

Chapter 1

Introduction to Matrix Product States

A. About this Module

This is really supposed to be a practical module: We will develop the tools to use tensor networks to solve many-body problems. There are 6 "lectures". Bring your laptop – we will spend part of each lecture coding. Part of your prep for each lecture will be installing and setting up appropriate software. The bulk of the module will focus on 1D, and the technique known as "Density Matrix Renormalization Group" (DMRG). You will write your own DMRG code from scratch. Each lecture will have some reading associated with it – not too much – at most an hour of your time. There will be also homeworks. I am not going to bother grading the homeworks. To get credit for the module you need to complete the activities listed as "for credit." You are probably smart enough to know if it is worth your while to complete the other activities.

B. Reading for Friday

Please read sections 1-4 of Roman Orus' paper "A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States," arXiv:1306.2164. This is pretty easy reading – no equations. It is mostly motivation and background. By reading this, you will save me to have to give a big song and dance about how interesting/important things are, and instead focus on the mechanics. It also reinforces some of the notation we will be using

C. Other Prep for Friday

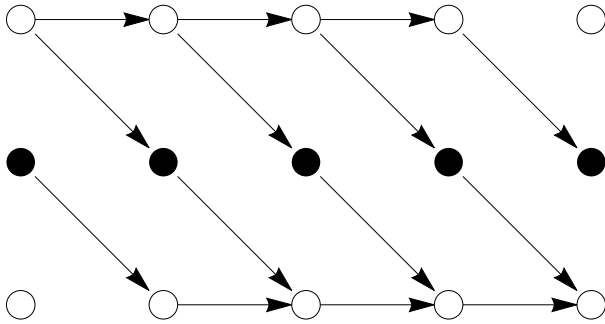
We will use Mathematica for some of the numerical activities. A 1 year student license for Mathematica is \$24 – do a web search for "Mathematica student license Cornell." Alternatively you can buy a copy for \$161. Please install it and bring your laptop. I won't go through the basics of using Mathematica – but you should pick up the syntax from the activities.

D. Thinking locally

Consider a single particle hopping around on a 1D lattice. For concreteness, imagine there are 5 sites. To specify the state of the system, you need five numbers: $\psi_1, \psi_2, \psi_3, \psi_4, \psi_5$, where $|\psi_j|^2$ is the probability for the particle to be on site j . Now consider the two-particle problem. You now need 5^2 complex numbers: ψ_{ij} . The Hilbert space grows quite quickly. Even if you imagine our particles are hard-core bosons – meaning you can't put more than one particle on any given site, the Hilbert space is large. Our two-particle state now needs $5 * 4/2$ numbers ($\psi_{ii} = 0$ and $\psi_{ij} = \psi_{ji}$).

The entire space of possible many-body states is spanned by the 2^5 states where each site is designated as empty or occupied. One way to solve the many-body problem is to write the Hamiltonian as a $2^5 \times 2^5$ matrix on this space, and then diagonalize it. Clearly that would be expensive for larger systems. We can improve things by noting that this $2^5 \times 2^5$ matrix is block diagonal: only basis states with the same number of particles are connected.

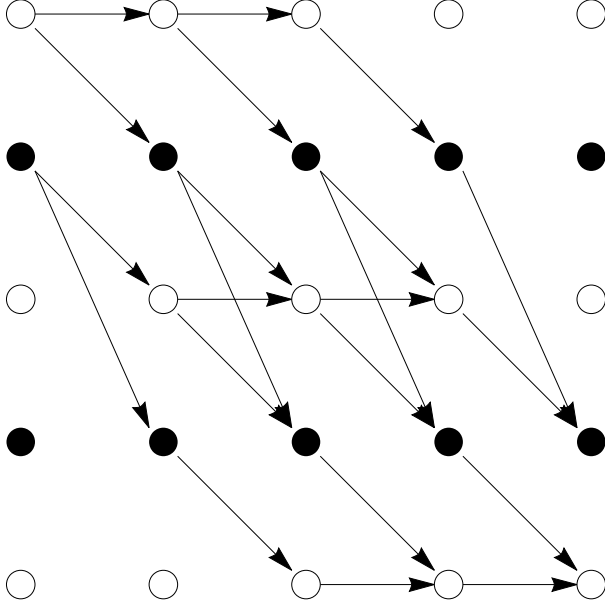
Let's algorithmically find the 1-particle basis states, starting at the left. This way of thinking will be the seed of our technique: Site 1 can be either occupied or empty. If it is empty then site 2 can be either occupied or empty. On the other hand, if site 1 was occupied, then site 2 must be empty. We get a decision tree:



If you count them, there are 5 routes through this network, corresponding to our original counting of 5 states.

Come up with the equivalent construction for 2 particle states.

You should have come up with something like:



This time if you count the number of paths you should get 10 – the same as our more conventional approach. The generalization to more sites and more particles should be obvious. You should be able to easily turn this into a computer algorithm for constructing n-particle states.

I am now going to use this structure to write the most general 1-particle state in a somewhat unusual way, namely

$$|\psi\rangle = \begin{pmatrix} |0\rangle_1 & \psi_1|1\rangle_1 \end{pmatrix} \begin{pmatrix} |0\rangle_2 & \psi_2|1\rangle_2 \\ 0 & |0\rangle_2 \end{pmatrix} \begin{pmatrix} |0\rangle_3 & \psi_3|1\rangle_3 \\ 0 & |0\rangle_3 \end{pmatrix} \begin{pmatrix} |0\rangle_4 & \psi_4|1\rangle_4 \\ 0 & |0\rangle_4 \end{pmatrix} \begin{pmatrix} \psi_5|1\rangle_5 \\ |0\rangle_5 \end{pmatrix} \quad (1.1)$$

The way to read this is that the upper row of the matrix corresponds to paths where no previous sites have particles on them, and the lower row to paths where a previous site has a particle. The left column corresponds to cases where there will be a particle to the right, and the right column to cases where there will not be. Multiplying out the matrices gives the standard definition of the ψ_j 's.

The analogous 2-particle construction is a bit more obscure,

$$|\psi\rangle = \begin{pmatrix} |0\rangle_1 & |1\rangle_1 \end{pmatrix} \begin{pmatrix} |0\rangle_2 & |1\rangle_2 & & \\ & |0\rangle & \psi_{12}|1\rangle_2 & \\ & & & \end{pmatrix} \begin{pmatrix} |0\rangle_3 & |1\rangle_3 & & \\ & & |0\rangle_3 & \psi_{23}|1\rangle_3 \\ & & & |0\rangle_3 & \psi_{13}|1\rangle_3 \\ & & & & |0\rangle_3 \end{pmatrix} \\ \times \begin{pmatrix} |1\rangle_4 & & & \\ & |0\rangle_4 & & \psi_{34}|1\rangle_4 \\ & & |0\rangle_4 & \psi_{24}|1\rangle_4 \\ & & & |0\rangle_4 & \psi_{14}|1\rangle_4 \\ & & & & |0\rangle_4 \end{pmatrix} \begin{pmatrix} \psi_{45}|1\rangle_5 \\ \psi_{35}|1\rangle_5 \\ \psi_{25}|1\rangle_5 \\ \psi_{15}|1\rangle_5 \\ |0\rangle_5 \end{pmatrix}$$

Nonetheless, if you multiply out these matrices, you will get the most general 2-particle state.

These are examples of what are known as “Matrix Product States” – the simplest of the tensor network states. For a broad class of 1D systems, the many-body states are efficiently written in this language.

E. Spin Models

As may be clear to you, our model of particles hopping on a lattice is equivalent to a spin model: We can take the presence of a particle to represent an up-spin, and the absence a down-spin. Much of the machinery we will be talking about was developed in the context of spin chains. If you are not familiar with spin chains, be sure to do the homework problems.

From classical spin systems, you should be familiar with a few canonical states of spin chains, such as an idealized ferromagnetic state $|\uparrow\uparrow\cdots\uparrow\rangle$ or an idealized antiferromagnet $|\uparrow\downarrow\uparrow\downarrow\cdots\rangle$. These are both product states – and have infinite correlation lengths. Matrix product states are one way to produce paramagnets with a finite correlation length. For example, consider

$$|\psi\rangle = \cdots \begin{pmatrix} |\uparrow\rangle & \epsilon|\downarrow\rangle \\ \epsilon|\uparrow\rangle & |\downarrow\rangle \end{pmatrix} \begin{pmatrix} |\uparrow\rangle & \epsilon|\downarrow\rangle \\ \epsilon|\uparrow\rangle & |\downarrow\rangle \end{pmatrix} \begin{pmatrix} |\uparrow\rangle & \epsilon|\downarrow\rangle \\ \epsilon|\uparrow\rangle & |\downarrow\rangle \end{pmatrix} \cdots \quad (1.2)$$

If I multiply out these matrices I get a state in conventional notation. This state has the property that if $|\epsilon| < 1$, then neighboring spins tend to point in the same direction, but spins that are far apart will be uncorrelated. Essentially ϵ penalizes domain walls. Our goal for most of the remainder of the class will be to figure out how to calculate the properties of wavefunctions like this.

F. Notation

We are going to be doing some intense tensor manipulations, and life will be much easier if we first develop some graphical notation. I will largely follow Orus’ notation, but will add a couple minor tweaks.

There are a few approaches to linear algebra. The typical physicist approach is that a vector is a real tangible quantity, and the components of a vector are a set of numbers which you get by projecting that vector onto an orthonormal basis. In the language of quantum mechanics

$$\psi_j = \langle j|\psi\rangle. \quad (1.3)$$

Computers are not good with abstract concepts, so the thing we are going to be manipulating are the components ψ_j . If we want to maintain the power we had with our abstract representation, we should store not only the list of numbers ψ_j , but also what the basis is. For concreteness, suppose I have two spin 1/2 objects. The basis for each is $|\uparrow\rangle_1, |\downarrow\rangle_1$ and $|\uparrow\rangle_2, |\downarrow\rangle_2$. The full Hilbert space could be written $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$. If I want to store the wavefunction, I would store the four numbers

$$\psi_{\sigma_1, \sigma_2} = \langle \sigma_1 \sigma_2 | \psi \rangle. \quad (1.4)$$

We will graphically denote these four numbers by a diagram such as

$$\langle \sigma_1 \sigma_2 | \psi \rangle = \psi_{\sigma_1, \sigma_2} = \begin{array}{c} \boxed{\psi} \\ \uparrow \quad \uparrow \\ \sigma_1 \quad \sigma_2 \end{array} \quad (1.5)$$

The two legs correspond to the indices. They are labeled so that you know what the basis is, and which particle is which. The arrows tell us that this is the components of a “bra.” If I were to draw the equivalent “ket,” I would instead draw

$$\langle \psi | \sigma_1 \sigma_2 \rangle = \psi_{\sigma_1, \sigma_2}^* = \begin{array}{c} \sigma_1 \quad \sigma_2 \\ \uparrow \quad \uparrow \\ \boxed{\psi^*} \end{array} \quad (1.6)$$

In this notation, I am free to move around where the legs are – the only significant thing are their label, and the direction of the arrow. The symbols and arrows tell us which contractions are allowed. You can only find the overlap between states that act on the same spins. Overlaps are made by connecting lines with the same indices, in the correct directions. For example, the norm of our state is

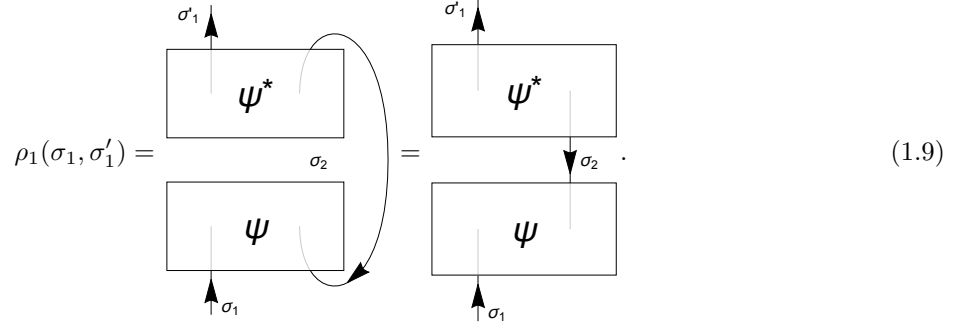
$$\langle \psi | \psi \rangle = \sum_{\sigma_1 \sigma_2} \langle \psi | \sigma_1 \sigma_2 \rangle \langle \sigma_1 \sigma_2 | \psi \rangle = \sum_{\sigma_1 \sigma_2} \psi_{\sigma_1, \sigma_2}^* \psi_{\sigma_1, \sigma_2} = \begin{array}{c} \boxed{\psi} \\ \uparrow \quad \uparrow \\ \sigma_1 \quad \sigma_2 \\ \downarrow \quad \downarrow \\ \boxed{\psi^*} \end{array} \quad (1.7)$$

This is just good solid notation. In addition to inner products, we can do outer products, for example, making the density matrix for this state is the rank 4 tensor

$$\rho_{\sigma'_1 \sigma'_2}^{\sigma_1 \sigma_2} = \langle \sigma_1 \sigma_2 | \psi \rangle \langle \psi | \sigma'_1 \sigma'_2 \rangle = \begin{array}{c} \sigma_1 \quad \sigma_2 \\ \uparrow \quad \uparrow \\ \boxed{\psi^*} \\ \downarrow \quad \downarrow \\ \boxed{\psi} \\ \uparrow \quad \uparrow \\ \sigma_1 \quad \sigma_2 \end{array} \quad (1.8)$$

This notation is particularly convenient for taking partial traces. For example, I can represent the reduced

density matrix for the first site by the diagram

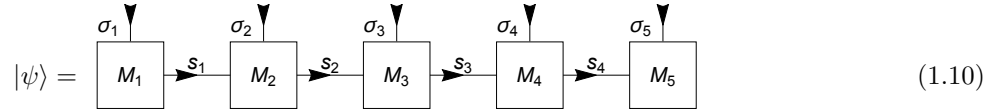


$$\rho_1(\sigma_1, \sigma'_1) = \begin{array}{c} \begin{array}{|c|} \hline \psi^* \\ \hline \end{array} \\ \begin{array}{|c|} \hline \psi \\ \hline \end{array} \end{array} = \begin{array}{c} \begin{array}{|c|} \hline \psi^* \\ \hline \end{array} \\ \begin{array}{|c|} \hline \psi \\ \hline \end{array} \end{array} \quad (1.9)$$

When everything is completely unambiguous, we can leave off the labels – as with most notation, knowing when to skip details is important.

G. Graphical structure of finite Matrix Product States

In our graphical language, a matrix product state on 5 sites is



$$|\psi\rangle = \begin{array}{c} \sigma_1 \downarrow \\ \boxed{M_1} \end{array} \xrightarrow{s_1} \begin{array}{c} \sigma_2 \downarrow \\ \boxed{M_2} \end{array} \xrightarrow{s_2} \begin{array}{c} \sigma_3 \downarrow \\ \boxed{M_3} \end{array} \xrightarrow{s_3} \begin{array}{c} \sigma_4 \downarrow \\ \boxed{M_4} \end{array} \xrightarrow{s_4} \begin{array}{c} \sigma_5 \downarrow \\ \boxed{M_5} \end{array} \quad (1.10)$$

In this "open boundary" MPS there are two rank 2 tensors and three rank 3 tensors. The indices s_j are referred to as "link indices."

For the wavefunction in Eq. (1.1), write explicit expressions for M_1, \dots, M_5 .

Obviously

$$M_1 = \left(\begin{array}{cc} |0\rangle_1 & \psi_1 |1\rangle_1 \end{array} \right), \quad (1.11)$$

but what does that mean? The most pedantic approach is to simply enumerate the matrix elements. The index σ_1 can take on two values: 0 or 1. The index s_1 can also take on two values, lets call them a and b , to keep things transparent. Therefore there are four numbers used to specify M_1 . Only two are non-zero, namely:

$$M_1^{0a} = 1 \quad (1.12)$$

$$M_1^{1b} = \psi_1. \quad (1.13)$$

Another notation which I will use is to keep track of which index is set to which value by using a script, so writing

$$M_1^{\sigma_1 s_1^a} = 1 \quad (1.14)$$

$$M_1^{\sigma_1 s_1^b} = \psi_1. \quad (1.15)$$

Note, if I use this notation, I could use 0, 1 in place of a, b without introducing any ambiguity. Of course the proliferation of indices can also be confusing.

Regardless, we can move on to the next matrix

$$M_2 = \begin{pmatrix} |0\rangle_2 & \psi_2|1\rangle_2 \\ 0 & |0\rangle \end{pmatrix}. \quad (1.16)$$

This is a rank-3 tensor – I have to specify $\sigma_2 = 0, 1$, $s_1 = a, b$, and $s_2 = c, d$. The non-zero entries are

$$M_2^{0ac} = 1 \quad (1.17)$$

$$M_2^{0bd} = 1 \quad (1.18)$$

$$M_2^{1ad} = \psi_2. \quad (1.19)$$

Suppose we want to calculate the norm of a matrix product state. That would be something like

$$\langle \psi | \psi \rangle = \begin{array}{ccccccccc} & \boxed{M^*_1} & \xleftarrow{t_1} & \boxed{M^*_2} & \xleftarrow{t_2} & \boxed{M^*_3} & \xleftarrow{t_3} & \boxed{M^*_4} & \xleftarrow{t_4} & \boxed{M^*_5} \\ \sigma_1 \downarrow & & & \sigma_2 \downarrow & & \sigma_3 \downarrow & & \sigma_4 \downarrow & & \sigma_5 \downarrow \\ \boxed{M_1} & \xrightarrow{s_1} & \boxed{M_2} & \xrightarrow{s_2} & \boxed{M_3} & \xrightarrow{s_3} & \boxed{M_4} & \xrightarrow{s_4} & \boxed{M_5} \end{array} \quad (1.20)$$

This is a scalar, because all of the indices are contracted. In principle the contraction is straightforward, but there are some subtleties that we will discuss next day, regarding numerical efficiency.

H. Graphical structure of Infinite MPS

Lets now return to our paramagnet:

$$|\psi\rangle = \dots \begin{pmatrix} |\uparrow\rangle & \epsilon|\downarrow\rangle \\ \epsilon|\uparrow\rangle & |\downarrow\rangle \end{pmatrix} \begin{pmatrix} |\uparrow\rangle & \epsilon|\downarrow\rangle \\ \epsilon|\uparrow\rangle & |\downarrow\rangle \end{pmatrix} \begin{pmatrix} |\uparrow\rangle & \epsilon|\downarrow\rangle \\ \epsilon|\uparrow\rangle & |\downarrow\rangle \end{pmatrix} \dots \quad (1.21)$$

Lets take the thermodynamic limit, and imagine we have an infinite number of spins.

Find the norm of this state.

The trick is to first construct the “transfer matrix”

$$E_I = \begin{array}{ccc} \xrightarrow{s'} & \boxed{M} & \xrightarrow{s} \\ & \uparrow & \\ \xrightarrow{t'} & \boxed{M^*} & \xrightarrow{t} \end{array} \quad (1.22)$$

This is a rank 4 tensor, each index taking on one of two values. It is however, natural to group the left two indices and the right two indices to write this as a 4×4 matrix. That is, we create a matrix whose rows correspond to $st = 00, 01, 10, 11$. One usually just does this in one's head, but one can also formally add graphical “combiners”

$$E_I = \begin{array}{c} \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \end{array} \\ \text{---} \text{---} \text{---} \end{array} \quad (1.23)$$

where the rank 3 tensors have non-zero elements:

$$\Gamma_{v_{00}}^{s_0 t_0} = \Gamma_{v_{10}}^{s_1 t_0} = \Gamma_{v_{01}}^{s_0 t_1} = \Gamma_{v_{11}}^{s_1 t_1} \quad (1.24)$$

It is just a trivial relabeling of two indices as one composite index. In fact, when we write computer programs, we will make functions which exactly do that.

The reason for calculating E_I is that if we use periodic boundary conditions the norm is just $\langle \psi | \psi \rangle = \text{Tr}(E_I)^N$. That means if λ is the largest eigenvalue of E_I , the norm is $\langle \psi | \psi \rangle = \lambda^N$. The state can then be normalized by dividing each M by $1/\sqrt{\lambda}$.

I. HW 1 – Due : Jan 29

Problem 1. (For Credit) Consider the “spin singlet”: $\uparrow\downarrow - \downarrow\uparrow$. Write this as a matrix product state. Hint the matrices are 1×2 and 2×1 (so I guess this could be called a “vector product state.”)

Problem 2. (For Credit) The “GHZ” or “cat” state of five spins is $|\uparrow\uparrow\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\downarrow\downarrow\rangle$. Write this as a matrix product state.

Problem 3. (For Credit) The “W”-state is $|\uparrow\downarrow\downarrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\downarrow\downarrow\uparrow\rangle$, Write this as a matrix product state. Hint: It is the same as what we used for the single-particle state.

Problem 4. (For Credit) The ferromagnetic 1D transverse field Ising Model is a spin model, defined by a Hamiltonian

$$H = \sum_j [-J\sigma_z^j\sigma_z^{j+1} - h\sigma_x^j]. \quad (1.28)$$

Here σ_z and σ_x are the regular Pauli matrices. We can get a simple understanding of how this model works through a variational calculation. The simplest variational wavefunction we can use is a product: $|\psi\rangle = \otimes_j |\psi_j\rangle$ – where $|\psi_j\rangle$ is a two-component spinor. In the \hat{z} basis we can parametrize $|\psi_j\rangle$ as

$$|\psi_j\rangle = \cos(\theta/2)|\uparrow\rangle + \sin(\theta/2)|\downarrow\rangle, \quad (1.29)$$

so that

$$\langle\psi_j|\sigma_z^j|\psi_j\rangle = \cos(\theta) \quad (1.30)$$

$$\langle\psi_j|\sigma_x^j|\psi_j\rangle = \sin(\theta). \quad (1.31)$$

4.1. Show that up to boundary terms (which are irrelevant in the thermodynamic limit),

$$\langle\psi|H|\psi\rangle = -JN \cos^2\theta - hN \sin(\theta), \quad (1.32)$$

where N is the total number of sites.

Problem 4. cont...

4.2. Write $x = \sin(\theta)$. Minimize $\langle\psi|H|\psi\rangle$ with respect to x (with the constraint that $-1 < x < 1$). Make a plot of the magnetization $m = \langle\sigma_z^j\rangle = \sqrt{1-x^2}$ as a function of the ratio h/J .

You should see two phases – a “ferromagnetic” phase where $m \neq 0$, and a “paramagnetic” one where $m = 0$. This is the simplest example of what is referred to as a “quantum phase transition.” Note, the mean-field theory over-estimates the stability of the ordered phase, so you should not take the numbers too seriously.

Problem 5. Challenge – not for Credit

If I have N hard core bosons on M sites, and use the algorithm in Sec. D for parameterizing the states, how big is the m 'th matrix? As we will see later, this is related to the entanglement entropy of a generic state. Hint: the m 'th matrix in this product has d_1 rows and d_2 columns. Express d_1 and d_2 as sums over binomial coefficients. Separately consider the cases $m < N$ and $m > N$.

For $m < N$ you should be able to do the sums. For $m > N$ there is no closed form.