \frac{1}{2}h(\sigma_i + \sigma_{i+1})),$$

where we assumed periodic boundary conditions, i.e $\sigma_0 = \sigma_N$. Let now $\mathcal{Z}$ be the partition function of that model.

$$\mathcal{Z} = \sum_{\sigma_i=\pm 1} \prod_{i=0}^{N-1} \exp\left[\beta[J\sigma_i\sigma_{i+1} + \frac{1}{2}h(\sigma_i + \sigma_{i+1})]\right)$$

We can now define the matrix T, such that :

$$T_{\sigma\sigma'} = \exp \beta(J\sigma\sigma' + \frac{1}{2}h(\sigma + \sigma'))$$

Rewriting $\mathcal{Z}$ in terms of T, we find :

$$\mathcal{Z} = \sum_{\sigma_i=\pm 1} T_{\sigma_0\sigma_1} T_{\sigma_1\sigma_2} ... T_{\sigma_{N-1}\sigma_0},$$

which we can rewrite :

$$\mathcal{Z} = \mathrm{Tr}(T^N) \tag{10}$$

Equation (10) has strictly the same form as the result obtained before in Theorem 3.2. We have therefore managed to rewrite the 1D Ising model with an external field as a tensor network model.

To push the analysis forward, let us look at the expected value of one spin in this tensor network representation.

$$< \sigma_p >= \frac{\sum_{\sigma_i=\pm 1} T_{\sigma_0\sigma_1} ... T_{\sigma_{p-1}\sigma_p} \sigma_p T_{\sigma_p\sigma_{p+1}} ... T_{\sigma_{N-1}\sigma_0}}{\mathrm{Tr}(T^N)}$$

Introducing the third Pauli matrix S as :

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$< \sigma_p >= \frac{\mathrm{Tr}(T^p S T^{N-p})}{\mathrm{Tr}(T^N)} = \frac{\mathrm{Tr}(T^N S)}{\mathrm{Tr}(T^N)}$$

Let us now look at a two point correlation function to conclude the construction of the full tensor network model of the 1D Ising model.

$$< \sigma_0 \sigma_p >= \frac{\sum_{\sigma_i = \pm 1} \sigma_0 T_{\sigma_0 \sigma_1} ... T_{\sigma_{p-1} \sigma_p} \sigma_p T_{\sigma_p \sigma_{p+1}} ... T_{\sigma_{N-1} \sigma_0}}{\mathrm{Tr}(T^N)}$$

$$< \sigma_0 \sigma_p >= \frac{\mathrm{Tr}(S T^p S T^{N-p})}{\mathrm{Tr}(T^N)}$$

We finally see that our tensor network model gives us symmetries[14] for any correlation function, and more generally for any observable function.

3.2.2 Tensor product model in 2-dimension

After having derived matrix product results in one dimension, we try to extend it to two dimensions keeping the Ising model as our toy model. Things tend to be somehow tedious in two-dimension as we have to choose a lattice. In this example, we will define our lattice to be a triangular lattice as in figure 3 and we will follow the approach of [10].

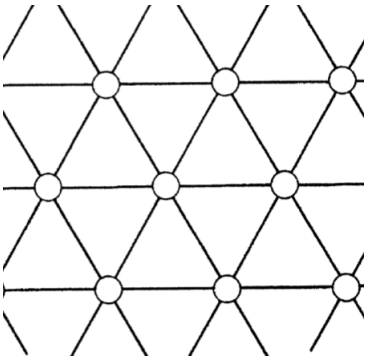


Figure 3: Triangular lattice

We consider the usual Hamiltonian and the related partition function :

$$\mathcal{H} = -J \sum_{<i,j>} \sigma_i \sigma_j$$

$$\mathcal{Z} = \sum_{\{\sigma\}} \prod_{\triangle_{ijk}} \exp\left[\beta \frac{J}{2} (\sigma_i \sigma_j + \sigma_j \sigma_k + \sigma_k \sigma_i)\right], \tag{11}$$

where the sum is made over all spin configurations and the product is made over all triangular unit cells. The factor $1/2$ comes from the fact that an edge

[14] As our tensors are simply matrices and the trace is invariant by circular permutation, symmetries are translation plus periodicity or rotation if we consider that a 1D lattice with periodic boundary conditions is a circular compactification of the real line.

is shared by two triangles.

We can go a little bit further by defining edge spins $\{\Sigma\}$ in the following manner : $\Sigma_{ij} = \sigma_i\sigma_j$. These edge spins keep the spin structure by allowing only ± 1 values, the interpretation lies now on whether site spins are parallel or anti-parallel. They have a closure relation which is deduced from the structure of the initial spins :

$$\Sigma_{ij}\Sigma_{jk}\Sigma_{ki} = 1 \tag{12}$$

Our goal is now to modify the partition function to use only edge spins which we can do by first introducing the following identity :

$$\sum_{\{\Sigma\}} \prod_{<ij>} \delta(\Sigma_{ij} - \sigma_i\sigma_j) = 1, \tag{13}$$

where the sum is made over all edge spin configurations and the product over nearest neighbors. It means that there is a unique configuration of $\{\Sigma\}$'s that represents correctly the configuration of $\{\sigma\}$'s in the lattice, it is possible for an edge is shared by exactly two sites.

Combining equations (11) and (13), we get :

$$\mathcal{Z} = \sum_{\{\sigma\}} \sum_{\{\Sigma\}} \prod_{<ij>} \delta(\Sigma_{ij} - \sigma_i\sigma_j) \prod_{\triangle_{ijk}} \exp\left[\beta\frac{J}{2}(\Sigma_{ij} + \Sigma_{jk} + \Sigma_{ki})\right]$$

As our goal is to manipulate only edge spins, we can exchange the two sums. This tell us that we now have to work one the following term :

$$\sum_{\{\sigma\}} \prod_{<ij>} \delta(\Sigma_{ij} - \sigma_i\sigma_j) \tag{14}$$

There has not been a rigorous treatment of equation (14) in the literature. There was an attempt in [10], but no proof was stated, we will try to bring a sketch of a proof here. We assume here a periodic boundary condition for the Ising model. A way to get this condition is to compactify the triangular lattice. Compactifying is really easy for 1-dimensional systems as the real line gets mapped to the circle S^1, but for 2-dimensional systems, it is not obvious at all that we have to compactify on the sphere S^2 or on the torus $S^1 \times S^1$. These topological spaces are really different for our problem, and this is the main key of a rigorous treatment of equation (14). The difference between the sphere and the torus for our problem lies in the Euler characteristic. This notion comes from algebraic topology and states that if we can compactify the triangular lattice on a torus, then we have the following relation :

$$N_{edges} = N_{triangles} + N_{sites}, \tag{15}$$

where N designates the number of edges, triangles or sites. Had it been possible to compactify the triangular lattice on the sphere, equation (15) would be modified in the following way :

$$N_{edges} + 2 = N_{triangles} + N_{sites} \tag{16}$$

There is a clear difference between equations (15) and (16) which tells us that we now have to prove whether indeed the triangular lattice can be compactified on a torus without any topological defect[15]. If we want to pave a surface with T triangles, we have $3T$ edges, but an edge is shared by exactly two triangles, so we effectively have $3T/2$ edges. There are also $3T$ sites, but as each site is shared by six triangles, we effectively have $T/2$ sites. This little reasoning proves that equation (15) is the good one and consequently that the compactification on a sphere is not possible without introducing defects. What is left to prove is the existence of a compactification that was just assumed at this point. This was done rigorously in [12] [16].

Looking back at equation (13), we see that we have to verify the conditions of N_{edges} delta functions, but according to equation (15) we equivalently have $N_{triangles} + N_{sites}$ conditions to fulfil. As we are summing over all $\{\Sigma\}$ configurations in equation (13), there are configurations where the closure relation (12) is not verified. Yet, these configurations are unphysical, and the delta functions prevent them to play a role in the partition function by giving them a zero weight. There are clearly as many closure relations as $N_{triangles}$, so we might replace the first $N_{triangles}$ conditions by a projection operator which image is the physical space where the closure relation is verified. It is common in the literature, in [10] for instance, to choose this projector to be :

$$P(\Sigma_{ij}, \Sigma_{jk}, \Sigma_{ki}) = \frac{1 + \Sigma_{ij}\Sigma_{jk}\Sigma_{ki}}{2},$$

where the operator P leads 1 for a physical state and 0 for an unphysical state. Other choices can be done, for that reason, we will be as general as possible, keeping the operator P in our equations.

Coming back to equation (14), we see that summing over all $\{\sigma\}$ configurations let us obtain the last N_{sites} conditions up to the global $\mathbb{Z}_2$-symmetry $\{\sigma\} \mapsto \{-\sigma\}$[17]. All arguments combined, we find that that equation (14) can be rewritten as :

$$\sum_{\{\sigma\}} \prod_{<ij>} \delta(\Sigma_{ij} - \sigma_i\sigma_j) = \prod_{\triangle_{ijk}} 2P(\Sigma_{ij}, \Sigma_{jk}, \Sigma_{ki}),$$

where the 2 stands for the global $\mathbb{Z}_2$-symmetry. Finally, we can reformulate the partition function to be in the following shape :

$$\mathcal{Z} = \sum_{\{\Sigma\}} \prod_{\triangle_{ijk}} 2P(\Sigma_{ij}, \Sigma_{jk}, \Sigma_{ki}) \exp\left[\beta \frac{J}{2}(\Sigma_{ij} + \Sigma_{jk} + \Sigma_{ki})\right]$$

As we now have a result depending on all edge spins of one triangle and we multiply the result for every triangle, we can as well denote by i the index of

[15] A topological defect would be for instance a site on the triangular lattice that has five neighbors instead of six.

[16] There are also empirical evidence of such compactification in the field of molecular physics where carbon nanotori were formed from graphene sheets and carbon nanotubes. Numerical predictions are mentionned in [13] and experimental results in [14].

[17] This point is, up to our knowledge, not discussed in the literature. The symmetry might be broken in the presence of an external non zero magnetic field.

a triangle, and $\Sigma_{x(i)}, \Sigma_{y(i)}, \Sigma_{z(i)}$ the three edge spins of this triangle. Changing the notations, we now find for the partition function :

$$\mathcal{Z} = \sum_{\{\Sigma\}} \prod_i 2P(\Sigma_{x(i)}, \Sigma_{y(i)}, \Sigma_{z(i)}) \exp\left[\beta \frac{J}{2}(\Sigma_{x(i)} + \Sigma_{y(i)} + \Sigma_{z(i)})\right]$$

This allows us to define a tensor T such that :

$$T_{x(i),y(i),z(i)} = 2P(\Sigma_{x(i)}, \Sigma_{y(i)}, \Sigma_{z(i)}) \exp\left[\beta \frac{J}{2}(\Sigma_{x(i)} + \Sigma_{y(i)} + \Sigma_{z(i)})\right]$$

We can finally rewrite the partition function as a tensor network state, ending here this section.

$$\mathcal{Z} = \sum_{\{\Sigma\}} \prod_i T_{x(i),y(i),z(i)}$$

3.3 Discussion on lattices and planar dual lattices

We have seen in the construction of tensor network models that the order of a tensor is equal to the number of neighbors each sites has in the interior of the lattice. In dimension two, we would expect this number to be 4 for a square lattice, 6 for a triangular lattice and 3 for a honeycomb lattice. Therefore, according on which lattice the problem is stated, we will have more or less difficulties to construct and manipulate tensor networks.

As lattices and planar dual lattices can have different topology - as we will see, the planar dual lattice of a triangular lattice is a honeycomb lattice, on which each site has less neighbors than on the original lattice - it might be interesting to see whether we can formulate tensor networks either in one lattice or in the other one.

Definition 3.10 (Lattice and planar dual lattice)**.** For any planar lattice, one can define its geometrical dual by tracing the perpendicular bisector of each edge and defining the dual sites as the intersection of these bisectors. Alternatively, one can define the center of each unit cell of a regular lattice and join all these centers.

An example of dual lattice construction is given in figure 4. We see that the square lattice is a self dual whereas the triangular and honeycomb lattices are dual of each other.

Within the proof of existence of tensor networks in section 3.1, we placed a tensor on every site of a lattice. But also, when we presented examples of tensor networks in section 3.2, we placed the tensors either for the 1-dimensional Ising model at the middle of each edge, or for the 2-dimensional Ising model at the center of each triangular unit cell, and accordingly to the definition of a planar dual lattice, we placed the tensors at the sites of the dual lattices. We have therefore shown two methods[18] to construct tensor networks, one of which should prove to be more useful than the other in some situations as discussed in this section.

[18]This also indicates that we do not have uniqueness of the networks. It would certainly be useful to study whether this non-uniqueness arises because of a gauge symmetry or because the problem was ill-posed at some point.

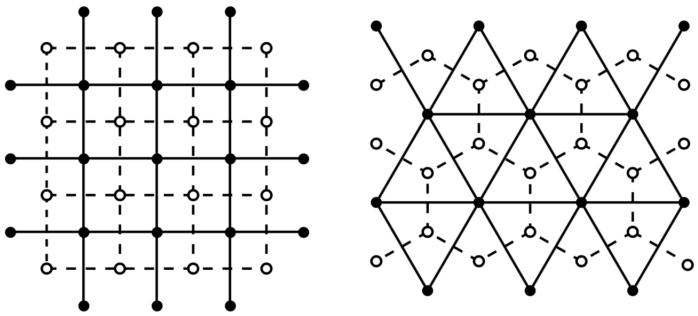


Figure 4: Construction of planar dual lattices, taken from [3]

4 Tensor network renormalization

While in the previous renormalization schemes we saw, we were following the flow of Hamiltonians, the new paradigm in tensor network renormalization consists in following the flow of tensors. The introduction of tensor network models enables the use of various tools coming from linear algebra : Perron-Frobenius theorem, Singular Value Decomposition (SVD) to mention some. Several renormalization schemes were proposed, the following sections are dedicated to discuss some of the most recent results.

4.1 Levin-Nave tensor renormalization group

The first tensor network renormalization results were published by Levin and Nave in [15]. We will follow their approach and question some parts of their construction.

We restrict ourselves to a tensor network model on the 2-dimensional honeycomb lattice with N sites and periodic boundary conditions as presented in figure 5.

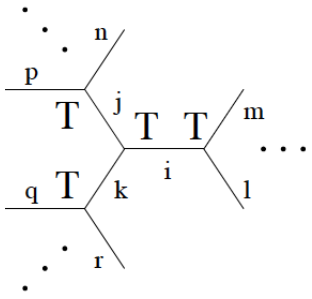
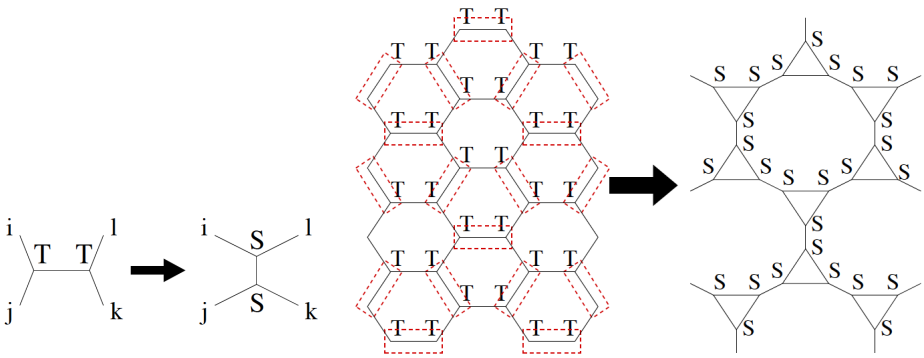


Figure 5: Tensor network model on the 2-dimensional honeycomb lattice, taken from [15]

To keep the conventions introduced when constructing tensor network states, the tensor of the model is (T_{ijk}). It is of order three as every site has three neighbors and we will assume each index can take D values. To respect the symmetries of the lattice, namely rotations by $\frac{2\pi}{3}$ and $\frac{4\pi}{3}$, T should be cyclically symmetric : $(T_{ijk}) = (T_{jki}) = (T_{kij})$.

Recalling that the partition function can be written, $\mathcal{Z} = \mathrm{Tr}\, T^N$, Levin and Nale proposed a scheme called tensor renormalization group to compute the partition function. The procedure is isotropic and coarse-graining - in the spirit of blocking renormalization - and besides using ideas from early ages renormalization, it introduces concepts coming from quantum information.

Tensor renormalization group is done in two steps. The first step is to find a tensor S to transform locally the lattice in the way presented in figure 6a and and more globally as in figure 6b, it is supposed not to affect the partition function.



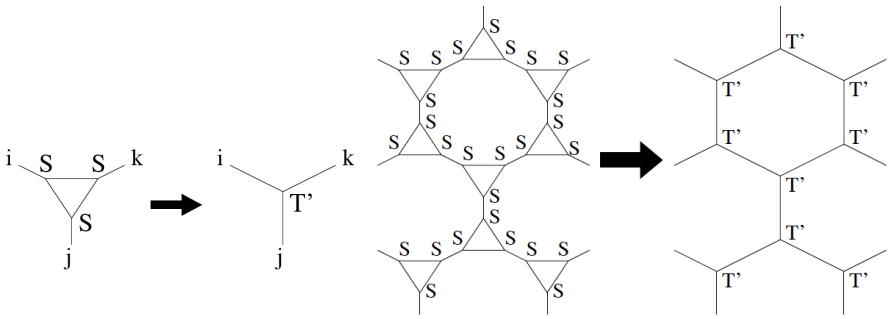
(a) Local lattice transformation (b) Global lattice transformation

Figure 6: First step of tensor renormalization group, taken from [15]

The second step consists on contracting the tensors S that are nearest neighbors as presented in figures 7a and 7b. The second step enables the definition of a new tensor T', and as we have to contract three S tensors to get a tensor T', we deduce that the number of sites in the lattice has been divided by three and that the new expression of the partition function is given by $\mathcal{Z} = \mathrm{Tr}\, T'^{\frac{N}{3}}$. This suggests us that we no longer are following a Hamiltonian flow, but rather a tensor flow.

As it is expected from the formulation "find a tensor S...", the first step of the renormalization will be approximate whereas the second one will be exact. Comparing figures 6b and 7b, we can clearly see that the result of this transformation is to obtain a coarse grained honeycomb lattice.

Now that we introduced graphical ideas for tensor renormalization group, let us construct the flow theoretically. The first part is the construction of the tensor S, suppose such a tensor exists, figure 6a tells us the following equation



(a) Local lattice transformation (b) Global lattice transformation

Figure 7: Second step of tensor renormalization group, taken from [15]

should be verified :

$$\sum_{n=1}^{D} S_{lin} S_{jkn} = \sum_{m=1}^{D} T_{ijm} T_{klm} \tag{17}$$

As it is easier to manipulate matrices than tensors, we introduce the $D^2 \times D^2$ matrix M such that :

$$M_{li,jk} = \sum_{m=1}^{D} T_{ijm} T_{klm}$$

and consider the tensor S as a $D^2 \times D$ matrix $S_{li,n}$. The problem than gets written :

$$M = SS^T \tag{18}$$

We managed to reduce the first step of the renormalization to a linear algebra problem. All the difficulties actually come from this linear algebra problem. M can have a rank on the whole range of $[0, D^2]$ whereas SS^T has a rank in $[0, D]$. In the most general case, we can assume that equation (18) has no exact solution, we will have to approximate the solution.

The way Levin and Nave proceeded to find an approximation was to define an optimization problem :

$$S = \underset{S \in \mathcal{M}(\mathbb{R}^D, \mathbb{R}^{D^2})}{\operatorname{argmin}} \|M - SS^T\|^2, \tag{19}$$

where $\|.\|$ is some matrix norm. Levin and Nave did not make a particular choice of norm, and we will try to discuss that more in details in section 4.2.

The idea behind equation (19) is somehow to factorize the matrix M. There exists two tools in linear algebra that allow us to factorize a matrix or a tensor in the most general case : Singular Value Decomposition (SVD) and Schmidt decomposition, we will come back in section 4.2 on precise mathematical details of these tools. An SVD on M gives us the existence of Λ, a diagonal positive matrix, U and V, unitary matrices, in $\mathcal{M}_{D^2}(\mathbb{C})$ such that $M = U\Lambda V^\dagger$.

One can than truncate Λ to keep the D biggest singular values of M and to reduce the rank of M to exactly D and remove the related columns of U and V. It allows to define $S^A = U\sqrt{\Lambda}$ and $S^B = V^*\sqrt{\Lambda}$ up to a $GL_D(\mathbb{C})$ gauge invariance[19].

Coming back to our original tensor T, we managed to build two tensors S^A and S^B such that :

$$\sum_{n=1}^{D} S^A_{lin} S^B_{jkn} = \sum_{m=1}^{D} T_{ijm} T_{klm} \tag{20}$$

Comparing equations (17) and (20), we see that our construction allows to transform the lattice as in figure 6a but the price to pay is the creation of two sublattices A and B. Finally we define a tensor T' in the following way :

$$T'_{ijk} = \sum_{p,q,r=1}^{D} S_{krp} S_{jqr} S_{irp}$$

Let us state a few remarks before ending this section. Something not mentionned in [15] is whether the triangle in figure 7a is composed of three tensors of type A or B or a combination of these (two A's and one B or one A and two B's). Something else that was not checked is the cyclically symmetry of the new tensor T', it does not seem trivial that the renormalization scheme conserved that property, all the more that we broke the rotational symmetry of the original lattice by introducing sublattices. Another cause of symmetry breaking is the gauge invariance, as we could easily choose a gauge for each sublattice or for each site. Finally, as we divide the number of sites by three at each renormalization iteration, we will need $\mathcal{O}(\log_3(N))$ iterations to have only one site on the lattice.

4.2 Optimization, truncated SVD and entanglement

This section is dedicated to question part of Levin-Nave renormalization scheme and to highlight the role of quantum information inside this renormalization scheme.

4.2.1 Optimality of truncated SVD

The work in this section is meant to show why the SVD appears in the calculations of Levin and Nave in [15], in fact we will see that whenever we need to lower the rank of a matrix while minimizing a cost, the norm of the difference of the matrices, we necessarily have to go through an SVD. Let us first go back to equation (19). We need to define a norm on the matrix space. Even if we work on finite dimensional spaces and all norms are equivalent in these spaces, optimization problems have different solutions depending on the norm considered.

[19] In fact, equivalently, $\forall \Phi \in GL_D(\mathbb{C}), M = U\sqrt{\Lambda}\Phi\Phi^{-1}\sqrt{\Lambda}V^\dagger$. The first step of the renormalization therefore conserves the gauge invariance we noticed while constructing tensor network models.

Definition 4.1 (Hilbert-Schmidt inner product and norm)**.** For D in $\mathbb{N}$ and A, B in $\mathcal{M}_D(\mathbb{R})$, we define the Hilbert-Schmidt inner product $\langle A, B\rangle_{HS} = \mathrm{Tr}(A^T B)$ and the Hilbert-Schmidt norm $\|A\|^2_{HS} = \mathrm{Tr}(A^T A)$.

If $A = (A_{ij})$, then $\|A\|^2_{HS} = \sum_{i,j=1}^{D} \mid A_{ij} \mid^2$.

If A is symmetric, calling λ_i its eigenvalues, then $\|A\|^2_{HS} = \sum_{i=1}^{D} \mid \lambda_i \mid^2$

Without stating any proof, we will now introduce the SVD and discuss its variants.

Theorem 4.2 (Singular Value Decomposition)**.** *For n,m in* $\mathbb{N}$ *and A in* $\mathcal{M}(\mathbb{R}^m, \mathbb{R}^n)$ *there exists a factorization called singular value decomposition.*

$$\exists\, U \in O(n),\ V \in O(m),\ S \in \mathcal{M}(\mathbb{R}^m, \mathbb{R}^n) \text{ s.t. } \quad A = USV^T$$

where S is diagonal positive and its coefficients are called the singular values. The singular values are the square roots of the eigenvalues of both AA^T *and* $A^T A$*. There exists rank(A) strictly positive singular values.*

Remark 4.3. *Generally, it us useful to get the singular values in descending order so that* $S_{11} \geq S_{22} \geq ... \geq S_{\min(n,m)\,\min(n,m)}$.

Remark 4.4. *There are other possibilities for the dimensions of U, S and V. Suppose rank(A)=p, we can choose S to be in* $\mathcal{M}(\mathbb{R}^p, \mathbb{R}^p)$*, U in* $\mathcal{M}(\mathbb{R}^p, \mathbb{R}^n)$ *and V in* $\mathcal{M}(\mathbb{R}^m, \mathbb{R}^p)$*.*

Remark 4.5. *The SVD tells us that if rank(A)=p, then A is the sum of p matrices of rank 1.* $A = \sum_{i=1}^{p} S_{ii} \underline{U_i}\, \underline{V_i}^T$*, where* $\underline{U_i}$ ($\underline{V_i}$) *designate the ith column of U (V).*

Remark 4.6. *If we truncate the singular values matrix to* χ *values instead of rank(A), we will call* $A^{(\chi)}$ *the approximate matrix.*

Now that we properly introduced a norm on our matrix space and the SVD, we can come back to tensor renormalization group. Observing that M is symmetric, equation (19) describes the following minimization problem :

$$\min_{S \in \mathcal{M}(\mathbb{R}^D, \mathbb{R}^{D^2})} \mathrm{Tr}(M - SS^T)^2 \tag{21}$$

We are now going to prove that $SS^T = M^{(D)}$ minimizes equation (21).

Theorem 4.7 (Low rank approximation - Eckart-Young-Mirsky theorem)**.** *Let us consider the low rank approximation problem, namely :*

$$\min_{N \in \mathcal{M}(\mathbb{R}^{D^2}, \mathbb{R}^{D^2})} \|M - N\|^2_{HS}, \qquad \text{subject to } \mathrm{rank}(N) \leq D$$

The solution to this problem is $M^{(D)}$ *if rank(M)>D, otherwise it is M.*

Proof. Let us use the SVD of M. There exists U and V in $O(D^2)$, S in $\mathcal{M}_{D^2}(\mathbb{R})$ such that $M = USV^T$. Let us define A as $A = U^T N V$ and let p be the rank

of M. Then $\|M-N\|_{HS}^2 = \|S-A\|_{HS}^2$ by rotational invariance of the norm. Expressing the norm in terms of matrices coefficients we have :

$$\|S-A\|_{HS}^2 = \sum_{i=1}^{p}(S_{ii}-A_{ii})^2 + \sum_{i=p+1}^{D^2} A_{ii}^2 + \sum_{i,j=1,j\neq i}^{D^2} A_{ij}^2$$

We can cancel the second and third sum by tuning all coefficients to zero. We now have to deal with the first sum remembering that rank(N)=rank(A)$\leq D$. As the singular values are ordered, we might want to suppress from the sum the contribution of the largest singular values. This leaves us two cases, either $p \leq D$ in which case we can choose A=S and then $\|S-A\|_{HS} = 0$, or $p > D$ and we can only cancel the D first singular values and $\|S-A\|_{HS}^2 = \sum_{i=p+1}^{D}(S_{ii})^2$. In the first case N=M and in the latter case, A is diagonal with the first D singular values of M, and consequently $N = M^{(D)}$. □

This proof ends the mathematical construction of Levin-Nave renormalization scheme and shows why the SVD method is naturally introduced to minimize the error introduced by the renormalization, in the least square sense.

4.2.2 Renormalization and entanglement

In [15], Levin and Nave claim that fundamental ideas from quantum information lie inside their new renormalization scheme. We will try to understand in this section what these ideas actually are.

Quantum information in usually linked to statistical physics through the notion of entropy. We will see now how we can reintroduce the notion of entropy in the context of tensor network renormalization.

Definition 4.8 (Von Neumann entropy)**.** Let A be a matrix of rank p which singular values are $\{S_1, ..., S_p\}$. We normalize the singular values $S_i \mapsto S_i/\sqrt{\sum_j S_j^2} = S_i'$ and define the entropy of this matrix as :

$$\mathcal{S}(A) = -\frac{1}{\ln p}\sum_{i=1}^{p} S_i'^2 \ln S_i'^2,$$

where we impose the factor $1/\ln p$ to have an entropy in the interval [0,1].

In quantum mechanics, the Von Neumann entropy generally measures the entanglement between two particles which can be seen as a quantum correlation. In classical physics, increasing the value of the entropy implies a loss of information in the system whereas in quantum physics, increasing the entropy implies an increase in these correlations. In our renormalization scheme, we should carefully control how the entropy evolves at each step. If the entropy increases, intuitively we also increased some correlations which is in contradiction with the ideas of renormalization : we want to suppress irrelevant correlations, and enforce them.

Let us come back to Levin and Nave scheme at the point where they proceed to a local lattice transformation as presented in figure 8. The matrix Λ

represents the singular values of M and can be interpreted in terms of the Von Neumann entropy as a measure of the entanglement between the new lattice sites where the matrices U and V are defined. It would be interesting to evaluate numerically the evolution of $\mathcal{S}(\Lambda)$ as a function of the renormalization step to determine whether we expect a breakdown from the scheme.

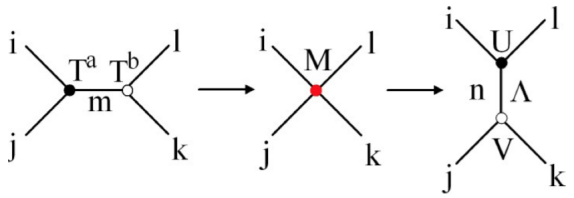


Figure 8: Lattice transformation in Levin and Nave renormalization scheme

It was pointed out in [16] that the Levin-Nave tensor renormalization group is unable to suppress irrelevant local correlations as the scheme admits a fixed point called *Corner Double Line* which conserves local correlations. Evenbly proposes in [17] a new renormalization scheme to overcome this difficulty, we will come back later to this.

4.2.3 Choice of an error to minimize

We question in this section the proposition of Levin and Nave in [15] for the optimization problem presented in equation (19). They want to minimize a local truncation error in the lattice. Yet, we do not know what physical implications this minimization has. Particularly, we may wonder what is the error induced in the partition function or in the free energy.

Like we did in section 1.2, we can also modify the minimization problem of Levin and Nave by adding different constraints. The current problem consists of minimizing the Hilbert-Schmidt norm subject to a rank constraint. The constraint we could add is the conservation of the Von Neumann entropy.

To proceed to that, it might be useful to relate the Von Neumann entropy to the Hilbert-Schmidt inner product. Let M be a matrix with the following SVD decomposition $M = U\Lambda V^T$, let M' be equal to $\mathrm{M}/\|M\|_{HS}$. We define the logarithm of $\tilde{M}$ as $\ln(\tilde{M}) = U \ln(\tilde{\Lambda}) V^T$, where we allow the value $\ln 0$. Then this identity follows :

$$\begin{aligned} \langle \tilde{M}\tilde{M}^T, \ln(\tilde{M}\tilde{M}^T) \rangle_{HS} &= \mathrm{Tr}(\tilde{M}\tilde{M}^T \ln(\tilde{M}\tilde{M}^T)) \\ &= \mathrm{Tr}(U\tilde{\Lambda}V^T V\tilde{\Lambda}^T U^T U \ln(\tilde{\Lambda}\tilde{\Lambda}^T)U^T) \\ &= -\mathcal{S}(M), \end{aligned} \tag{22}$$

where we suppressed the ambiguity of $\ln 0$ by stating $0 \ln 0 = 0$. This identity could prove to be useful if we want to express the entropy constraint as an inner product for an optimization problem[20].

[20] Additionally, if we want to express the rank of a matrix as norm, it is interesting to introduce the L^0-norm of the vector of singular values.

Coming back to figure 8, we can define two matrices $M^{(1)}$ and $M^{(2)}$ on the original and modified lattice in the following way :

$$M^{(1)}_{ji,lk} = \sum_{m=1}^{D} T_{ijm} T_{klm}$$

$$M^{(2)}_{li,jk} = \sum_{m=1}^{D} T_{ijm} T_{klm}$$

We can as well define the entropy of both matrices. The entropy of the matrix $M^{(1)}$ is interpreted as the entanglement between two sites characterized by the tensors T. If we want the entropy to be conserved during our renormalization scheme, we should modify (21) in the following manner :

$$\min_{S \in \mathcal{M}(\mathbb{R}^D, \mathbb{R}^{D^2})} \mathrm{Tr}(M^{(2)} - SS^T)^2, \qquad \text{subject to } \mathcal{S}(SS^T) = \mathcal{S}(M^{(1)}) \tag{23}$$

The interpretation of equation (23) is that the truncated SVD will not be the optimal solution to our problem anymore. We will rather have to choose which singular values of $M^{(2)}$ we will keep to construct the matrix S so to conserve the entropy. The problem now becomes a combinatorial one, as we have to choose D singular values among $\mathrm{rank}(M^{(2)})$ to satisfy the entropy constraint, if different combinations verify the constraint, we will choose the one that minimizes the Hilbert-Schmidt norm. As we want to proceed to several renormalization steps, it might be useful to find a generic numerical algorithm that is perhaps faster than looking at $\binom{\mathrm{rank}(M^{(2)})}{D}$ different combinations. Recalling the method used in section 1.2, we have a way to add the constraint in the minimization problem. We propose a scheme to modify equation (23) based on this method. We define a Lagrange multiplier λ and set the new minimization problem to be :

$$\min_{S \in \mathcal{M}(\mathbb{R}^D, \mathbb{R}^{D^2})} \mathrm{Tr}(M^{(2)} - SS^T)^2 + \lambda[\mathcal{S}(SS^T) - \mathcal{S}(M^{(1)})] \tag{24}$$

We might also rewrite equation (24) only in terms of Hilbert-Schmidt inner product using the relation derived in equation (22). We can transform our minimization problem on a matrix space into a problem on a vector space of singular values. Calling Λ_i (respectively S_i) the singular values of $M^{(2)}$ (respectively SS^T), if we suppose that (Λ_i) are in descending order and that $\mathrm{rank}(M^{(2)})$=p>D, this yields :

$$\min_{S_1 \ldots S_p} \sum_{i=1}^{p} (\Lambda_i - S_i)^2 - \lambda[\frac{1}{\ln D} \sum_{i=1}^{p} \frac{S_i^2}{\sum_j S_j^2} \ln \frac{S_i^2}{\sum_j S_j^2} + \mathcal{S}(M^{(1)})], \tag{25}$$

subject to forcing $(p - D)$ singular values among $(S_1 \ldots S_p)$ to be equal to zero which imposes $\mathrm{rank}(SS^T)$ to be equal to D[21].

[21] Once again we can introduce the L^0-norm of the singular values vector and impose its value to be equal to D.

A last remark on the scheme we proposed would be to ask ourselves whether we should also have conservation of the rank of $M^{(1)}$, namely rank(SS^T)=rank($M^{(1)}$). This concludes this section where we proposed an improvement to Levin and Nave scheme.

4.3 Evenbly-Vidal tensor network renormalization

After Levin-Nave tensor renormalization group appeared, several ameliorations were developped. A very recent model entitled Tensor Network Renormalization was introduced by Vidal and Evenbly in [16] and [17], they claimed that their scheme does not accumulate short-range correlations as Levin-Nave scheme does thanks to the use of disentanglers. We will follow in this section their approach to give a review of the most advanced properties of tensor network renormalization.

Vidal and Evenbly restrict themselves to the 2-dimensional Ising model on the square lattice and transform it to a tensor network as presented in figure 9. The tensor A is defined to verify $A_{ijkl} = e^{\beta J(\sigma_i\sigma_j+\sigma_j\sigma_k+\sigma_k\sigma_l+\sigma_l\sigma_i)}$ and as previously, the partition function is equal to the trace of the tensor product of the whole network. Denoting $\mathcal{F}$ the 2×2 block of tensors, like in the Levin-Nave scheme, the idea is to modify each subnetwork (as in figure 10) to coarse-grain the global lattice (as in figure 11) at the expense of approximation error. The error is quantified as the Hilbert-Schmidt norm of the difference between the block $\mathcal{F}$ and the approximate block $\tilde{\mathcal{F}}$[22].

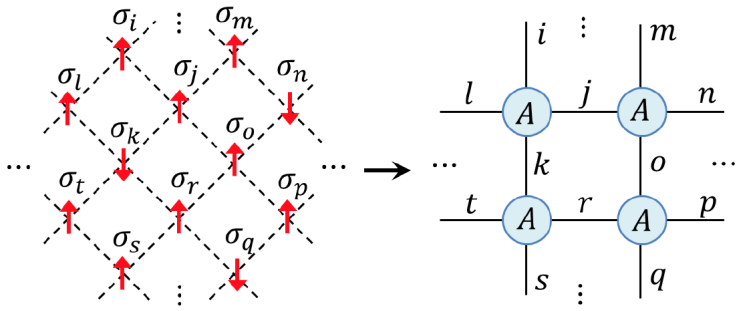


Figure 9: The partition function of the 2-dimensional Ising model on the square lattice can be expressed as a network of 4-dimensional tensors A. Each tensor encodes for the Boltzmann weight in a given square, taken from [17].

As seen in section 4.1 the approximation can be done through a truncated SVD. The innovation in Evenbly-Vidal tensor network renormalization lies in the introduction of a more general class of approximation called projective truncations. It requires the definition of a projector on P acting on block of tensors that verifies $PP^\dagger = P^2 = P$. For instance, a matrix $P = ww^\dagger$ with $w^\dagger w = \mathbb{1}$ is such a projector. We then formally define $\tilde{\mathcal{F}}$ to be equal to $P\mathcal{F}$. The use of a

[22]Evenbly gives a graphical representation of the Hilbert-Schmidt norm in [17] which turns the renormalization scheme into a graphical problem.

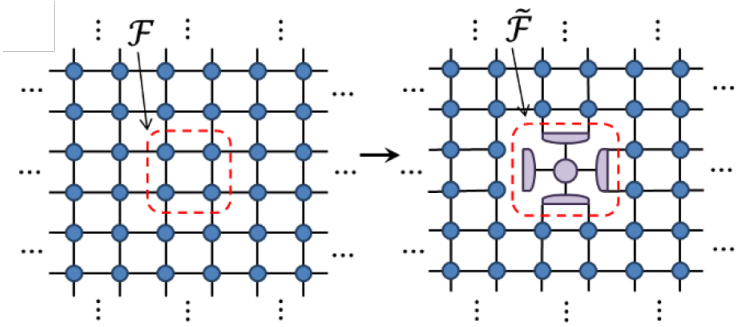


Figure 10: Transformation of a subnetwork $\mathcal{F} \mapsto \tilde{\mathcal{F}}$, taken from [17].

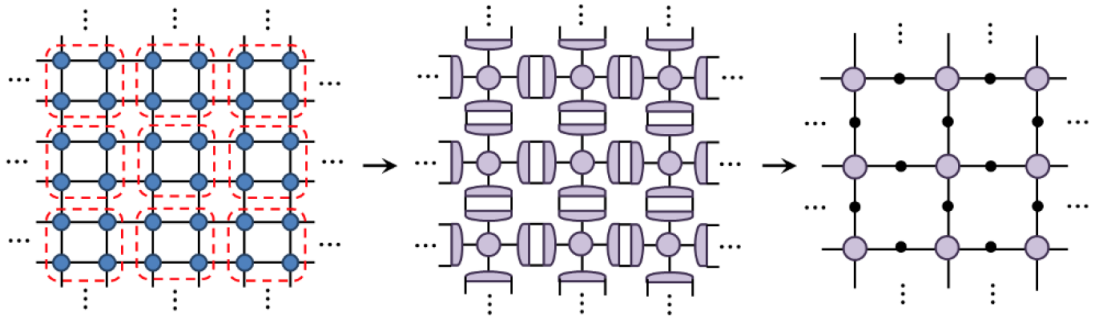


Figure 11: Coarse-graining step made through intermediate tensor contraction, taken from [17].

projector allows us to make the following observation :

$$\langle \tilde{\mathcal{F}}, \tilde{\mathcal{F}} \rangle_{HS} = \langle \mathcal{F}, \tilde{\mathcal{F}} \rangle_{HS}$$

The error ϵ gets simplified as follows :

$$\epsilon^2 = \|\mathcal{F} - \tilde{\mathcal{F}}\|_{HS}^2 = \|\mathcal{F}\|_{HS}^2 - \|\tilde{\mathcal{F}}\|_{HS}^2$$

We finally see that in order to minimize the error, we have to maximize the norm of $\tilde{\mathcal{F}}$.

Now that we have the abstract idea of the coarse graining scheme, let us go into the mathematical details. An edge inside the tensor network is shared by only two tensors, it represents an index that can take $\mid \Omega \mid$ values. A tensor restricted on one edge is a linear form on the vector space $\mathbb{V} = \mathbb{R}^{|\Omega|}$. We define two tensors u and v such that :

$$u : \mathbb{V} \otimes \mathbb{V} \to \mathbb{V} \otimes \mathbb{V}, \qquad u^\dagger u = \mathbb{1}_{\mathbb{V}} \otimes \mathbb{1}_{\mathbb{V}}$$

$$v : \mathbb{V} \to \mathbb{V} \otimes \mathbb{V}, \qquad v^\dagger v = \mathbb{1}_{\mathbb{V}}$$

We call u a disentangler, inserting a tensor $u^\dagger u$ does not change the partition function but a proper choice is supposed to remove short-range correlations. We

call v an isometry, it is aimed at combining two indices into one, inserting $vv^\dagger$ induces a projection operation from a space of dimension $\mid \Omega \mid^2$ to a space of dimension $\mid \Omega \mid$ at the expense of a truncation error in the partition function. The notion of isometries was first introduced by Levin and Nave in [15] but disentanglers are what is really new in the Evenbly-Vidal approach. We now show graphically how to do the insertions, we use two different isometries v and w.

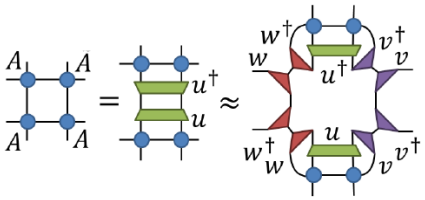


Figure 12: Insertion of disentanglers and isometries, taken from [16].

Proceeding to this local transformation, we realize the abstract scheme presented in figure 10, we then have to apply this step to the whole lattice. Evenbly and Vidal described the full scheme in figure 13. We can make several comments on the scheme. It is anisotropic in the construction of the disentanglers and isometries, but isotropic in the sense after a renormalization scheme, the dimension of the network has been reduced by the same amount in every direction, it raises the issue of whether a scheme that preserves symmetries is better. We also raise the matter of whether choosing a gauge for the tensor network can be useful to restore symmetries, or to be in an eigenvector base.

The delicate point now is to choose which disentangler and isometry to use. For that we come back to the idea of projective truncations and we try to minimize the Hilbert-Schmidt norm of the error tensor, defined in figure 14. The full algorithm designed for this problem in presented in [17]. The ideas presented in the discussion of section 4.2 could also be applied here to improve the choice of disentangler and isometry.

The renormalization scheme can then be iterated several times, inducing a flow in the space of tensors. The process will be stopped when we reach a fixed point in the space tensor up to a gauge transformation, and the partition function will then be the contraction of this single tensor.

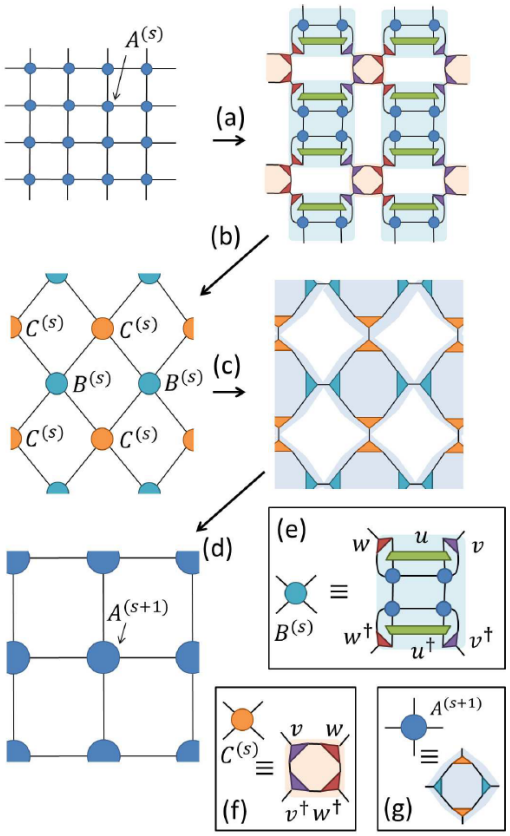


Figure 13: A full step of renormalization is presented here, from a tensor $A^{(s)}$ to a coarse grained tensor $A^{(s+1)}$. The operation (c) consists in doing an SVD of the intermediate tensor variables B and C. Taken from [16].

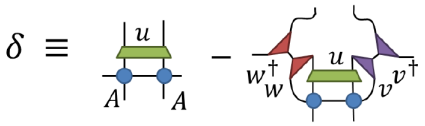


Figure 14: Definition of the truncation error tensor, taken from [16].

4.4 Local rescaling on a lattice

Very recently, it was found [18] that tensor network renormalization could be used to proceed to local rescaling on a lattice, and in particular conformal rescaling and could help to study vacancies inside a lattice.

The idea of implementing a scale transformation on a lattice comes from the early age of renormalization when Wilson's renormalization implied a global rescaling. As a renormalization is designed to drive the system to its critical point where it falls into the universality of scale invariance, rescaling should not in principle modify the critical properties of this system.

Where the recent result [18] brings a new insight lies on the notion of local rescaling. If a model presents a scale invariance at criticality, this invariance should be local.

Let us now introduce the procedure of local rescaling. We consider a square lattice with a translation invariance model. As proved within the construction of tensor network models, the partition function can be represented as the trace of a single tensor. As proved in paragraph 4.3, Tensor Network Renormalization allows to perform a global discrete rescaling :

$$(x, y) \longmapsto (\lambda x, \lambda y),$$

where (x,y) denotes the lattice site labels. Generally λ is equal to $1/2$ on square lattices, as in figure 11. Yet, the construction of the coarse grained tensor network is made through local tensor replacements. We can therefore think that by making inhomogeneous replacements we can have a local rescaling :

$$(x, y) \longmapsto (\lambda_1(x, y)\, x, \lambda_2(x, y)\, y)$$

Let us now analyze the example provided in [18]. We consider a square lattice with the 2D translational invariant Ising model. Firstly, we are going to map the square lattice into a cylinder. We choose the following coordinates transformation[23] :

$$(x, y) \longmapsto (s, \theta)$$

$$s = \ln \sqrt{x^2 + y^2}, \qquad \theta = 2 \arctan \frac{y}{x + \sqrt{x^2 + y^2}} \tag{26}$$

In complex notations, if $z = x + iy = re^{i\theta}$, then $z \longmapsto w = \ln z = s + i\theta$. Since this transformation leads to $\mathrm{d}x^2 + \mathrm{d}y^2 = e^{2s}(\mathrm{d}\theta^2 + \mathrm{d}s^2)$, we have to rescale our coordinates by :

$$\lambda(x, y) = \frac{1}{\sqrt{x^2 + y^2}} = e^{-s},$$

so that (s,θ) now parametrizes a cylinder of unit radius[24] as presented in figure 15. We can now see that the chosen origin of the lattice defines the lower boundary of the cylinder and the infinite boundary of the square lattice is now the upper boundary of the cylinder.

Let us briefly state why this representation might be useful. If we are at criticality, scale invariance transforms into translation invariance on the cylinder. In addition to that if we introduce one vacancy inside our lattice, at the origin for instance, thanks to the cylindrical mapping, this vacancy is now a boundary condition at $s = -\infty$[25].

[23] Actually the transformation chosen here differs from the one in [18], as the authors propose a scheme where $\theta = \arctan \frac{y}{x}$. This scheme is not bijective whereas (26) is bijective. Simple trigonometry proves that $\theta = \arg(x + iy)$.

[24] That way, we transform the metric into $\mathrm{d}x^2 + \mathrm{d}y^2 = \mathrm{d}\theta^2 + \mathrm{d}s^2$ which is the metric of a cylinder with unit radius.

[25] This might play a role when studying bulk-boundary correspondence.

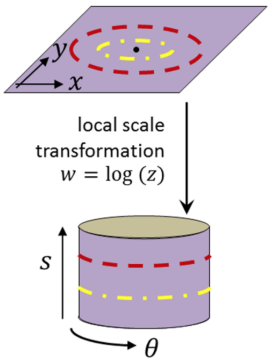


Figure 15: The Euclidean plane is mapped into a cylinder of unit radius, a translation $s \longmapsto s + a$ on the cylinder is equivalent to a rescaling $z \longmapsto e^a z$ on the Euclidean plane, taken from [18].

The second step of the calculation is to prove that Tensor Network Renormalization can actually create such a map. Following the conventions of section 4.3, we show graphically how to apply the Evenbly-Vidal scheme to proceed to local rescaling in figure 16. A block of 2×2 tensor is first removed from the lattice, we then choose the origin of the lattice to be at the center of the puncture and finally the Evenbly-Vidal scheme is applied to the remaining network. A first iteration of the coarse-graining transformation rescales the whole lattice by a factor $1/2$ except in a close neighborhood of the puncture where the scale is not changed. The new network conserves the property of having a puncture at the origin consisting of a block of 2×2 tensors and in addition to that two concentric rows of tensors appear. We can now reapply the scheme to this network, there will be again a $1/2$ global rescaling except in a close neighborhood of the new puncture.

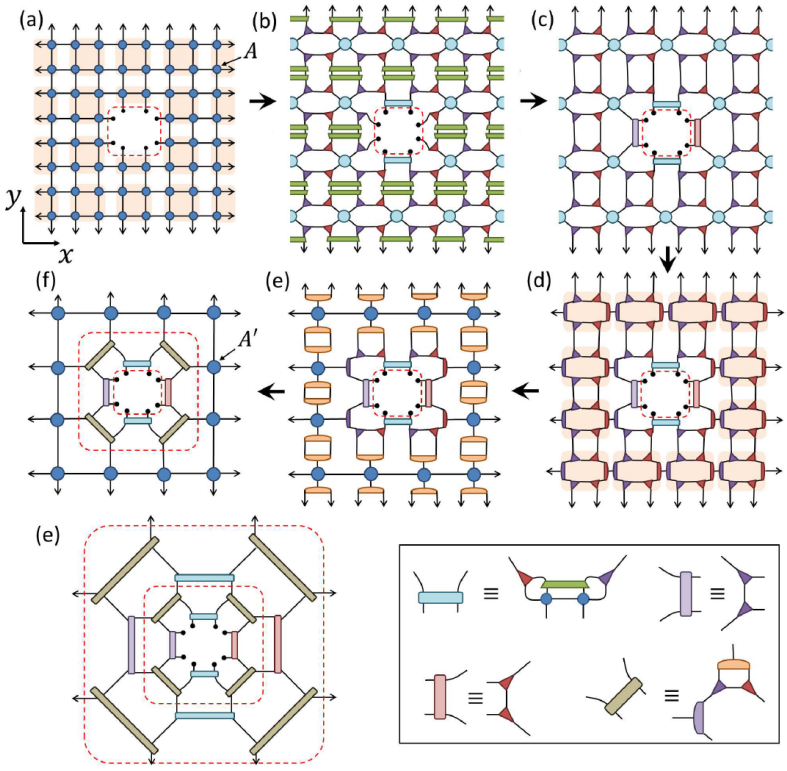


Figure 16: Tensor network renormalization scheme to map a planar lattice to a cylindrical lattice. Figure (f) is obtained after one iteration and figure (e) after two iterations, taken from [18].

By induction, we see that after a large number of iterations, any region located at a distance $\sqrt{x^2+y^2} \simeq 2^t$ of the initial puncture has been rescaled by a factor $(1/2)^t \simeq 1/\sqrt{x^2+y^2}$ [26]. This allows us to use our cylindrical representation of the tensor network, as presented in figure 17. We then restore the four initial tensor missing at the origin at the lattice. We define the concentric row of tensors appearing at each renormalization step $\mathcal{R}_t$, t being the step at which the row is introduced. Graphically we show that the partition function of the system is now :

$$\mathcal{Z} = \langle v_0 \mid \mathcal{R}_{\tilde{t}} ... \mathcal{R}_1 \mathcal{R}_0 \mid \mathbb{1} \rangle,$$

where $\tilde{t}$ accounts for the finite size of the lattice. The row of tensors allows to implement of change of scale in the lattice in the same way that a translation in the cylinder accounts for a rescaling in the plane. The first row $\mathcal{R}_0$ applied on the 2×2 initial tensors yields the first coarse-grained tensors and we get the notion of scale changing recursively. The concentric rows have the semi-group property, which gives us the knowledge of their spectral property. At

[26] In a discrete lattice like the square lattice, we could change the topology used to account for symmetries, here for instance the L^2-norm is not well adapted to the problem, whereas the L^1-norm or the L^∞-norm would account for the square geometry.

criticality, we expect the concentric tensors not to depend on the scale at which they are applied and therefore finding the partition function of the model is just transformed into a problem of finding a base of eigenvectors and the related eigenvalues of the concentric rows. This concludes this section where we have shown that tensor network renormalization can also proceed to local rescaling on a lattice, allowing to capture new features of scale invariance at criticality.

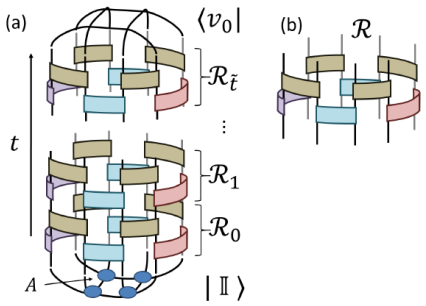


Figure 17: Cylindrical representation of the tensor network. At one boundary of the cylinder lie four initial tensors that are denoted by the ket $|\mathbb{1}\rangle$, at the other boundary lie a set of edges accounting for the boundary conditions - periodic for instance - denoted by the bra $\langle v_0|$, taken from [18].

4.5 Outlooks

This final section aims at suggesting further research direction. Up to our knowledge, all the models investigated within the tools of tensor network renormalization are purely deterministic without randomness. Models with impurities, disorders, random couplings or random external fields carry a great interest and the framework of tensor network renormalization could help to improve the understanding we have of them. Conversely, Random Matrix Theory has been extensively developed and applied to physics, for instance in [19], an adaptation of some results to a generalized Random Tensor Theory might be the key to solve modern challenges in condensed matter physics.

For 1-dimensional systems, the partition function in a matrix network model is fully determined by the spectral properties of the matrices. It would be interesting to investigate whether these properties can be extended to tensor network models in higher dimension. Some particularly interesting properties are the non-degeneracy of eigenvalues and their analyticity. The smoothness of an eigenvalue, seen as a function of the temperature, determines if a phase transition occurs.

Extending the work of this essay towards models with three-body interactions could also help to develop new results in tensor networks. Besides that, the tensor networks models we built were based on discrete lattice physics. This physics can sometimes be seen as a discretization of a continuous theory. It would also be interesting to investigate whether we can map back the discrete tensor models to continuous models.

References

[1] Kenneth G Wilson. The renormalization group: Critical phenomena and the kondo problem. *Reviews of Modern Physics*, 47(4):773, 1975.

[2] Leo P Kadanoff. *Statistical physics: statics, dynamics and renormalization*. World Scientific, 2000.

[3] Geoffrey Grimmett. *Probability on graphs: random processes on graphs and lattices*, volume 1. Cambridge University Press, 2010.

[4] Pankaj Mehta and David J Schwab. An exact mapping between the variational renormalization group and deep learning. *arXiv preprint arXiv:1410.3831*, 2014.

[5] Thomas C Schelling. Dynamic models of segregation†. *Journal of mathematical sociology*, 1(2):143–186, 1971.

[6] Efi Efrati, Zhe Wang, Amy Kolan, and Leo P Kadanoff. Real-space renormalization in statistical mechanics. *Reviews of Modern Physics*, 86(2):647, 2014.

[7] Antoines Georges, Marc Mézard, and Rémi Monasson. Théorie statistique des champs : des polymères aux transitions de phase, 2009.

[8] Pierre-Gilles de Gennes. Exponents for the excluded volume problem as derived by the wilson method. *Physics Letters A*, 38(5):339–340, 1972.

[9] Hugo Duminil-Copin and Stanislav Smirnov. The connective constant of the honeycomb lattice equals $\sqrt{2+\sqrt{2}}$. *Annals of Mathematics*, 175(3):1653–1665, 2012.

[10] HH Zhao, ZY Xie, QN Chen, ZC Wei, JW Cai, and T Xiang. Renormalization of tensor-network states. *Physical Review B*, 81(17):174411, 2010.

[11] José A Cuesta and Angel Sánchez. General non-existence theorem for phase transitions in one-dimensional systems with short range interactions, and physical examples of such transitions. *Journal of statistical physics*, 115(3-4):869–893, 2004.

[12] Amos Altshuler. Construction and enumeration of regular maps on the torus. *Discrete Mathematics*, 4(3):201–217, 1973.

[13] I-Ling Chang and Jiu-Wen Chou. A molecular analysis of carbon nanotori formation. *Journal of Applied Physics*, 112(6):063523, 2012.

[14] Jie Liu, Hongjie Dai, Jason H Hafner, Daniel T Colbert, and Richard E Smalley. Fullerene'crop circles'. *Nature*, 385:780–781, 1997.

[15] Michael Levin and Cody P Nave. Tensor renormalization group approach to two-dimensional classical lattice models. *Physical Review Letters*, 99(12):120601, 2007.

[16] Glen Evenbly and Guifre Vidal. Tensor network renormalization. *arXiv preprint arXiv:1412.0732*, 2014.

[17] Glen Evenbly. Algorithms for tensor network renormalization. *arXiv preprint arXiv:1509.07484*, 2015.

[18] G. Evenbly and G. Vidal. Local scale transformations on the lattice with tensor network renormalization. *Phys. Rev. Lett.*, 116:040401, Jan 2016.

[19] David S Dean, Pierre Le Doussal, Satya N Majumdar, and Grégory Schehr. Universal ground-state properties of free fermions in a d-dimensional trap. *EPL (Europhysics Letters)*, 112(6):60001, 2015.