

Advanced Computational Physics

**Time Evolution of  
one-dimensional spin systems,  
using Matrix Product States**

SS 2015

H. G. EVERTZ

Commented version, 19 Dec 2022

# 1 Matrix Product States (MPS)

DMRG Density Matrix Renormalization Group

The state space of many-particle models grows exponentially with the number of particles involved. Such models can therefore be treated exactly only for very small systems.

In recent years, a new representation for the states of one-dimensional systems has been developed, so-called Matrix Product States (MPS), which permit very precise and efficient approximations even for very large systems. They have allowed for, e.g., calculations of the ground state energy of Heisenberg models on several hundred sites to 10 digit precision. Recently, the non-equilibrium time evolution of such models has also become accessible.

This chapter provides a first introduction and examples for MPS.

## 1.1 Heisenberg spin chain and equivalent particle model

Throughout we shall treat the one-dimensional spin  $\frac{1}{2}$  Heisenberg model with open boundary conditions:

$$\hat{H} |\uparrow\downarrow\rangle = J_{xy} |\downarrow\uparrow\rangle - J_z |\uparrow\downarrow\rangle \quad \hat{H} = \sum_{j=1}^{L-1} J_{xy} (\hat{S}_j^x \hat{S}_{j+1}^x + \hat{S}_j^y \hat{S}_{j+1}^y) + J_z \hat{S}_j^z \hat{S}_{j+1}^z \quad (1.1)$$

alleine: Ising

$$H_{up\_i up\_i+1} = J_z (\hbar/2) up\_i (\hbar/2) up\_j+1 \\ H_{up\_i down\_i+1} = (J_{xy}/2) down\_i up\_i+1 - J_z (\hbar/2)^2 up\_i down\_i+1 = \sum_{j=1}^{L-1} \frac{1}{2} J_{xy} (\hat{S}_j^+ \hat{S}_{j+1}^- + \hat{S}_j^- \hat{S}_{j+1}^+) + J_z \hat{S}_j^z \hat{S}_{j+1}^z \quad (1.2)$$

Sup in z-Basis:  $(\begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix})$

$\uparrow_{up\_j down\_j+1} \leftrightarrow \downarrow_{down\_j up\_j+1}$  --> Faktor plusminus  $(\hbar/2)^2$

with  $\hat{S}_j^\pm = \hat{S}_j^x \pm i \hat{S}_j^y$ , thus  $\hat{S}_j^+ |\downarrow\rangle_j = |\uparrow\rangle_j$  and  $\hat{S}_j^- |\uparrow\rangle_j = |\downarrow\rangle_j$ . The representation of  $S_j^\alpha$  in the z-Basis at site  $j$  is  $\frac{1}{2}\sigma^\alpha$  (we leave out  $\hbar$ ). Site  $L$  and site 1 do not interact, i.e. we have open boundary conditions.

$$Sup/Up = 0$$

A chain of  $L$  sites has  $2^L$  basis states, which are  $|s_1, s_2, \dots, s_L\rangle$  with  $s_j = \uparrow, \downarrow$  in the z-basis.

In case of isotropic couplings  $J_{xy} = J_z =: J$  we get  $\hat{H} = J \sum_j \hat{S}_j \hat{S}_{j+1}$ . This model is also a good approximation of the strongly repulsive Hubbard model ( $U \gg t$ ) at half filling.

On a one-dimensional chain, the Heisenberg model is equivalent to a model of "spinless fermions", if one identifies a down-spin at site  $j$  with an empty site, and an up spin with an occupied site:

$$|0\rangle_j = |\downarrow\rangle_j \quad \text{and} \quad |1\rangle_j = |\uparrow\rangle_j. \quad \text{C}_j \int S_j = |0\rangle$$

The Hamiltonian then becomes

$$\hat{H} = \sum_{j=1}^{L-1} t (c_j^\dagger c_{j+1} + c_j c_{j+1}^\dagger) + V \hat{n}_j \hat{n}_{j+1} - \frac{V}{2} (\hat{n}_j + \hat{n}_{j+1}) + \frac{1}{4} \quad (1.3)$$

$|01\rangle \leftrightarrow |10\rangle$

with hopping  $t = J_{xy}/2$ , occupation number operator  $\hat{n}_j = \hat{S}_j^z + \frac{1}{2}$ , and "Coulomb repulsion"  $V = J_z$ . When one has a fixed particle number, the last 2 terms in  $\hat{H}$  add up to a constant, i.e. they are just a shift of energy.

The models are equivalent since there is only next-nearest neighbor hopping, so that particles never switch places and thus the Pauli fermion-sign never appears.<sup>1</sup>  
We will occasionally switch between the spin-language and the particle language.

## 1.2 Matrix Product State ansatz

A general state of the Heisenberg chain is

$$|\psi\rangle = \sum_{s_1, s_2, \dots, s_L} c_{s_1, s_2, \dots, s_L} |s_1, s_2, \dots, s_L\rangle \quad (1.4)$$

with  $2^L$  complex numbers as coefficients.

We will now write the coefficients in a different way, as a product of matrices, with one matrix for every lattice-site. This can always be done exactly (see later) when the matrices are chosen big enough, namely up to  $2^{L/2} \times 2^{L/2}$ . We will later see that much smaller matrices ( $O(100) \times O(100)$ ) already provide an extremely good approximation to physically relevant states.

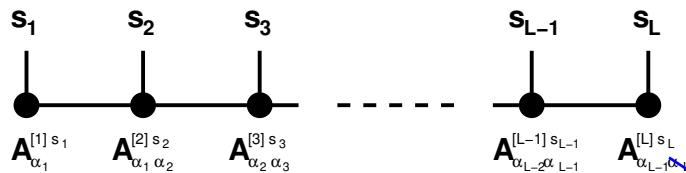
Ansatz:

s1 = up oder down

$$|\psi\rangle = \sum_{s_1, s_2, \dots, s_L} \sum_{\alpha_1, \alpha_2, \dots, \alpha_{L-1}} A_{\alpha_1}^{[1]s_1} A_{\alpha_1, \alpha_2}^{[2]s_2} A_{\alpha_2, \alpha_3}^{[3]s_3} \dots A_{\alpha_{L-2}, \alpha_{L-1}}^{[L-1]s_{L-1}} A_{\alpha_{L-1}}^{[L]s_L} |s_1, s_2, \dots, s_L\rangle. \quad (1.5)$$

- The  $A$  can be taken to be square matrices (except for the first and last  $A$ , which are vectors).
- Matrix size:  $\chi \times \chi$ , where  $\chi$  is some number.
- $\alpha_j = 1 \dots \chi$  are the matrix indices.
- The upper index  $[j]$  numbers the lattice sites. There can be different matrices  $A$  at each site.
- At each site  $j$ , there are two matrices,  $A^{[j]\uparrow}$  and  $A^{[j]\downarrow}$ , corresponding to the values  $s_j = \uparrow$  and " $s_j = \downarrow$ " in the basis vector  $|s_1, s_2, \dots, s_L\rangle$ .
- Note that for a given state, **the matrices  $A$  are not unique**: one can replace any pair of matrices  $A^{[j]} A^{[j+1]}$  by  $(A^{[j]} X) (X^{-1} A^{[j+1]})$ , with any invertible matrix  $X$ .

It is very helpful to denote this ansatz for the coefficients in a graphical way:



<sup>1</sup> A difference does however appear when one calculates fermionic observables like  $\langle c_i c_j^\dagger \rangle$ , which can be different from  $\langle \hat{S}_i^- \hat{S}_j^+ \rangle$ . Such observables will not appear in these lecture notes.

when i and j  
are not nearest  
neighbors

## 1.2.1 Examples

### • Single basis state (Product state)

A basis state like  $|\downarrow\downarrow\uparrow\downarrow\rangle$  is called "product state", since it can be written as a product  $|\downarrow\rangle_1 |\downarrow\rangle_2 |\uparrow\rangle_3 |\downarrow\rangle_4$  and it does not contain any linear combination.

$$\begin{aligned} |\psi\rangle &= | \downarrow \downarrow \uparrow \downarrow \uparrow \downarrow \rangle \\ j &= \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ A_j^\uparrow &= \begin{matrix} 0 & 0 & 1 & 0 & 1 & 0 \end{matrix} \\ A_j^\downarrow &= \begin{matrix} 1 & 1 & 0 & 1 & 0 & 1 \end{matrix} \end{aligned}$$

The "matrices"  $A$  are just single numbers here. (In the numerical project, we will do time evolutions starting from such simple product states).

### • Singlet, L=2 sites

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( | \uparrow \downarrow \rangle - | \downarrow \uparrow \rangle \right)$$

$$\begin{aligned} A_j^\uparrow &= (1, 0), \quad - \begin{pmatrix} 0 \\ 1 \end{pmatrix} / \sqrt{2} \\ A_j^\downarrow &= (0, 1), \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} / \sqrt{2} \end{aligned}$$

### • Nonlocal singlet

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}} \left( | \downarrow \downarrow \uparrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \rangle - | \downarrow \downarrow \downarrow \downarrow \uparrow \downarrow \downarrow \downarrow \downarrow \downarrow \rangle \right) \\ A_j^\uparrow &= 0, 0, (1, 0), \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, - \begin{pmatrix} 0 \\ 1 \end{pmatrix} / \sqrt{2}, 0, 0 \\ A_j^\downarrow &= 1, 1, (0, 1), \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} / \sqrt{2}, 1, 1 \end{aligned}$$

Notes:

- 1) The A-matrices of these three examples already satisfy the normalization condition eq (4.5) on page 12.
- 2) The row resp. column vectors can be moved to the left- resp. right-most site. See also examples on next page (3a)
- 3) All these A-matrices can be extended to any larger size (chi times chi), by taking the values given here in the upper left hand corners of the matrices and filling the rest of the matrices with zeros

Examples continued  
on page 3a !

## 2 Singular Value Decomposition (SVD)

The Singular Value Decomposition (SVD) is an extremely important and versatile tool from linear algebra. It is related to the familiar eigenvalue decomposition (see below). However, it is in general a different decomposition, and it is valid **for every real or complex matrix!** Unfortunately, the SVD is often unknown to physicists, whereas it is extremely widespread elsewhere and used for e.g. image processing, signal processing, optimizations, etc. For this reason, we will treat it in some detail.

**Further examples:** Let us examine the normalized state

$$|\psi\rangle = a|1000\rangle + b|1010\rangle + c|0010\rangle + d|1001\rangle \quad (\text{with } 0 \text{ for down-spin, } 1 \text{ for up-spin, and } |a|^2 + |b|^2 + |c|^2 + |d|^2 = 1)$$

(1) An immediate representation as an MPS with diagonal(!) matrices is:

(These A-matrices are already (left-)normalized, eq. (4.4):  
 $\sum_s A^s \text{dagger } A^s = 1$ )

$$A^\downarrow: (0111) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$$

$$A^\uparrow: (1000) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

(The "Nonlocal singlet" on the previous page  
can be written analogously)

o Alternatively, one can move the coefficients a,b,c,d into the column on the left.

Then the matrices are called "B" and are "right-normalized":  $\sum_s B^s \text{dagger } B^s = 1$

o Similarly, when one has L sites, the analogous sum of L basis states can be written with  $L \times L$  matrices A.

o In the same way, any sum of n basis states (with arbitrary numbers of spin up) can be written with  $n \times n$  matrices.

Thus a generic linear combination of all  $2^L$  basis states can immediately be written with exponentially large (but diagonal) matrices of sizes  $2^L \times 2^L$

o For convenience in programming, one can extend the row resp. column vector into a matrix, by filling up with zeroes.

Then the product of matrices for all sites is again a matrix, but it is all zero except for the(1,1) component, which is the desired coefficient.

(2) This same state  $|\psi\rangle$  can also be written with much smaller matrices of dimension  $2 \times 2$  (!) (similar for any number of terms in the sum)

$$A^\downarrow: (10) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A^\uparrow: (0a) \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & c \\ 0 & 0 \end{pmatrix} \begin{pmatrix} d \\ 0 \end{pmatrix}$$

(These matrices are not yet normalized.

Procedures for normalization and for reducing the matrix sizes: see comments on top of page 12, bottom of page 13.)

## 2.1 Basic SVD

Every real or complex  $n \times m$  matrix  $M$  can be decomposed like

$$M = U D V^\dagger \quad (2.1)$$

with a diagonal matrix  $D$  that contains only *positive real numbers* (or zero), which are called the *singular values* of  $M$ . It is of dimension  $N = \min(n, m)$ . Furthermore,

$$U^\dagger U = 1, \quad V^\dagger V = 1, \quad D = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}, \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_N = 0. \quad (2.2)$$

The number  $r$  of non-zero entries  $\lambda$  is called the *rank* of the matrix  $M$ .

- Case  $m \leq n$ : 
$$\boxed{M} = \boxed{U} \quad \boxed{D} \quad \boxed{V^\dagger}$$

- Case  $m \geq n$ : 
$$\boxed{M} = \boxed{U} \quad \boxed{D} \quad \boxed{V^\dagger}$$

- When  $M$  is real, then  $U$  and  $V^\dagger$  can also be chosen real.
- When  $M$  is real and quadratic, then  $U$  and  $V^\dagger$  are rotations (basis transformations), and  $D$  scales the directions in the intermediate basis.
- This version of the SVD, with (in general) non-square matrices  $U$  and  $V^\dagger$ , is called a "thin SVD".
- When two or more singular values are equal, then the SVD in this subspace is non-unique.
- In addition, for every direction  $j$ , one can multiply  $U$  by a phase factor  $\exp(i\varphi_j)$  and  $V^\dagger$  by  $\exp(-i\varphi_j)$  without changing  $M$ .
- The values  $\lambda_i$  are uniquely determined.
- The matrices  $U$  and  $V^\dagger$  are only determined up to phase factors:

$$\lambda_j = e^{i\varphi_j} \lambda_j e^{-i\varphi_j}$$

which can be moved into  $U$  and  $V^\dagger$ <sup>1</sup>

---

<sup>1</sup>When several  $\lambda_j$  are identical, the phase factors become unitary matrices.

## 2.2 SVD as a sum of outer products

The singular value decomposition  $M = UDV^\dagger$  can equivalently be written as

$$M_{ij} = \sum_k \lambda_k U_{ik} (V^\dagger)_{kj} = \sum_k \lambda_k u_i^{(k)} v_j^{(k)}. \quad (2.3)$$

Here  $u_i^{(k)} = U_{ik}$  and  $v_j^{(k)} = V_{ik}$ ,  $i = 1, \dots, m$ ;  $k = 1, \dots, N = \min(m, n)$  are the column vectors of  $U$  and  $V$  (not of  $V^\dagger$ ). The property  $U^\dagger U = V^\dagger V = \mathbb{1}$  implies that they are orthonormal. In vector notation, the SVD reads

$$M_{ij} = \sum_k \lambda_k \vec{u}^{(k)} \times \vec{v}^{(k)T} = \sum_k \left( \begin{array}{c} \vec{u}^{(k)} \\ \vec{v}^{(k)T} \end{array} \right) \times \left( \begin{array}{c} \vec{v}^{(k)} \\ \vec{u}^{(k)T} \end{array} \right) \quad (2.4)$$

This is a sum of *outer products* of the vectors  $\vec{u}^{(k)}$  and  $\vec{v}^{(k)}$ , weighted with  $\lambda_k$ .

The vectors  $\vec{u}^{(k)}$  and  $\vec{v}^{(k)}$  are called left and right *singular vectors* of  $M$ . The SVD together with  $V^\dagger V = \mathbb{1}$  implies

$$M \vec{v}^{(k)} = \lambda_k \vec{u}^{(k)} \quad \text{and} \quad M^\dagger \vec{u}^{(k)} = \lambda_k \vec{v}^{(k)}.$$

## 2.3 Connection to the eigenvalue decomposition

- An SVD exists for every matrix, with positive singular values, whereas an eigenvalue decomposition does not exist for every matrix, and eigenvalues need not be real or positive.
- When  $M$  is quadratic and all eigenvalues are  $\geq 0$ , then the eigenvalue decomposition  $M = UDU^\dagger$  is the same as the SVD.
- The SVD of an arbitrary matrix  $M$  provides the eigenvalue decomposition for the hermitian matrices  $M^\dagger M$  and  $MM^\dagger$ :

$$M^\dagger M = \underbrace{(UDV^\dagger)^\dagger}_{= UDV} \underbrace{UDV^\dagger}_{= VD^\dagger U^\dagger} = V D^\dagger U^\dagger UDV^\dagger = V \begin{pmatrix} |\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix} V^\dagger. \quad (\lambda_i |^2 = \lambda_i^2)$$

Similarly,

$$MM^\dagger = UDV^\dagger (UDV^\dagger)^\dagger = UDV^\dagger VD^\dagger U^\dagger = U \begin{pmatrix} |\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix} U^\dagger.$$

## 2.4 Representation of the SVD with square matrices

When  $M$  is not quadratic, then either  $U$  or  $V^\dagger$  is not quadratic in the SVD  $M = UDV^\dagger$ . Alternatively, one can write the SVD with *unitary* quadratic matrices  $\tilde{U}$  and  $\tilde{V}^\dagger$ :

$$M = UDV^\dagger = \tilde{U} \tilde{D} \tilde{V}^\dagger \quad (2.5)$$

This can be interpreted as a basis transformation by  $\tilde{V}^\dagger$ , a weighting of directions by  $\tilde{D}$ , and another basis transformation by  $\tilde{U}$ . When  $M$  is  $m \times n$  dimensional, then  $\tilde{U}$  is  $m \times m$ ,  $\tilde{D}$  is  $m \times n$ , and  $\tilde{V}^\dagger$  is  $n \times n$ .

- Case  $m \leq n$ :

$$\boxed{\begin{array}{c} M \\ = \end{array}} \quad \boxed{\begin{array}{c|c} U & D \mid 0 \ 0 \ 0 \end{array}} \quad \boxed{\begin{array}{c} V^\dagger \\ \hline (\text{rest}) \end{array}}$$

In this case,  $\tilde{U} = U$ . The lower rows of  $\tilde{V}^\dagger$  contain extra eigenvectors, beyond those in  $V^\dagger$ . They do not contribute to  $M$  because of the zeroes in  $\tilde{D}$ . Since the eigenvectors in  $\tilde{V}^\dagger$  are orthogonal, the application of  $M$  to such a vector gives zero, i.e. they belong to the *null space* of  $M$ . (The directions  $j$  beyond the rank  $r$ , with vanishing singular value  $\lambda_{j>r} = 0$  also belong to the null space). When considering the action of  $M$  on the full vector space, this null space can be ignored (see below).

- Case  $m \geq n$ :

$$\boxed{\begin{array}{c} M \\ = \end{array}} \quad \boxed{\begin{array}{c|c} U & (\text{rest}) \end{array}} \quad \boxed{\begin{array}{c} D \\ \hline 0 \end{array}} \quad \boxed{V^\dagger}$$

Now  $\tilde{V} = V$ .

## 2.5 Pseudoinverse

We first discuss the case of a square matrix  $M$ . Formally, the inverse is

$$M^{-1} = (\tilde{U}\tilde{D}\tilde{V}^\dagger)^{-1} = \tilde{V}\tilde{D}^{-1}\tilde{U}^\dagger, \quad (2.6)$$

since  $\tilde{U}$  and  $\tilde{V}$  are unitary. But the matrix  $M$  can contain singular values  $\lambda_{j>r} = 0$ . In these directions  $j$ ,  $M$  does not act, and the inverse  $\tilde{D}^{-1}$  would contain infinity.

It is much better to *exclude this null space* completely also from the inverse, i.e. to set  $\tilde{D}^{-1}$  to zero there. This is called the *pseudoinverse*:

$$\lambda_j \mapsto \frac{1}{\lambda_j}, \quad \text{but} \quad 0 \mapsto 0. \quad (2.7)$$

In practice, one maps singular values to zero when they are below some threshold (e.g.  $10^{-15}$ ). Using the pseudoinverse,  $M^{-1}M$  becomes

$$M^{-1}M = \tilde{V}\tilde{D}^{-1}\tilde{U}\tilde{U}^\dagger\tilde{D}\tilde{V}^\dagger \stackrel{\text{dagger}}{=} \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} = \dots = MM^{-1} \quad (2.8)$$

in which only (up to) the first  $r$  components correspond to the unity matrix, while the rest vanishes.

The same considerations apply when  $M$  is not square,  $n > m$  or  $m > n$ . Then  $M^{-1}M$  is an  $n \times n$  matrix, and  $MM^{-1}$  is  $m \times m$ . They are both of the form eq. 2.8.

## 2.6 Applications

The computational cost of the SVD is  $\mathcal{O}(\max(n, m)^3)$ , similar to a matrix diagonalization.

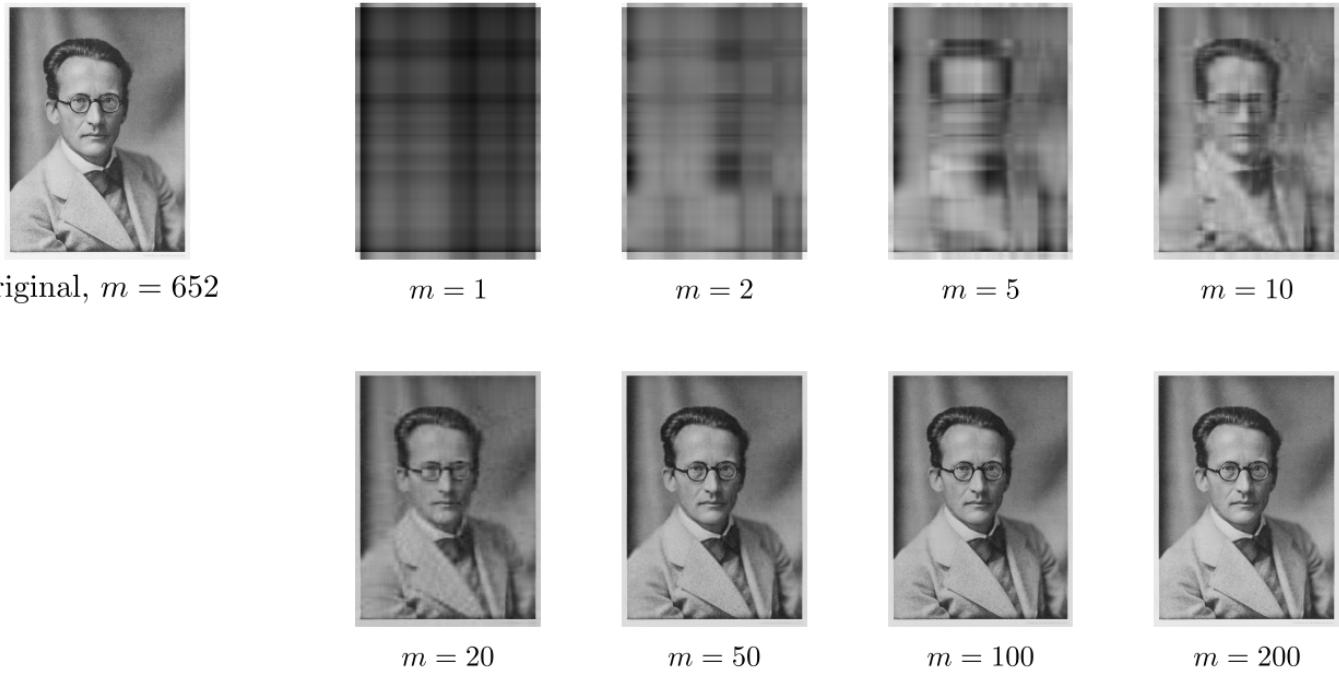


Figure I.B.8. Compressed images<sup>108</sup> of Erwin Schrödinger. Shown are various compressions where the original image was decomposed with an SVD into three matrices and multiplied back again.

Example of truncation by SVD; from

**Florian Maislinger,  
Non-Equilibrium Approaches to Strongly Interacting Many Body Systems**

Ph.D. thesis, TU Graz 2021.

Many applications of the SVD involve *truncation*: one replaces small singular values by zero. This provides an *approximation* to  $M$ , which is usually very good. Yet it often contains *far* fewer components, providing for much lower computational cost. Applications include *image processing*, signal processing, optimizations, strongly correlated many body systems, and many more.

## 3 Schmidt decomposition of a quantum state; Entanglement

### 3.1 Schmidt decomposition

Consider any quantum-mechanical system and two arbitrarily chosen subsystems  $A$  and  $B$ , for instance the left and right side of a one dimensional chain with some arbitrary split.

Let  $|j\rangle_A$  be the orthonormal basis states of subsystem  $A$ , and  $|k\rangle_B$  those of subsystem  $B$ . Then a general state of the total system is

(pure state)

$$|\Psi\rangle = \sum_{j,k} c_{jk} |j\rangle_A |k\rangle_B, \quad (3.1)$$

where  $c_{jk}$  are coefficients.

We now regard  $c_{jk}$  as a matrix and look at its singular value decomposition:

$$(c_{jk}) = \tilde{U} \tilde{D} \tilde{V}^\dagger, \quad \text{with } \tilde{U} \text{ and } \tilde{V} \text{ unitary.} \quad (3.2)$$

Written in matrix components, this becomes  $c_{jk} = \sum_{\alpha=1}^{\chi} \lambda_\alpha \tilde{U}_{j\alpha} (\tilde{V}_{\alpha k})^\dagger$ , where  $\chi$  is the rank of the matrix  $c_{jk}$ , the so called Schmidt-rank. Since  $\tilde{U}$  and  $\tilde{V}$  are unitary it is possible to perform two basis transformations:  $|A\rangle_\alpha := \sum_j \tilde{U}_{j\alpha} |j\rangle_A$  and  $|B\rangle_\alpha := \sum_k (\tilde{V}^\dagger)_{\alpha k} |k\rangle_B$  and express the state  $|\Psi\rangle$  in the new basis:

Single sum !

$$|\Psi\rangle = \sum_{\alpha=1}^{\chi} \lambda_\alpha |A\rangle_\alpha |B\rangle_\alpha, \quad \text{with } \chi \leq \min(\dim(A), \dim(B)). \quad (3.3)$$

This Schmidt-decomposition of a general state  $|\Psi\rangle$  always exists!

The normalization  $\langle \Psi | \Psi \rangle = 1$  implies

$$\sum_{\alpha} \lambda_{\alpha}^2 = 1.$$

Note: a unitary transformation *within a subsystem* does not change the values lambda\_alpha !

E.g.:  $|k\rangle = \sum_p W_{k,p} |p\rangle$  with a unitary matrix  $W$  and a new basis  $|p\rangle$  in subsystem  $B$

(3.4)

$$\begin{aligned} \text{Thus } |B\rangle_{\alpha} &= \sum_k (V^\dagger)_{\alpha k} |k\rangle \\ &= \sum_k (V^\dagger W)_{\alpha k} |k\rangle \\ &= \sum_p (V^\dagger W)_{\alpha p} |p\rangle \end{aligned}$$

i.e.  $(V^\dagger)$  is replaced by  $(V^\dagger W)$  in the SVD.

### 3.2 Reduced Density Matrix of a pure state

The total density matrix of a *pure* state is just:  $\hat{\rho} = |\Psi\rangle \langle \Psi|$

With the results from the previous section this can be written as:

$$\hat{\rho} = \sum_{\alpha=1}^{\chi} \sum_{\beta=1}^{\chi} \lambda_{\alpha} \lambda_{\beta} |A\rangle_{\alpha} |B\rangle_{\alpha} \langle A|_{\beta} \langle B|$$

In the state  $|\psi\rangle$ :

$$\langle O \rangle = \langle \psi | O | \psi \rangle$$

$$= \text{trace} (|\psi\rangle \langle \psi| O)$$

$$= \text{trace} (\rho O)$$

The *reduced density matrix* for the subsystem  $A$  is defined by taking the trace over the subsystem  $B$ :

$$\hat{\rho}_A = \text{tr}_B \hat{\rho} = \sum_{\gamma} \gamma \langle B | \hat{\rho} | B \rangle_{\gamma}$$

and since the bases  $|A\rangle$  and  $|B\rangle$  are orthonormal, one gets:

$$\hat{\rho}_A = \sum_{\gamma=1}^{\chi} \lambda_{\gamma}^2 |A\rangle_{\gamma} \langle A|$$

For an observable  $O$   
which only acts on subsystem  $A$ :

$$\begin{aligned}
 \langle O \rangle &= \langle \psi | O | \psi \rangle \\
 &= \text{sum\_a sum\_b lambda\_a lambda\_b} \\
 &\quad \langle A_b | \langle B_b | O | A_a \rangle \text{ (3.5)} \\
 &= \text{sum\_a sum\_b lambda\_a lambda\_b} \\
 &\quad \langle B_b | B_a \rangle \langle A_b | O | A_a \rangle \\
 &= \text{sum\_a lambda\_a}^2 \langle A_b | O | A_a \rangle \\
 \\
 \langle O \rangle &= \text{trace\_AB rho O} \\
 &= \text{trace\_A ( rho\_A O )} \\
 &= \text{sum_gamma lambda_gamma}^2 \\
 &\quad \langle A_gamma | O | A_gamma \rangle
 \end{aligned} \tag{3.6}$$

This is a sum over eigenvalues<sup>1</sup>  $\lambda_{\gamma}^2$  times a corresponding projection operator  $|A\rangle_{\gamma} \langle A|$ .

- When there is only one term in the sum,  $\chi = 1$ , then  $\hat{\rho}_A$  is a pure state and the Schmidt decomposition is simply  $|\Psi\rangle = |A\rangle |B\rangle$ , i.e. the total system is in a product state.  
Example:  $|\Psi\rangle = |\uparrow\rangle_A |\downarrow\rangle_B = |\uparrow\downarrow\rangle$
- When there is more than one term in the sum,  $\chi > 1$ , the reduced density matrix is a *mixed state*, and the eigenvalues  $\lambda_{\gamma}^2$  correspond to the probability to be in the state  $|A\rangle_{\gamma}$ .  
Example: The total state is a singlet  $|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$  (see below) [Entangled state](#)

The reduced density matrix for sub-system  $B$  is defined completely analogous to the sub-system  $A$  case, which results in the same Schmidt rank  $\chi$ .

$$\hat{\rho}_B = \sum_{\gamma=1}^{\chi} \lambda_{\gamma}^2 |B\rangle_{\gamma} \langle B| \tag{3.7}$$

### 3.3 Entanglement

The entanglement is similar in definition to the entropy of a statistical system which is

$$S = - \text{tr}(\hat{\rho} \ln \hat{\rho}). \tag{3.8}$$

Note that the entropy of a pure state is zero since:

$$\hat{\rho} = |\Psi\rangle \langle \Psi| \hat{=} (1) \Rightarrow \ln \hat{\rho} = 0 \text{ (see below)} \tag{3.9}$$

#### 3.3.1 Von Neumann entanglement entropy

The von Neumann entropy of a subsystem  $A$  is defined as the entropy of its reduced density matrix:

$$S_A := - \text{tr}_A(\hat{\rho}_A \ln \hat{\rho}_A) \tag{3.10}$$

It is a measure of the entanglement of the two subsystems  $A$  and  $B$ .

Since the reduced density matrix  $\hat{\rho}_A$  is diagonal in the basis  $|A\rangle$ : (given by the Schmidt-decomposition)

$$\hat{\rho}_A = \sum_{\gamma=1}^{\chi} \lambda_{\gamma}^2 |A\rangle_{\gamma} \langle A|,$$

---

<sup>1</sup>Note that the lecture notes for introductory quantum mechanics called the eigenvalues  $\lambda_{gamma}$ , without squares.

the von Neumann entropy reduces to

$$S_A = - \sum_{\gamma=1}^{\chi} \lambda_{\gamma}^2 \ln \lambda_{\gamma}^2 = S_B \quad (3.11)$$

Thus the entanglement entropy of subsystem  $A$  relative to  $B$  is the same as that of subsystem  $B$  relative to  $A$ .

$S_A$  takes its maximum possible value when all  $\lambda_{\gamma}$  are of equal value, i.e.,  $\lambda_{\gamma}^2 = \frac{1}{\chi}$  (since  $\sum_{\gamma} \lambda_{\gamma}^2 = 1$ ). Then the reduced density matrix is just a normalized identity matrix:

$$\hat{\rho}_A = \frac{1}{\chi} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \quad (3.12)$$

and  $S_A$  is given by:

$$S_A = -\chi \left( \frac{1}{\chi} \ln \frac{1}{\chi} \right) = \ln \chi \quad (3.13)$$

### 3.3.2 Examples for a two site system

#### • Product state $|\Psi\rangle = |\uparrow_A \uparrow_B\rangle = |\uparrow\rangle_A |\uparrow\rangle_B$

- Schmidt decomposition:  $|\Psi\rangle = |\uparrow\rangle_A |\uparrow\rangle_B = \sum_{\alpha=1}^1 1 |\uparrow\rangle_A |\uparrow\rangle_B$
- Reduced density matrix:  $\hat{\rho}_A = |\uparrow\rangle_A \langle \uparrow|$ , which is a pure state.
- Entropy:  $S_A = S_B = -1^2 \ln 1^2 = 0$ , i.e. there is no entanglement!

#### • Singlet: $|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle_A |\downarrow\rangle_B - |\uparrow\rangle_B |\downarrow\rangle_A)$

- Schmidt decomposition: already done!  $|\Psi\rangle = \sum_{\alpha=1}^2 \lambda_{\alpha} |A\rangle_{\alpha} |B\rangle_{\alpha}$ , where

$$\begin{aligned} |A\rangle_1 &= |\uparrow\rangle_A & , |A\rangle_2 &= |\downarrow\rangle_A \\ |B\rangle_1 &= |\downarrow\rangle_B & , |B\rangle_2 &= -|\uparrow\rangle_B \\ \lambda_1 &= \frac{1}{\sqrt{2}} & , \lambda_2 &= +\frac{1}{\sqrt{2}} \end{aligned}$$

- Reduced density matrix:

$$\begin{aligned} \hat{\rho}_A &= \sum_{\alpha=1}^2 \lambda_{\alpha}^2 |A\rangle_{\alpha} \langle A| = \frac{1}{2} (|\uparrow\rangle_A \langle \uparrow| + |\downarrow\rangle_A \langle \downarrow|) \\ &\stackrel{\textcolor{red}{\lambda_1 = \lambda_2}}{=} \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_A = \frac{1}{2} \mathbb{1}_A \quad \rightarrow \text{ mixed state} \end{aligned}$$

Entanglement

- Entropy:  $S_A = -2(\frac{1}{2} \ln \frac{1}{2}) = \ln 2$  This is the MAX possible entanglement ==> the two spins cannot also be entangled with anything else!

# 4 Derivation of MPS

## 4.1 Exact representation of a state by an MPS

The coefficients of any pure state

$$|\Psi\rangle = \sum_{s_1 \dots s_L} c_{s_1 \dots s_L} |s_1 \dots s_L\rangle$$

can be written as a “Matrix Product State” by going through the system site by site and performing Schmidt decompositions, i.e. basis transformations, at each site.

### • First site

We treat the coefficients  $c_{s_1, (s_2 \dots s_L)}$  as a matrix with row index  $s_1$  and column index  $(s_2 \dots s_L)$  and apply a SVD:

$$c_{s_1, (s_2 \dots s_L)} = \sum_{\alpha_1=1}^2 \underbrace{U_{s_1 \alpha_1}^{[1]}}_{2 \times 2 \text{ matrix}} \lambda_{\alpha_1}^{[1]} V_{\alpha_1 (s_2 \dots s_L)}^\dagger \quad \text{U\_tilde, V\_tilde !!}$$

The upper index [1] denotes the first lattice site. Since the index  $s_1$  has only two values,  $s_1 = \uparrow$  and  $s_1 = \downarrow$ , the matrix  $U^{[1]}$  is  $(2 \times 2)$  dimensional. We split it into two  $(1 \times 2)$ -matrices  $A^{[1]\uparrow}$  and  $A^{[1]\downarrow}$  for the two spin components  $\uparrow$  and  $\downarrow$ .

$$\begin{array}{ccc} \boxed{U_{s_1 \alpha_1}^{[1]}} & = & \boxed{\begin{array}{c} s_1^\uparrow \\ \hline s_1^\downarrow \end{array}} \\ \alpha_1 = 1, 2 & & \alpha_1 = 1, 2 \\ & & \boxed{\begin{array}{c} A_{\alpha_1}^\uparrow \\ \hline A_{\alpha_1}^\downarrow \end{array}} \\ & & \alpha_1 = 1, 2 \end{array}$$

$U$  is a unitary matrix. It contains a basis transformation from the basis  $s_1 = \uparrow, s_1 = \downarrow$  to a new basis with indices  $\alpha_1 = 1, 2$ . Similarly,  $V^\dagger$  defines a basis transformation on the vectors  $|s_2, s_3, \dots, s_L\rangle$ , i.e. we get the Schmidt decomposition (Vdagger tilde)

$$|\Psi\rangle = \sum_{\alpha_1=1}^2 \lambda_{\alpha_1}^{[1]} |\Phi_{\alpha_1}^L\rangle |\Phi_{\alpha_1}^R\rangle, \quad (4.1)$$

where  $L(R)$  denote the left(right) subsystem.

### • Second site

We now regard  $\lambda_{\alpha_1} V_{\alpha_1 (s_2 \dots s_L)}^\dagger$  (including the diagonal matrix  $\lambda$  !) as a matrix element (matrix)  $(\alpha_1 s_2) (s_3 \dots s_L)$  with row index  $(\alpha_1 s_2)$  and column index  $(s_3 \dots s_L)$  and decompose it with a SVD:

$$\lambda_{\alpha_1} V_{\alpha_1 (s_2 \dots s_L)}^\dagger = \sum_{\alpha_2=1}^4 \underbrace{U_{(\alpha_1 s_2) \alpha_2}^{[2]}}_{4 \times 4 \text{ matrix}} \lambda_{\alpha_2}^{[2]} V_{\alpha_2 (s_3 \dots s_L)}^\dagger$$

The summation index  $\alpha_2$  now goes up to 4, because of the possible combinations of  $\alpha_1 = \{1, 2\}$  and  $s_2 = \{\uparrow, \downarrow\}$ <sup>1</sup>

The  $4 \times 4$  matrix  $U^{[2]}$  is again unitary and a basis transformation from  $\{\alpha_1 s_2\}$  to  $\{\alpha_2\}$  which we split up into two  $2 \times 4$  matrices  $A^{[2]\uparrow}$  and  $A^{[2]\downarrow}$  for the two spin indices  $s_2$ :

$$U_{\alpha_1 s_2 \alpha_2}^{[2]} =: \begin{array}{c} A_{\alpha_2}^{\uparrow} \\ \hline A_{\alpha_2}^{\downarrow} \end{array} \quad \alpha_2 = 1 \dots 4$$

### • After iteration up to site $j$

After a sequence of  $j$  such steps we get the following representation of the state  $|\Psi\rangle$ :

$$|\Psi\rangle = \sum_{s_1 \dots s_L} \sum_{\alpha_1}^2 \sum_{\alpha_2}^4 \sum_{\alpha_3}^8 \dots U_{s_1 \alpha_1}^{[1]} U_{(\alpha_1 s_2) \alpha_2}^{[2]} \dots U_{(\alpha_{j-1} s_j) \alpha_j}^{[j]} \lambda_{\alpha_j}^{[j]} V_{\alpha_j (s_{j+1} \dots s_L)}^{\dagger} |s_1 \dots s_L\rangle \quad (4.2)$$

This is also a Schmidt decomposition of  $|\Psi\rangle$ , between sites  $j$  and  $j + 1$ .

By splitting up each  $U$ -matrix into two distinct matrices for  $\uparrow$  and  $\downarrow$  spin indices, we can also write

$$U_{s_1 \alpha_1}^{[1]} U_{(\alpha_1 s_2) \alpha_2}^{[2]} U_{(\alpha_2 s_3) \alpha_3}^{[3]} \dots = A_{\alpha_1}^{[1] s_1} A_{\alpha_1 \alpha_2}^{[2] s_2} A_{\alpha_2 \alpha_3}^{[3] s_3} \dots$$

### • Exact MPS representation of $|\Psi\rangle$

Continuing until the last lattice site, we find that indeed an arbitrary state  $|\Psi\rangle$  can be represented exactly by a Matrix Product State:

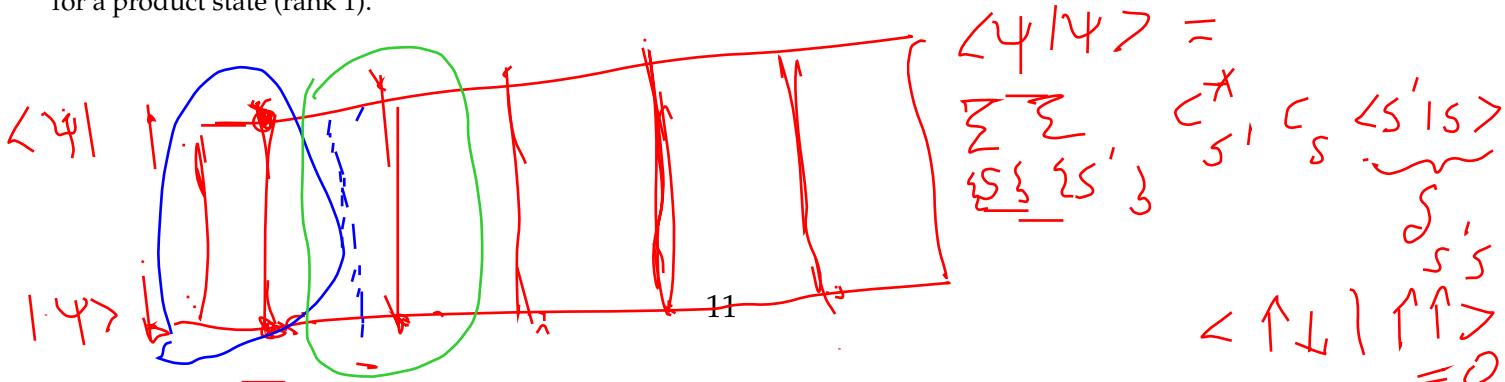
$$|\Psi\rangle = \sum_{s_1 \dots s_L} \sum_{\{\alpha_i\}} A_{\alpha_1}^{[1] s_1} A_{\alpha_1 \alpha_2}^{[2] s_2} A_{\alpha_2 \alpha_3}^{[3] s_3} \dots A_{\alpha_{L-2} \alpha_{L-1}}^{[L-1] s_{L-1}} A_{\alpha_{L-1}}^{[L] s_L} |s_1 \dots s_L\rangle \quad (4.3)$$

Graphical representation:



The range of the intermediate indices  $\alpha_i$  is equal to the rank of the corresponding SVD. Note that between sites  $L - 1$  and  $L$ , this rank is at most 2, and between sites  $L - 2$  and  $L - 1$ , it is at most 4. Thus, an exponentially large maximum rank of up to  $2^{L/2}$  is reached in the middle of the chain for a general state.

<sup>1</sup>The actual range of  $\alpha$  can be smaller, when the coefficients of  $|\Psi\rangle$  are such that the SVD has a lower rank, e.g. for a product state (rank 1).



## 4.2 Normalization

Each of the matrices  $U^{[j]}$  comes from an SVD and therefore satisfies  $U^{[j]\dagger} U^{[j]} = 1$ , which also provides for normalized basis transformations. In terms of the matrices  $A^{[j]}$  this becomes

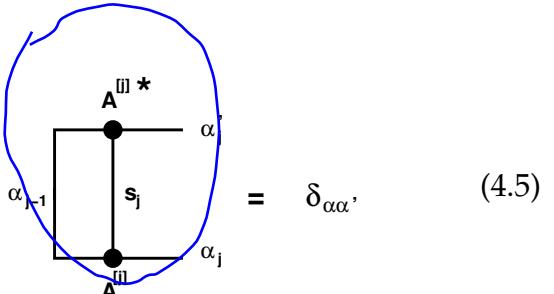
The matrices A are called "left-normalized"

$$\sum_{s_j} A^{[j]s_j\dagger} A^{[j]s_j} = 1 \quad (\text{chi x chi}) \quad (4.4)$$

Examples: pages 3 and 3a

Written in matrix components, this equation reads

$$\sum_{s_j, a_{j-1}} A_{\alpha_{j-1}s_j}^* A_{a_{j-1}\alpha_j} = \delta_{\alpha_j a'_j} \quad \text{or graphically: } \quad (4.5)$$



In the graphical representation, closed lines imply a summation.

The normalization of the whole MPS,  $\langle \Psi | \Psi \rangle = 1$ , can now be deduced in a simple way, by applying the graphical form of eq. 4.5 site by site.

N.B. Similarly, one can start from the right-most site and work to the left, obtaining a sequence of matrices  $V^\dagger$ . This can be split (left half, right half) into two matrices  $B^\text{up}$  and  $B^\text{down}$ .

The eq  $V^\dagger V = 1$  can be rewritten as  $\sum_s B^\dagger B = 1$  similar to eq (4.4). The  $B$ -matrices are called "right-normalized".

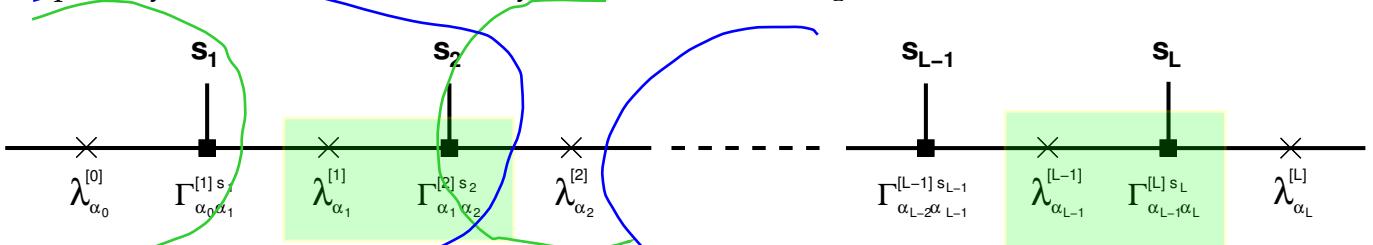
## 4.3 Canonical form of an MPS

Each Schmidt decomposition in the derivation of the MPS also gave us the singular values  $\lambda_\alpha$ , i.e. information about the reduced density matrix at that step. We now write this information explicitly in the MPS. We take the singular values  $\lambda_\alpha$  out of the A-matrices; this defines new matrices " $\Gamma$ ":

$$A_{\alpha_{j-1}s_j}^{[j]} =: \lambda_{\alpha_{j-1}}^{[j-1]} \Gamma_{\alpha_{j-1}s_j}^{[j]} \quad \text{and similarly } B = \text{Gamma lambda} \quad (4.6)$$

When the A-matrices are given, one can obtain the  $\Gamma$  matrices by multiplying A with the *pseudoinverse* of  $\lambda$ .

Graphically, the "MPS state" (actually the coefficient  $c_{s_1s_2\dots s_L}$ ) now becomes



We denote A-matrices by circles, and  $\Gamma$  matrices by squares. The  $\lambda^{[j]}$  are located between sites  $j$  and  $j + 1$ . They are the singular values of the Schmidt decomposition at that bond,  $c_{s_1s_2\dots s_L} = U\lambda V^\dagger$ . In a product state, all  $\Gamma^{[i]}$  are numbers 0 or 1, and all  $\lambda^{[i]} = 1$ .

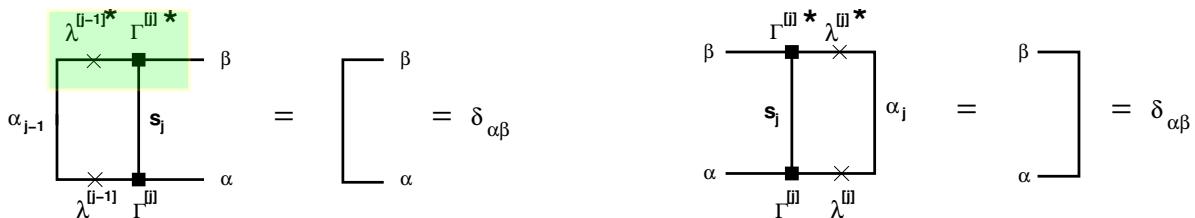
Since the first "matrix"  $A_{\alpha_1}^{[1]s_1}$  is just a vector, the first index  $\alpha_0$  on  $\Gamma_{\alpha_0\alpha_1}^{[1]s_1}$  is superfluous. We still write  $\Gamma_{\alpha_0\alpha_1}^{[1]s_1}$  as a matrix, for easier notation later on, by letting the index  $\alpha_0$  only have one value  $\alpha_0 = 1$  and  $\lambda_{\alpha_0}^{[0]} \equiv 1$ . Similarly,  $\alpha_L$  has only the one value  $\alpha_L \equiv 1$  and  $\Gamma_{\alpha_{L-1}\alpha_L}^{[L]s_L}$  is actually a vector. At the right hand side we have introduced a final  $\lambda_{\alpha_L}^{[L]} \equiv 1$ .

Examples: See excercise sheet

**Normalization** The normalization eq. 4.5 now becomes an equation for  $\Gamma$  and  $\lambda$ , shown below on the left hand side:

$$\sum_{s_j} (\Gamma^{[j]s_j})^\dagger (\lambda^{[j-1]})^2 \Gamma^{[j]s_j} = \mathbb{1} \quad (4.7)$$

$$\sum_{s_j} \Gamma^{[j]s_j} (\lambda^{[j]})^2 (\Gamma^{[j]s_j})^\dagger = \mathbb{1}$$



(This is equivalent to  $V^\dagger V = 1$ ,  
also equiv to the normalization of the B-matrices)

A similar normalization can also be shown to hold when summing over the second matrix index of  $\Gamma^{[j]}$ . This normalization is displayed on the right side of the equation above.

## 5 Operations on MPS

### 5.1 Truncation

In order to achieve an efficient representation of a state, with relatively small matrices, we can approximate the state by **discarding** small singular values  $\lambda_\alpha$ . This is done most easily in the canonical representation. When all  $\lambda_{\alpha > \alpha_0}^{[j]}$  are discarded, the matrices  $\Gamma^{[j]}$  and  $\Gamma^{[j+1]}$  can be *truncated* correspondingly beyond matrix index  $\alpha_0$ . One can either discard values below a certain **threshold**  $\epsilon$  (e.g.  $10^{-10}$ ), which results in a varying matrix dimension, or one can set a maximum matrix dimension  $\chi_{max}$  beyond which all singular values are discarded.

The quality of the approximation is related to how much of the reduced density matrix  $\rho = \text{diag}(\lambda_1^2, \lambda_2^2, \dots)$  is discarded. This can be quantified by the so-called *discarded weight* (which should stay below a small threshold  $\ll 10^{-6}$ ).

$$w = 1 - \sum_{\alpha=1}^{\chi_{max}} \lambda_\alpha^2 \quad (5.1)$$

Normalization:  $\text{sum}_i (\lambda_i)^2 = 1$

In order to keep the state normalized, we have to re-normalize the remaining  $\lambda_\alpha$ :

$$\lambda_\alpha \rightarrow \lambda_\alpha / \sqrt{1-w}, \quad \text{so that} \quad \sum_{\alpha} \lambda_\alpha^2 = 1. \quad (5.2)$$

The normalizations of the A-matrices and the  $\Gamma$  matrices are unaffected by the truncation, except that the matrix  $\mathbb{1}$  in eq. 4.4 is also truncated.

Better: check quality of truncation by looking at convergence of results with lower and lower threshold. MUST be done for quantitative results.

N.B.: Procedure for **truncation of a normalized MPS** given in the form A A A A or lambda Gamma lambda:

- (1) perform a sequence of SVDs and truncate the lambdas (again, see Schollwöck review 2011, p.129 (use commented version, on the webpage of this class) ), or
- (2) Perform a time evolution step using TEBD (see below), e.g. with Hamiltonian  $H=0$

## 5.2 How large do the matrices need to be ?

The discarded weight  $w$  is small when the singular values  $\lambda_\alpha$  decay quickly. This is the case when the entanglement entropy eq. 3.11,  $S_A = - \sum_{\gamma=1}^{\chi} \lambda_\gamma^2 \ln \lambda_\gamma^2$  is small. We saw earlier that the maximum entanglement entropy of a reduced density matrix of size  $\chi$  is  $\ln \chi$ . Thus one can estimate that one needs matrices of size about

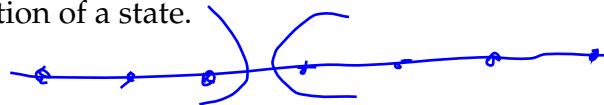
$$\chi_{max} \sim e^{S_A} \quad (5.3)$$

$\uparrow \downarrow - \leftarrow \rightarrow$

for a good representation of a state.

### • 1D

Chain



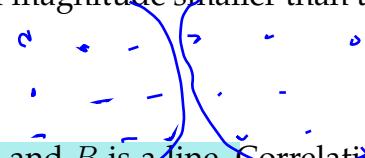
In 1D, the border of two subsystems  $A$  and  $B$  is just a point. All entanglement between  $A$  and  $B$  must go through this point. It is related to the spin correlations between  $A$  and  $B$ . When the Hamiltonian contains only local couplings, these correlations are limited in spatial extent. One can show that for the ground state of Hamiltonians with local couplings one needs only

(but: time evolution after some **excitation**  
will result in growing entanglement, sometimes linear in time!)  
== restriction on time that can be simulated

$$\chi_{max} \sim \ln \xi \quad (5.4)$$

where  $\xi$  is the spatial correlation length of the system. When  $\xi$  is finite, a finite value of  $\chi_{max}$  is therefore sufficient, even on very large systems. Even when the system is “critical”,  $\xi = \infty$ ,  $\chi_{max} \sim \ln L$  is sufficient. This is the **reason why Matrix Product States work so well in one-dimensional physical systems**.

Note that for large systems, this  $\chi_{max}$  is many orders of magnitude smaller than the generic  $2^{L/2}$  behavior found earlier for an arbitrary state.



### • 2D

In two dimensions, the border between subsystems  $A$  and  $B$  is a line. Correlations and entanglement can go across any point of this line. Since the length of the border in 2D is proportional to subsystem size  $L$ , the von-Neumann entropy  $S_A$  is also proportional to  $L$  and thus  $\chi_{max}$  has to grow exponentially:

$$S_A \sim L \rightarrow \chi_{max} \sim e^L \quad (5.5)$$

This is the reason why the MPS-Ansatz fails for dimensions  $D > 1$ <sup>1</sup> where an exponentially growing computational effort is required to accurately represent  $|\Psi\rangle$  by an MPS.

A special case occurs when 1D chains are coupled. The required matrix dimension then grows exponentially in the number of coupled chains, but when this number remains small (e.g. 2 or 3 or 4), the resulting matrix dimension remains within reach.

<sup>1</sup>In  $D=3$ , the boundary is an area, and the general relation  $\chi_{max} \sim L^{D-1}$  is therefore called the **area law**, even though it actually refers to a boundary.

### 5.3 Expectation values of one-site operators

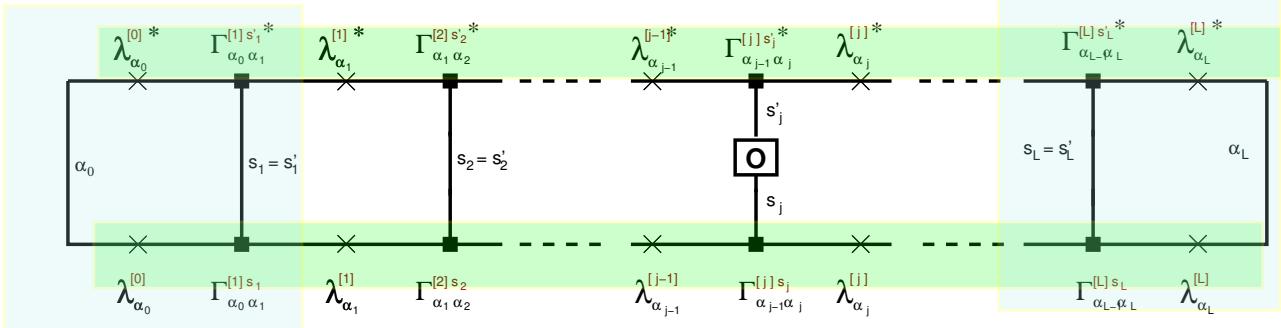
A very big advantage of the MPS representation of states is that **only a few local matrices are needed to calculate the effect of a local operator** on the state (see below), or its expectation value. We will here look at the expectation value of a one-site operator  $\hat{O}^{[j]}$  that acts only on the spin at site  $j$ . Example: the operator  $\hat{S}_z^{[j]}$ , with matrix elements (in z-basis)

$$(S_z) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (5.6)$$

The expectation value is

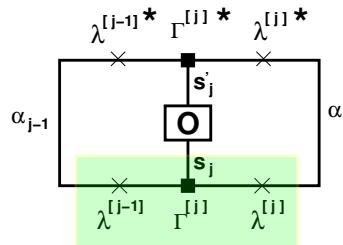
$$\langle \Psi | \hat{O}^{[j]} | \Psi \rangle = \sum_{\{s\}, \{s'\}} \langle s'_1 \dots s'_L | \dots \lambda^{*[j]} \Gamma^{*[j]s'_j} \lambda^{*[j-1]} \dots O_{s_j s'_j}^{[j]} \dots \lambda^{[j-1]} \Gamma^{[j]s_j} \lambda^{[j]} \dots | s_1 \dots s_L \rangle \quad (5.7)$$

In graphical representation, this becomes



At the interior vertical lines, the spins  $s_i'$  from the bra vector  $\langle \psi |$  and  $s_i$  from the ket vector  $| \psi \rangle$  meet. Since  $\langle s'_i | s_i \rangle = \delta_{s_i s'_i}$ , they have to be equal, except at the location  $j$  of the operator  $\hat{O}^{[j]}$ . Note that all lambda values are actually real, since they are singular values.

This expression can now be simplified by using the normalization eqs. 4.7, which iteratively cause all matrices from both ends of the chain up to site  $j$  to just contribute Kronecker deltas. The remaining contribution is, in graphical representation and as an equation



$$\langle \Psi | \hat{O}^{[j]} | \Psi \rangle = \sum_{s_j, s'_j} \langle s'_j | \lambda^{[j]} \Gamma^{*[j]s'_j} \lambda^{[j-1]} O_{s_j s'_j}^{[j]} \lambda^{[j-1]} \Gamma^{[j]s_j} \lambda^{[j]} | s_j \rangle . \quad (\text{with summations over the matrix indices alpha}) \quad (5.8)$$

This is the same as

$$\langle \Psi | \hat{O} | \Psi \rangle = \sum_{\alpha_{j-1} \alpha_j} \sum_{s, s'} \left( \lambda_{\alpha_{j-1}}^{[j-1]} \lambda_{\alpha_j}^{[j]} \right)^2 \left( \Gamma_{\alpha_{j-1} \alpha_j}^{[j]s'} \right)^* \Gamma_{\alpha_{j-1} \alpha_j}^{[j]s} \langle s' | \hat{O} | s \rangle \quad (5.9)$$

or by defining the (spatially local) matrix  $M$ :

$$M_{\alpha_{j-1} \alpha_j}^{s_j} := \lambda_{\alpha_j}^{[j-1]} \Gamma_{\alpha_{j-1} \alpha_j}^{[j]s_j} \lambda_{\alpha_j}^{[j]} . \quad (5.10)$$

the expectation value of  $\hat{O}$  can be calculated as:

$$\langle \psi | \hat{O}^{[j]} | \psi \rangle = \sum_{s, s'} \langle s' | \hat{O} | s \rangle \operatorname{tr} \left( M^{s'} \right)^\dagger M^s . \quad (\text{index } j \text{ left out}) \quad (5.11)$$

# 6 Time evolution

The time evolution of a state in the Schrödinger picture for a time-independent hamiltonian is given by

$$|\Psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}t}|\Psi(0)\rangle$$

## 6.1 Trotter Suzuki decomposition

For the Heisenberg-model the Hamilton operator  $\hat{H}$  can be expressed as:

$$\hat{H} = \sum_i \hat{H}_i \quad \text{with} \quad \hat{H}_i = \frac{J_{xy}}{2} \left[ \hat{S}_i^+ \hat{S}_{i+1}^- + \hat{S}_i^- \hat{S}_{i+1}^+ \right] + J_z \hat{S}_i^z \hat{S}_{i+1}^z$$

The problem now is that the Hamiltonians of two adjacent sites do not commute,  $[\hat{H}_i, \hat{H}_{i+1}] \neq 0$  and as a consequence the exponential  $e^{-i\hat{H}t}$  is not expressible as a product of local operators:

$$e^{-i\hat{H}t} \neq \prod_j e^{-i\hat{H}_j t}$$

But the Hamiltonians of next nearest neighbor sites do commute:  $[\hat{H}_i, \hat{H}_{i+2}] = 0$ .

It is therefore helpful to decompose  $\hat{H}$  into a sum of even and odd site Hamiltonians:

$$\hat{H} = \hat{H}_{even} + \hat{H}_{odd} = \sum_{j,odd} \hat{H}_j + \sum_{j,even} \hat{H}_j \quad \begin{array}{ccccccc} \text{---} & \text{---} & \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ | & | & | & | & | & | & | \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \quad (6.1)$$

such that

$$e^{-i\hat{H}_{even}t} = \prod_{j,even} e^{-i\hat{H}_j t} \quad \text{and} \quad e^{-i\hat{H}_{odd}t} = \prod_{j,odd} e^{-i\hat{H}_j t}. \quad (6.2)$$

Within  $\hat{H}_{even}$ , all terms commute, and also within  $\hat{H}_{odd}$ , but  $[\hat{H}_{even}, \hat{H}_{odd}] \neq 0$ .

Next we subdivide the time  $t$  into small “time slices” of length  $\Delta t$

$$e^{-i\hat{H}t} = \left( e^{-i\hat{H}\Delta t} \right)^{\frac{t}{\Delta t}} = \left( e^{-i(\hat{H}_{even} + \hat{H}_{odd})\Delta t} \right)^{\frac{t}{\Delta t}} \quad \text{exact}$$

Now we use the Baker-Hausdorff formula to get

$$e^{-i\hat{H}t} = \left\{ e^{-i\hat{H}_{even}\Delta t} e^{-i\hat{H}_{odd}\Delta t} \left( 1 + \mathcal{O}((\Delta t)^2 [\hat{H}_{even}, \hat{H}_{odd}]) \right) \right\}^{\frac{t}{\Delta t}}$$

Writing this product of  $t/\Delta t$  terms explicitly yields

is (at most) of order L: L commutators of two-site spin-operators

$$e^{-i\hat{H}t} = \left( e^{-i\hat{H}_{even}\Delta t} e^{-i\hat{H}_{odd}\Delta t} e^{-i\hat{H}_{even}\Delta t} \dots e^{-i\hat{H}_{\text{odd}}\Delta t} \right) (1 + \tilde{\mathcal{O}}(t \Delta t)) \quad (\text{still unitary: helps}) \quad (6.3)$$

We lost one order of  $\Delta t$  because of the  $(t/\Delta t)$  many factors.

(faster error increase with time t when using MPS with truncations at each time step)

The Trotter-Suzuki decomposition leads to a series of operators  $e^{-i\hat{H}_j t}$  which only act on two adjacent sites at once. The time evolution of the system is traced back to application of these 2-site operators.

The smaller the time step  $\Delta t$ , the smaller the error in the method. One can gain another order of  $\Delta t$  with almost no effort by the so-called 2nd order Trotter Suzuki approximation:

$$e^{-i\hat{H}\Delta t} = e^{-i\hat{H}_{even}\frac{\Delta t}{2}} e^{-i\hat{H}_{odd}\Delta t} e^{-i\hat{H}_{even}\frac{\Delta t}{2}} + O((\Delta t)^3), \quad (6.4)$$

For the time evolution of the state  $|\Psi\rangle$  this is the *same* as the first order approximation, because  $e^{-i\hat{H}_{even}\frac{\Delta t}{2}} e^{-i\hat{H}_{even}\frac{\Delta t}{2}} = e^{-i\hat{H}_{even}\Delta t}$ . The only difference occurs when measurements are performed: for the 2nd order approximation, measurements have to be performed after *half-time-steps*  $e^{-i\hat{H}_{even}\frac{\Delta t}{2}}$ .

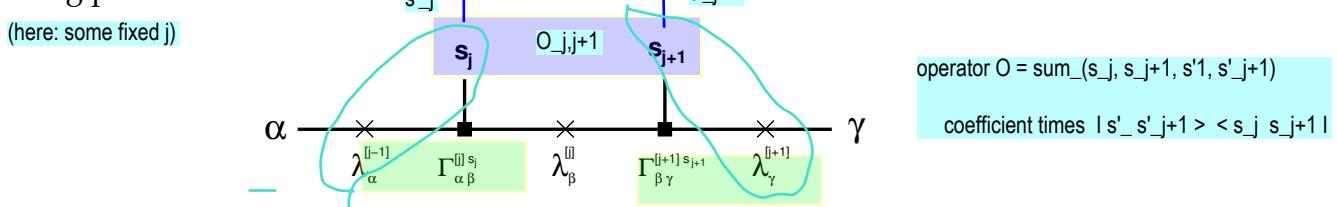
and very first and very last time steps: size Delta-t/2

## 6.2 Application of 2-site operators

In order to calculate the time evolution, we need to apply the 2-site operators  $\hat{H}_i$  to  $|\Psi\rangle$ . For a general 2-site operator, we want to calculate

$$|\Psi'\rangle = \hat{O}^{[j,j+1]} |\Psi\rangle. \quad (6.5)$$

The structure of MPS as products of matrices located on individual sites is again very helpful. Since  $\hat{O}^{[j,j+1]}$  acts on sites  $j$  and  $j+1$ , only the  $\Gamma$  matrices at these sites and the  $\lambda$  matrix in between are affected! Let  $\chi$  be the dimension of these matrices. The calculations turn out to be easier when one also includes the two  $\lambda$  matrices to the left and the right. We therefore look at eqs (3.3),(3.4) the following part of the MPS:



This object has spin indices  $s_j, s_{j+1}$  on which the operator will act, and free matrix indices  $\alpha$  and  $\gamma$ . It is the graphical representation of

$$\Theta_{\alpha\gamma}^{s_j s_{j+1}} := \sum_{\beta} \lambda_{\alpha}^{[j-1]} \Gamma_{\alpha\beta}^{[j] s_j} \lambda_{\beta}^{[j]} \Gamma_{\beta\gamma}^{[j+1] s_{j+1}} \lambda_{\gamma}^{[j+1]} \quad (6.6)$$

The application of  $\hat{O}$  on  $\Theta$  yields

$$\tilde{\Theta}_{\alpha\gamma}^{s'_j s'_{j+1}} := \sum_{s_j, s_{j+1}} \langle s'_j s'_{j+1} | \hat{O} | s_j s_{j+1} \rangle \Theta_{\alpha\gamma}^{s_j s_{j+1}} \quad (6.7)$$

We want to express  $|\Psi'\rangle$  as a normalized MPS, similar to  $|\Psi\rangle$ . We therefore need to write  $\tilde{\Theta}$  in the same form as the original  $\Theta$ , with new normalized matrices  $\tilde{\Gamma}_{\alpha\beta}^{[j] s'_j}$ ,  $\tilde{\lambda}_{\beta}$ , and  $\tilde{\Gamma}_{\beta\gamma}^{[j+1] s'_{j+1}}$ .

In order to get there, we first interpret  $\tilde{\Theta}$  as a matrix with two indices, a row index  $(\alpha s'_j)$  and a column index  $(\gamma s'_{j+1})$

$$\bar{\Theta}_{(\alpha s'_j), (\gamma s'_{j+1})} := \tilde{\Theta}_{\alpha\gamma}^{s'_j s'_{j+1}} \quad (6.8)$$

Next we perform an SVD on  $\bar{\Theta}$ :

(similar to derivation of MPS,  
see pages 10, 11)

$$\bar{\Theta}_{(\alpha s'_j), (\gamma s'_{j+1})} = \sum_{\beta=1}^{2\chi} U_{(\alpha s'_j)\beta} \tilde{\lambda}_{\beta} V_{\beta(\gamma s'_{j+1})}^{\dagger} \quad (6.9)$$

Because of the presence of  $s'_j$  and  $s'_{j+1}$  in the indices (with 2 values  $\uparrow, \downarrow$ ), this SVD has a Schmidt rank up to *twice the rank* of the original matrices  $\Gamma, \lambda$ . If we kept this increased rank, then the matrix dimensions would explode exponentially during the time evolution.

We therefore need to truncate the matrix dimensions, for example back to the original  $\chi$  (or a smaller value), by discarding the smallest singular values in  $\tilde{\lambda}$ . We then need to calculate the discarded weight  $w$  and re-normalize  $\tilde{\lambda}_\beta \rightarrow \tilde{\lambda}_\beta / \sqrt{1-w}$ . The matrices  $U$  and  $V^\dagger$  are also truncated at the new size.

Finally, we extract new  $\Gamma$  matrices, by splitting the unchanged outer matrices  $\lambda^{[j-1]}$  and  $\lambda^{[j]}$  off  $U$  and  $V^\dagger$ , by means of applying the pseudoinverse:

$$\tilde{\Gamma}_{\alpha\beta}^{[j]s'_j} = (\lambda_\alpha^{[j-1]})^{\text{inv}} U_{(\alpha s'_j)\beta}, \quad \tilde{\Gamma}_{\beta\gamma}^{[j+1]s'_{j+1}} = V_{\beta(\gamma s'_{j+1})}^\dagger (\lambda_\gamma^{[j+1]})^{\text{inv}} \quad (6.10)$$

As desired, this results in updated  $\chi \times \chi$  matrices  $\tilde{\Gamma}_{\alpha\beta}^{[j]s'_j}$ ,  $\tilde{\lambda}_\beta$ , and  $\tilde{\Gamma}_{\beta\gamma}^{[j+1]s'_{j+1}}$ .

## 6.3 Recipe for calculating the time evolution of a 1D many body state

1. Write the initial state as a canonical MPS.

For a product state, this is very simple: the “matrices”  $\Gamma^{[j]s_j}$  are just numbers, e.g. when  $s_j = \uparrow$ , then  $\Gamma^{[j]\uparrow} = 1$  and  $\Gamma^{[j]\downarrow} = 0$ . The “matrices”  $\lambda^{[j]}$  are all just the number 1.

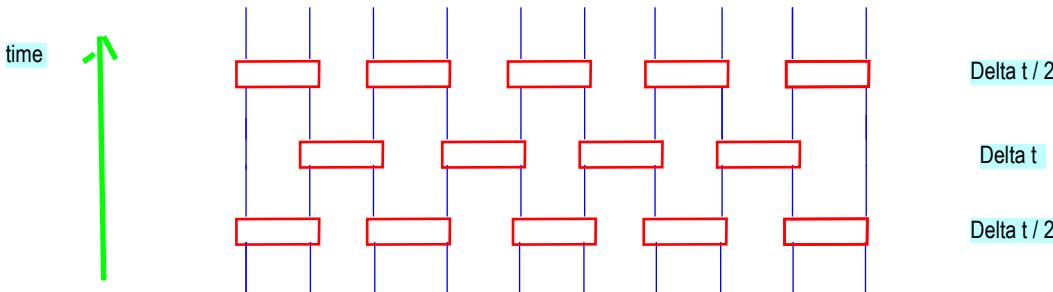
2. Calculate the matrix elements  $\langle s'_j s'_{j+1} | e^{-i\hat{H}_{j,j+1}\Delta t} | s_j s_{j+1} \rangle$  (and also for  $\Delta t/2$ ).

3. One time step (2nd order Trotter Suzuki):

- Apply  $e^{-i\hat{H}_{\text{even}}\Delta t/2}$  to the matrix product state, by applying all of the two-site operators  $H_i$  for even  $i$  to the MPS. This can be done in any sequence since these operators commute.
- Apply  $e^{-i\hat{H}_{\text{odd}}\Delta t}$ .
- Apply  $e^{-i\hat{H}_{\text{even}}\Delta t/2}$ .

4. After one or several time steps, perform measurements, e.g. of  $\langle S_z^{[j]} \rangle$ .

This algorithm, the so-called *Time Evolved Block Decimation (TEBD)*, was invented in 2004 by G. Vidal and in a closely related form (*tDMRG*) at the same time by White and Feiguin. These methods have provided a great leap forward in the investigation of large 1D many-body quantum systems, including strongly correlated ones, by finally permitting the precise and rather fast calculation of their time evolution from arbitrary initial states.



Exact solution of the tight binding model (i.e. only hopping, Coulomb repulsion V=0, or equiv. Jz=0) :

$$\hat{H} = -T \sum_j^L c_j^+ c_{j+1}^- + \underbrace{c_{j+1}^+ c_j^-}_{\text{harm. coupling of 1st term}}$$

(Here with periodic boundary conditions for simplicity.  
Can also be done with open boundary conditions)

"T" (=  $J_{xy}/2$ ) here is the "hopping amplitude",  
which was called "t" on page 1.  
Different name chosen here to avoid confusion.

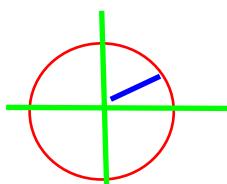
Fourier :  $c_j = \frac{1}{\sqrt{L}} \sum_k e^{ikj} \tilde{c}_k$ ,  $c_j^+ = \dots \tilde{c}_k^+$

$\hat{c}$  form. op.  $\tilde{c}_a^+ \tilde{c}_{a'}^- + \tilde{c}_{a'}^+ \tilde{c}_a^-$   
 $= \delta_{aa'}$

Inverse  $\tilde{c}_k = \frac{1}{\sqrt{L}} \sum_j e^{-ikj} c_j$   
with  $k_n = n \frac{2\pi}{L}$

Insert in  $\hat{H}$ :

$$\hat{H} = -T \frac{1}{L} \sum_{j, k, p} \left( e^{+ik(j+1)} e^{-ipj} \tilde{c}_p^+ \tilde{c}_k^- + h.c. \right)$$



$$\sum_j^L \underbrace{\sum_p}_{= L} = \sum_{k, p} \quad (AB) = B^+ A^+$$

$$= -T \sum_k e^{ik \cdot 1} \tilde{c}_k^+ \tilde{c}_k^- + e^{-ik \cdot 1}$$

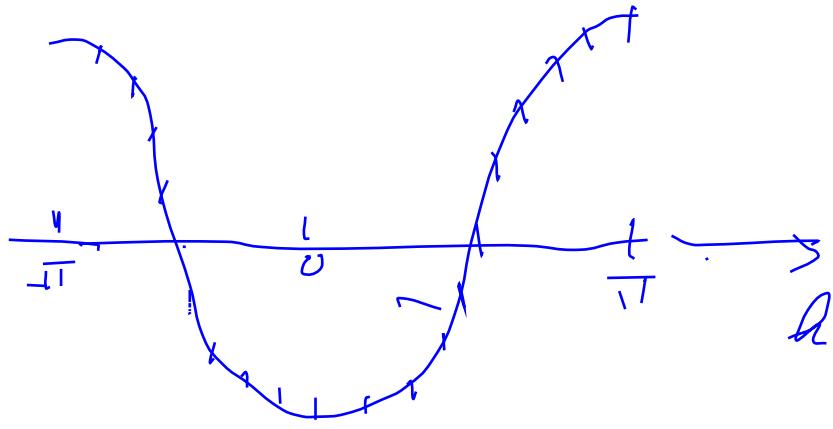
$$\tilde{c}_k^+ \tilde{c}_k^- \quad \tilde{n}_k$$

$$\hat{H} = \sum_k -2T \cos k \hat{n}_k$$

$$\hat{n}_k |0\rangle_k = 0 |0\rangle_k$$

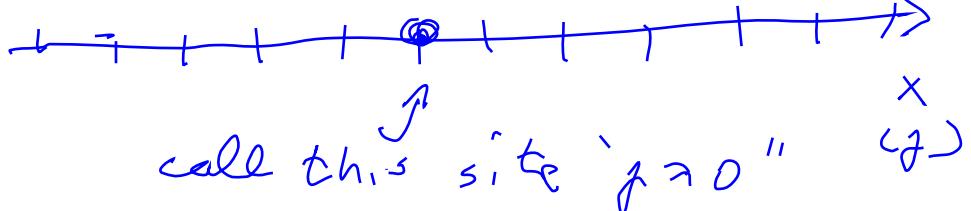
$$\hat{n}_k |1\rangle_k = 1 |1\rangle_k$$

diagonal in  
k-space!



Exact solution for a single initial particle (at fixed position) :

Initial config. :



$$|\psi(t=0)\rangle = c_j^+ |\underbrace{0\rangle}_{\text{empty state}} \quad (\text{not "ground state"})$$

$$= \frac{1}{\sqrt{L}} \sum_p e^{-i p j} \underbrace{\tilde{c}_p^+}_{j=0} |\underbrace{0\rangle}_{j=0}$$

Time evolution :

$$|\psi(t)\rangle = e^{-i \hat{H} t} |\psi(0)\rangle$$

$$= e^{-i \sum_h \epsilon_h \hat{n}_h t} \frac{1}{\sqrt{L}} \sum_p \tilde{c}_p^+ |\underbrace{0\rangle}_{j=0}$$

$$= \sum_p \frac{1}{\sqrt{L}} \underbrace{\frac{1}{\hbar} e^{-i \sum_h \hat{n}_h t}}_{\text{all } \hat{n}_h \text{ give zero, except } h=p} \tilde{c}_p^+ |\underbrace{0\rangle}_{j=0}$$

$$|\psi(t)\rangle = \frac{1}{\sqrt{L}} \sum_p e^{-i \epsilon_p t} \tilde{c}_p^+ |\underbrace{0\rangle}_{j=0}$$

Same structure as

$$c_1 \exp(-i E_1 t) |1\rangle + c_2 \exp(-i E_2 t) |2\rangle + \dots$$

$$\psi(x, t) = \langle x | \psi(t) \rangle$$

$$= \underbrace{\langle x |}_{\text{discrete "x" = site number}} \sum_p \tilde{c}^{-i\varepsilon_p t} \frac{1}{\sqrt{L}} \sum_j \tilde{c}^{+i\eta p} c_j^+ |0\rangle$$

particle at  $x$

particle at position  $j$

$$\langle x | c_j^+ | 0 \rangle = \delta_{jx}$$

$$\psi(x, t) = \frac{1}{L} \sum_p e^{+i(xp - \varepsilon_p t)}$$

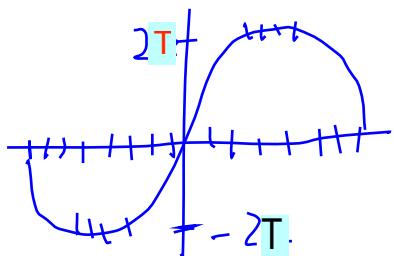
$\rightarrow \dots \rightarrow$   
Bessel-function  
see PRL 2012

$\rightarrow$  group velocity:

$$\text{at some } p : v_p = \frac{\partial \varepsilon_p}{\partial p}$$

(see QM chapter 4.7.2)

$$\text{Here: } \varepsilon_p = -2T \cos p \Rightarrow v_p = 2T \sin p$$



Many modes with  
 $|v_p|$  almost  $2T$ :

How "many" momenta contribute at a given  $\omega$ ?

Density of velocities:

$$g(\omega) = \sum_k \delta(\omega - 2T \sin k)$$

$$\Delta k = \frac{2\pi}{L}$$

$$\sum_k \Delta k \rightarrow \int dk$$

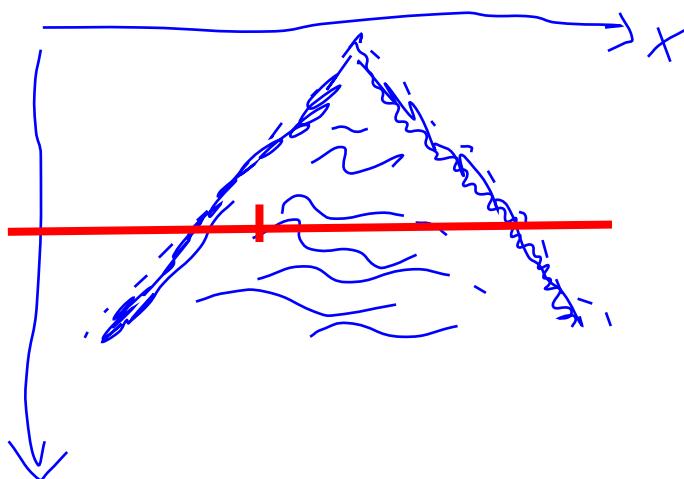
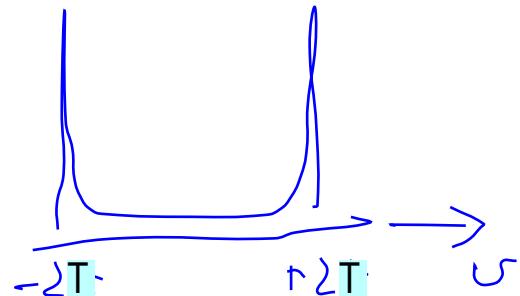
$$\approx \frac{L}{2\pi} \int dk \delta(\nu - \frac{2\pi k}{L} \sin \theta)$$

$$\delta(g(x)) = \sum_v \frac{\delta(x - x_v)}{|g'(x_v)|}$$

with zeroes:  $g(x_v) = 0$

$$g' \rightarrow \cos \theta = \sqrt{1 - \sin^2 \theta} \quad \hookrightarrow \nu$$

$$= \frac{L}{2\pi} \frac{1}{\sqrt{(2T)^2 - \nu^2}}$$



Exercise 17 Jan 2022: Start with a chain ( $L=50$  or  $100$  sites), for which the left half is occupied (spin up) and the right half is empty (spin down). Calculate the time evolution.

Observables:  $n(x,t)$  (same as  $\langle S_z(x,t) \rangle$  in spin language)

$S(t)$  (entanglement entropy, only at middle bond, between the two regions, up to fairly large times  $t=L$  or  $t=2L$ )

Cases:

- 1)  $J_z=0$ , all  $J_{xy}=1$
- 2)  $J_z=0$ ,  $J_{xy}=1$  on left half, not 1 (e.g. 0.5 or 2) on right half
- 3)  $J_z=0$ , all  $J_{xy}=1$  except for middle bond: =0.5 or =2
- 4)  $J_z=J_{xy}=1$  everywhere

Do you see qualitatively different functions of time for  $S(t)$  at large times?

Not that you still need large matrices, since matrix size goes like  $\exp(S)$