



The Abdus Salam
International Centre for Theoretical Physics


United Nations
Educational, Scientific
and Cultural Organization


International Atomic
Energy Agency



SMR 1666 - 8

**SCHOOL ON QUANTUM PHASE TRANSITIONS
AND
NON-EQUILIBRIUM PHENOMENA IN COLD ATOMIC GASES**

11 - 22 July 2005

***Quantum Many-Body Algorithms
based on
Quantum Information Theory***

Presented by:

Maciej Lewenstein

Institut de Ciències Fòniques, Barcelona

Quantum Many-Body Algorithms based on Quantum Information Theory

Maciej Lewenstein Armand Niederberger

June 21, 2005

1 Introduction

Physical systems composed of many particles are generally very difficult to study because the dimension of their associated Hilbert space is usually growing exponentially with the number of constituents. In this lecture, we review some fundamentals of Quantum Information Theory (QIT) in order to present newly proposed algorithms to simulate slightly entangled systems.

2 Classical versus quantum coin tossing

Let us start by considering by considering the tossing of an unbiased classical coin. We know that in half the cases where we toss such a coin, the outcome will be heads and in the other half, it will be tails. Furthermore, we know that, in principle, there is absolutely nothing probabilistic inherent to this experiment. If we were to include in our description the exact movement of the hand as well as the exact initial position and speed of both hand and coin, we would be able to predict the outcome of the result with certainty¹. Once we have included the appropriate initial conditions, the outcome of this experiment is completely determined by Newton's laws of motion. There is nothing that breaks classical determinism in this experiment. The probabilities are solely due to our incomplete knowledge of the system. Nevertheless, we can consider classical coin tossing as a prototype for a probabilistic classical experiment presenting two possible outcomes.

¹Of course, we can imagine situations, where additional parameters have to be included in our description in order to predict the outcome of the experiment. For example, there may be extreme winds or strong electromagnetic fields that influence the movement of the coin. In principle, however, there is *nothing inherently probabilistic* to this experiment.

Now, we would like to consider the tossing of *two* coins. A priori, for unbiased coins, the two outcomes are completely uncorrelated. The measurement of one coin does not influence the measurement of the other. If, however, we were to construct two coins – that is, two binary systems – for which the value of the first one influences the value of the other, the only way would be to include some communication between the two coins. For example, we could manufacture two binary random number generators that include wireless communication. We could then program them in such a way that whenever the first random number generator takes one value, the second one takes the other value. Note, however, that it is *not* possible to build such a classical system without kind of communication between the two experiments. In particular, if we were to measure one of these coins on Earth and the other one on Mars at the same time, the correlation would disappear since the signal emitted from the one apparatus could not reach the other before the experiment was completed. This concept of *signalling*, i.e. the fact that correlation is *not instantaneous* is absolutely fundamental for discussion of quantum correlation. The peculiar thing about the latter, commonly referred to as *entanglement*, is, that it is an experimental fact, that the correlation persists even if we exclude signalling.

So what kind of system would we consider in the quantum case? A standard example that is very suitable for intuitive argumentation is the polarisation of a single photon. Consider a vertically polarised photon. We know that measuring it with a polarising beam-splitter (PBS) that is oriented according to the horizontal-vertical basis will lead to a deterministic result. The detector corresponding to the vertical measurement of the photon will *always* fire². If, on the contrary, we were to set the PBS to the diagonal basis, the outcome of the measurement is completely random. It is important to understand that this randomness is inherent to the physical system. Above, we considered a classical coin whose measurement was random because of our incomplete knowledge of the system. This kind of randomness is referred to as *hidden variable*: the outcome of the experiment is completely deterministic, however, we cannot predict the outcome because we do not know the value of the hidden variable (for example, the movement of the hand). The example with the photon, in contrast, is totally different. There is no hidden variable that would allow us to predict what of at the PBS we will detect the photon. The outcome of the measurement is inherently random³.

²We neglect the fact that real (i.e. existing) detectors are not perfect.

³Of course, one could argue that this cannot be proven for single particles – and it is true, this claim *cannot be proven* for single particles. However, as we will see later, it can be proven for two-particle and multi-particle systems that predictions of hidden variable approaches fail to explain experimental observations. From this, combined with the fact that our claim is consistent with experiments we infer the justification of our point of view.

Talking about correlation is, obviously, talking about at least two systems. Well, with quantum systems such as photons, it is possible to realise composite systems that behave as if they felt what the other parts are doing. For example, it is possible to prepare photon pairs for which the result of the individual measurements are random but equal for both particles. With classical systems, we could try to imitate such correlation in two ways. The trivial possibility is to simply prepare the two coins for the same result. It is obvious that we would then always measure the same state of the two coins. Of course, if the state in which both photons are prepared is chosen arbitrarily, the outcome of the measurement of the coin would appear to be random but correlated to the other coin. This set-up corresponds to a measurement of a system presenting a hidden variable. There exist, however, mathematical tools that allow testing for hidden variables. Usually, these tests translate into inequalities that are satisfied for all classical correlation. In the present example, the test translates to the famous *Bell inequalities*. The violation of these inequalities proves that it is impossible to find a model with hidden variables that generates the same correlation. Experimental Bell inequality tests were performed in the seventies of the last century. They clearly proved that it is not possible to explain this quantum correlation via hidden variable models. The second, slightly more complicated way to imitate correlation is to consider high-tech coins that exchange signals. We then only need to include an electronic heads or tails display on both sides of the coin in order to reproduce the correlation. As soon as one of the coins is measured, it would then send a signal with the result to the other coin, which, in return, would display the corresponding face on both sides. This kind of correlation is generated *after* the measurement of the first coin via some signal. Through general relativity, however, we know that now information can be exchanged faster than with the speed of light. Consequently, if two very distant partners perform the experiment, the correlation should vanish. For example, it takes a light-beam several minutes to go from earth to mars. If two partners completed the experiments faster than that, there would be no way the two coins can exchange a signal. Therefore, this correlation would vanish. Here again, however, it was shown experimentally, that no signal could have been exchanged. In fact, the experiments showed that *if* there was a signal being exchanged between the two photons, it would have had to be *a lot* faster than the speed of light. Some experiments suggest for example at least 10^7 times faster than the speed of light. Taking the theory of general relativity as a fact, however, we must exclude the possibility of faster-than-light communication. Hence, we cannot explain this correlation via the exchange of (hidden) signals. In fact, this *Non-Signalling*⁴ also is a central ingredient of this quantum correlation,

⁴Note that it is *not* possible to use the present quantum correlation to transport information. From this point of view, it is perfectly consistent with general relativity.

usually referred to as *entanglement*.

3 Tensor product and entanglement

Let us now take it one step further and consider the formalism of quantum physics. First of all, let us denote the Hilbert space of the complete (composite) physical system by \mathcal{H}_{tot} . For simplicity, let us consider a system that is composed of two disjoint and distinguishable particles A and B described by their associated Hilbert spaces \mathcal{H}_A and \mathcal{H}_B . The Hilbert space structure of quantum mechanics, together with the superposition principle, tells us that a composite quantum system is described by the *tensor product* of the constituent's Hilbert spaces

$$\mathcal{H}_{\text{tot}} = \mathcal{H}_A \otimes \mathcal{H}_B. \quad (1)$$

That is, for any set of two basis $\{|\Phi_i\rangle_A\}$ and $\{|\Phi_j\rangle_B\}$ of the constituent's Hilbert spaces, the system's total Hilbert space is

$$\mathcal{H}_{\text{tot}} = \left\{ |\psi\rangle_{\text{tot}} \mid \exists \{c_{ij} \in \mathbb{C}\}, \sum_{i,j} |c_{ij}|^2 = 1 : |\psi\rangle_{\text{tot}} = \sum_{i,j} c_{ij} |\Phi_i\rangle_A |\Phi_j\rangle_B \right\}. \quad (2)$$

A direct consequence of this definition is that the dimension of the resulting Hilbert space \mathcal{H}_{tot} is not the sum of the dimensions of the subspaces but their product. For this reason, it is clear that the dimension of a Hilbert space associated to a large composite system does not grow linearly with the number of constituents but exponentially. While this fact is great for (theoretically possible) quantum computation, it is *the* major problem when trying to simulate condensed matter systems or even smaller quantum systems numerically.

Investigating the properties of the tensor product, we see that every combination of states from the subsystems, $|\phi\rangle_A$ and $|\varphi\rangle_B$ is represented by a state $|\psi\rangle_{\text{tot}}$ of the composite system. In fact, expanding to their bases, we write

$$|\phi\rangle_A = \sum_i a_i |\Phi_i\rangle_A \quad , \quad |\varphi\rangle_B = \sum_j b_j |\Phi_j\rangle_B \quad (3)$$

and

$$|\psi\rangle_{\text{tot}} = \sum_{i,j} c_{ij} |\Phi_i\rangle_A |\Phi_j\rangle_B. \quad (4)$$

With this in mind – and reviewing some common ways to write tensor a tensor product – we are able to derive the representation of this state in the

composite Hilbert space

$$|\psi\rangle_{\text{tot}} = |\phi\rangle_A \otimes |\varphi\rangle_B \equiv |\phi\rangle_A |\varphi\rangle_B \equiv |\phi\rangle |\varphi\rangle \equiv |\phi, \varphi\rangle = \quad (5)$$

$$= \left(\sum_i a_i |\Phi_i\rangle_A \right) \otimes \left(\sum_j b_j |\Phi_j\rangle_B \right) = \quad (6)$$

$$= \sum_{i,j} a_i b_j |\Phi_i\rangle_A |\Phi_j\rangle_B \quad (7)$$

Comparing equation (7) to the definition of a tensor product state (4), we see that it is enough to set

$$c_{ij} = a_i b_j. \quad (8)$$

in order to describe *any* composite state that is formed of two pure states in terms of the tensor product expansion. States that can be written as $|\psi\rangle_{\text{tot}} = |\phi\rangle_A |\varphi\rangle_B$ are called *product states*. Note that the maximal number of (complex) parameters describing any product state is $\dim(\mathcal{H}_A) + \dim(\mathcal{H}_B)$. Since $\dim(\mathcal{H}_A)$ and $\dim(\mathcal{H}_B)$ are natural numbers, we see that the set of all product states is only a – potentially marginal – subset of the whole tensor-space \mathcal{H}_{tot} because, as we have seen above,

$$\dim(\mathcal{H}_{\text{tot}}) = \dim(\mathcal{H}_A) \cdot \dim(\mathcal{H}_B) \geq \dim(\mathcal{H}_A) + \dim(\mathcal{H}_B). \quad (9)$$

From this we infer that there are there exist states of the composite system that *cannot* be described as a certain state in subsystem A and a certain state in B but only in terms of the whole system. These states have no classical analogue, the correlation resulting from this need for the composite description (i.e. the fact that it is not possible to write them as in equation (5)) is called *entanglement*. Therefore, a state that cannot be written as a product of pure states of the respecting subsystem is called an *entangled state*.

4 Schmidt decomposition

Now that we know how to write a composite system in quantum mechanical notation, let us come back to our initial example of an entangled state. In fact, the quantum correlation we described in Section 2 is nothing else but an intuitive characterisation of an entangled system. The example given in that section is even a very special entangled state: it's one of the four

maximally entangled states for a two-qubit⁵ system, a *singlet state*⁶. Using the mathematical tools introduced in Section 3, a system of two quantum coins that always show the same result is, for example, represented by

$$|\psi\rangle_{\text{tot}} = \frac{1}{\sqrt{2}} (|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B), \quad (10)$$

where the factor $1/\sqrt{2}$ is assuring the normalisation of the state. It can easily be verified that the measurement of either of the subsystems will always be completely random with same probabilities for result 0 and 1. Furthermore, we know that any quantum state collapses during a measurement to an eigenstate of the corresponding operator. Therefore, as soon as one qubit is measured, the other qubit will instantaneously collapse to the same value.

It is important to understand entangled states *cannot* be written as product states. For example, it is *not* possible to write $|\psi\rangle_{\text{tot}}$ as the tensor product of two states $|\phi\rangle_A$ of equation (10) and $|\varphi\rangle_B$. When considering small systems, it is, of course, possible to verify this claim by just trying all the possibilities. Furthermore, we can think that there are may be more complicated entangled states that require more than two terms. Generally speaking, every (pure) state $|\psi\rangle_{\text{tot}}$ can be written as

$$|\psi\rangle_{\text{tot}} = \sum_{i=1}^N \lambda_i |\Phi'_i\rangle_A |\Phi'_i\rangle_B, \quad (11)$$

where $N = \min(\dim(\mathcal{H}_A), \dim(\mathcal{H}_B))$, $\langle \Phi'_i | \Phi'_j \rangle = \delta_{i,j}$ (where $\delta_{i,j}$ is the Kronecker Delta-Function) and $\lambda_i \in \mathbb{R} \forall i$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ and $\sum_i \lambda_i^2 = 1$. Note that the basis vectors $\{|\Phi'_i\rangle\}$ are, in fact, summed over the same index (that is not a typo!) and that they are not, in general, the same as $\{|\Phi_i\rangle\}$ in Section 3. The decomposition of a system state into the minimum number of terms using orthonormal bases of the constituting systems is called *the Schmidt decomposition* (SD). The coefficients $\{\lambda_i\}$ are called *Schmidt coefficients* and $\{|\Phi'_i\rangle\}$ are the *Schmidt vectors*. There are different ways to prove the fact that we can always write a quantum state as in equation (11). One of the most elegant one is based on the singular value decomposition (SVD) known from linear algebra. In fact, since everything that follows relies on the Schmidt decomposition, we will derive SD from SVD here.

⁵For our needs, it is enough to think of a qubit as a polarized photon. In general, a qubit is a quantum physical two-level system that can take one of two reference values (denoting zero and one) as well as any linear superposition of these two states. In the case of photon polarisation, $|0\rangle$ may correspond to vertical polarisation and $|1\rangle$ to horizontal polarisation. Diagonal polarisation (with an angle of $\pm 45^\circ$) is usually described by $|\pm\rangle = \frac{1}{\sqrt{2}} (|0\rangle \pm |1\rangle)$.

⁶The four singlet states are: $\frac{1}{\sqrt{2}} (|0\rangle_A |0\rangle_B \pm |1\rangle_A |1\rangle_B)$ and $\frac{1}{\sqrt{2}} (|+\rangle_A |+\rangle_B \pm |-\rangle_A |-\rangle_B)$

The Theorem of the Singular Value Decomposition states that every linear operator, represented by a complex $m \times n$ matrix \mathbf{A} can be written (decomposed) as

$$\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{V}^+, \quad (12)$$

where

- $\mathbf{D} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ is a diagonal matrix of dimension $r = \min(m, n)$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$, where σ_i are the singular values of \mathbf{A} .
- \mathbf{U} and \mathbf{V} are complex unitary matrices of dimensions $m \times r$ and $r \times n$. The first r columns of \mathbf{V} are called the right singular vectors and the first r columns of \mathbf{U} the left singular vectors.

For the sake of clarity, we will denote the elements of \mathbf{U} and \mathbf{V} as u_{ij} and v_{ij} . The diagonal elements of \mathbf{D} will be denoted $d_{ii} \equiv \sigma_i$ as defined above. Keep in mind that $u_{ij}, v_{ij} \in \mathbb{C}$, $\forall i, j$ and that $d_{ii} \in \mathbb{R}$ with $d_{ii} \geq 0 \forall i$. Using these notations, we can re-write equation (12) as

$$a_{ij} = \sum_{k=1}^r u_{ik} d_{kk} (v_{jk})^*, \quad (13)$$

where $(v_{jk})^*$ is the complex conjugate of v_{jk} . Starting from here, recall that every state $|\psi\rangle_{\text{tot}}$ of the composite system can be written as

$$|\psi\rangle_{\text{tot}} \stackrel{(4)}{=} \sum_{i,j} c_{ij} |\Phi_i\rangle_A |\Phi_j\rangle_B \stackrel{(13)}{=} \sum_{i,j,k} u_{ik} d_{kk} (v_{jk})^* |\Phi_i\rangle_A |\Phi_j\rangle_B \quad (14)$$

$$= \sum_k d_{kk} \cdot \left(\sum_i u_{ik} |\Phi_i\rangle_A \right) \otimes \left(\sum_j (v_{jk})^* |\Phi_j\rangle_B \right) \quad (15)$$

It is now very natural to set

$$\lambda_k = d_{kk} \equiv \sigma_k \quad , \quad |\Phi'_k\rangle_A = \sum_i u_{ik} |\Phi_i\rangle_A \quad , \quad |\Phi'_k\rangle_B = \sum_j (v_{jk})^* |\Phi_j\rangle_B. \quad (16)$$

Because \mathbf{U} and \mathbf{V} are unitary matrices, the scalar products between vectors of the same basis before and after application of these operators remain the same. For this reason, if $\{|\Phi_i\rangle\}$ is an orthonormal basis, $\{|\Phi'_i\rangle\}$ will be one as well. Writing $N = r$, we can therefore conclude that

$$|\psi\rangle_{\text{tot}} \stackrel{(15)}{\stackrel{(16)}}{=} \sum_{i=1}^N \lambda_i |\Phi'_i\rangle_A |\Phi'_i\rangle_B,$$

q.e.d. \blacklozenge

The *Schmidt rank* χ is defined as the number of non-zero singular values. In other words, $\chi \leq N$ is the number of terms that will appear in the summation of the Schmidt Decomposition,

$$|\psi\rangle_{\text{tot}} = \sum_{i=1}^N \lambda_i |\Phi'_{i/A}\rangle |\Phi'_{i/B}\rangle = \sum_{i=1}^{\chi} \lambda_i |\Phi'_{i/A}\rangle |\Phi'_{i/B}\rangle. \quad (17)$$

The Schmidt rank χ will be of utmost importance for our simulation algorithm. If the subsystems are not of the same dimension, the collection of the $\{|\Phi_i\rangle_X\}$ of the larger-dimensional subspace do not span the whole subspace of the corresponding system. Mathematically speaking,

$$\text{if } N < \dim(\mathcal{H}_X), \quad \sum_{i=1}^N |\Phi_i\rangle_X \langle \Phi_i| \neq \mathbf{1}. \quad (18)$$

where the index X stands for the larger-dimensional subsystem A or B . The same argument holds even for the smaller-dimensional subsystem, if we only consider the first χ out of all N basis vectors $\{|\Phi_i\rangle\}$. The base may not be complete.

So far so good. We know that we can always decompose a pure state of a composite system into a linear combination of some (particular) basis vectors of the subspaces. But what is all this good for? Well, first of all, recall that a product state can always be written as $|\psi\rangle_{\text{tot}} = |\phi\rangle_A |\varphi\rangle_B$. Hence, for a product state, $\chi = 1$. From this perspective, we can therefore think of χ as some measure of the entanglement between the two subsystems⁷. Conversely, the more terms we need in the Schmidt decomposition, i.e. the higher the Schmidt rank χ , the more entanglement there is between the two subsystems.

Well then, now that we know why the Schmidt decomposition is so important, how can we compute it? After all, we guessing the right decomposition may be quite difficult! Consider the density matrix $\rho \equiv \rho_{\text{tot}}$ of the total system. For better readability, we will omit the index tot from now on. From equation (4) we can compute the corresponding density matrix for the pure state of the complete system

$$\rho = |\psi\rangle\langle\psi| \stackrel{(4)}{=} \left(\sum_{i,j} c_{ij} |\Phi_i\rangle_A |\Phi_j\rangle_B \right) \left(\sum_{i',j'} (c_{i'j'})^* \langle\Phi_{i'}|_A \langle\Phi_{j'}|_B \right), \quad (19)$$

⁷In fact, χ itself is not exactly suited for measuring the amount of entanglement as it does not scale linearly with the system size. A product state – i.e. a state without *any* entropy – has $\chi = 1$. Therefore, it is more suitable to take the entropy of the different Schmidt coefficients $H(\{\lambda_i\}) = -\sum_{i=1}^{\chi} \lambda_i \log(\lambda_i)$ as a measure of entanglement. In particular, this definition assures that the measure of a product state (one single λ which is equal to 1) is 0 and that the entanglement reaches its maximum if all terms of Schmidt decomposition contribute with the same amplitude.

which expands to

$$\rho = \sum_{i,j,i',j'} c_{ij} (c_{i'j'})^* |\Phi_i\rangle_A \langle \Phi_{i'}| \otimes |\Phi_j\rangle_B \langle \Phi_{j'}|. \quad (20)$$

Of course, it is not quite obvious how to get from this form to the Schmidt decomposition. Recalling, however, that the same ρ can be written in terms of the Schmidt decomposition of the state, we note

$$\rho \stackrel{(11)}{=} \left(\sum_{i=1}^N \lambda_i |\Phi'_i\rangle_A \langle \Phi'_i| \right) \left(\sum_{i'=1}^N \lambda_{i'} |\Phi'_{i'}\rangle_B \langle \Phi'_{i'}| \right), \quad (21)$$

which, again, can be expanded

$$\rho = \sum_{i,i'} \lambda_i \lambda_{i'} |\Phi'_i\rangle_A \langle \Phi'_{i'}| \otimes |\Phi'_{i'}\rangle_B \langle \Phi'_i|. \quad (22)$$

At this point, we have to introduce a new concept: the *partial trace*, denoted by tr_X . Recall that, in linear algebra, the trace of a $n \times n$ matrix M is usually defined as

$$\text{tr}(M) = \sum_{i=1}^n \langle \Phi_i | M | \Phi_i \rangle, \quad (23)$$

where $\{|\Phi_i\rangle\}$ is an orthonormal basis of the vector-space. Note that the definition of the trace, equation (23), implies that its value does not depend on the choice of the (orthonormal) base. In particular, if the matrix elements of M are denoted by m_{ij} ,

$$\text{tr}(M) = \sum_{i=1}^n m_{ii}, \quad (24)$$

Now, if our vector space is the Hilbert space of a composite system, the trace over an operator M is

$$\text{tr}(M) = \sum_{i=1}^n (\langle \Phi_i | \langle \Phi_i | M | \Phi_i \rangle | \Phi_i \rangle). \quad (25)$$

Written in this form, it appears quite natural to define the partial trace as the trace over a partial system. Hence, we write the trace over subsystem A as

$$\text{tr}_A(M) = \sum_{i=1}^n \langle \Phi_i | \langle \Phi_i | M | \Phi_i \rangle_A = \sum_{i=1}^n \langle \Phi'_i | \langle \Phi'_i | M | \Phi'_i \rangle_A, \quad (26)$$

because we can chose any orthonormal basis of \mathcal{H}_A . Equivalently, the trace over subsystem B is

$$\text{tr}_B(M) = \sum_{i=1}^n \langle \Phi_i | \langle \Phi_i | M | \Phi_i \rangle_B = \sum_{i=1}^n \langle \Phi'_i | \langle \Phi'_i | M | \Phi'_i \rangle_B. \quad (27)$$

Note that while the (total) trace as defined in equation (23) is a scalar, the partial trace is, in general, a matrix. In particular, the partial trace over a all but one subsystem of the density matrix of a composite system corresponds to the density matrix of that subsystem alone. Therefore, the physical interpretation of the partial trace corresponds to incomplete knowledge of the system. For example, imagine that the density matrix in equation (22) corresponds to a system of two particles that are in different labs. In this case, $\rho_A = \text{tr}_B(\rho)$ describes the state as the experimentalist in lab A would see the system and $\rho_B = \text{tr}_A(\rho)$ describes the particle in lab B . From this point of view, we will always *trace out* the parts of the system that we cannot access.

Apart from its physical significance, the partial trace can also be a powerful tool for computing Schmidt decompositions. Consider, again, equation (22). Because of the orthonormality of the bases, the density matrices corresponding to particles A and B are

$$\rho_A = \text{tr}_B(\rho) \stackrel{(22)}{\underset{(26)}{=}} \sum_i \lambda_i^2 |\Phi'_i\rangle_A \langle \Phi'_i|, \quad (28)$$

$$\rho_B = \text{tr}_A(\rho) \stackrel{(22)}{\underset{(27)}{=}} \sum_i \lambda_i^2 |\Phi'_i\rangle_B \langle \Phi'_i|. \quad (29)$$

These two expressions tell us how to compute either *one* of the basis required for the Schmidt decomposition of an arbitrary composed system of the form (20): all we have to do is to compute the partial trace of the density matrix and then find the eigensystem of the reduced density matrix. After all, we know that any matrix can be written in the form of its spectral representation, i.e. in terms of its eigensystem,

$$\mathbb{M}|\Phi'_i\rangle = e_i|\Phi'_i\rangle \iff \mathbb{M} = \mathbb{M} \underbrace{\left(\sum_i |\Phi'_i\rangle \langle \Phi'_i| \right)}_{=\mathbb{1}} = \sum_i e_i |\Phi'_i\rangle \langle \Phi'_i|. \quad (30)$$

For this reason, the set of the Schmidt coefficients $\{\lambda_i\}$ coincides with the set of the square roots of the eigenvalues of either of the reduced density matrices. The requirement for real Schmidt coefficients is fulfilled because of the fact that ρ is a hermitian matrix. Moreover, we know that the eigenvalues of ρ correspond to the probabilities to measure the system in state described by the corresponding eigenvector. Hence, it is enough to order the eigenvalues according to their amplitude to find the Schmidt coefficients

$$e_1 \geq e_2 \geq \dots \geq e_N \geq 0 \stackrel{(28),(29)}{\underset{(30)}{\implies}} \lambda_i = \sqrt{e_i}, \text{ with } \sum_{i=1}^N \lambda_i^2 = 1. \quad (31)$$

When computing the Schmidt vectors, note that one *cannot* compute the Schmidt basis of *both* subsystems with this method. Through equations (26) and (27), the basis vectors $\{|\Phi'_i\rangle\}$ are only defined up to a complex phase factor. Usually, the global phase of a state does not play any role in quantum mechanics, $|\psi\rangle \equiv e^{i\theta}|\psi\rangle$. Here, on the contrary, the global state will be formed by the (tensor) *product* of two basis states. Hence⁸,

$$|\psi\rangle \equiv e^{i\theta}|\psi\rangle = e^{i\theta} \sum_{j=1}^N \lambda_j |\Phi'_j\rangle_A |\Phi'_j\rangle_B \neq \sum_{j=1}^N \lambda_j (e^{i\alpha_j} |\Phi'_j\rangle_A) (e^{i\beta_j} |\Phi'_j\rangle_B), \quad (32)$$

because, in general, $e^{i\theta} \neq e^{i\alpha_j} e^{i\beta_j} \forall i, j$. With this in mind, we see the importance to compute *the other* Schmidt basis in another way – in a way that preserves the complex phase factor. A straightforward possibility is the projection onto the other basis,

$$\lambda_j |\Phi'_j\rangle_B =_A \langle \Phi'_j | \psi \rangle \quad \text{or} \quad \lambda_j |\Phi'_j\rangle_A =_B \langle \Phi'_j | \psi \rangle. \quad (33)$$

5 Simulation algorithm (Vidal)

The algorithms we would like to discuss is based on the idea, that we can try to approximate a composite state by a state that has an upper bound of entanglement. In practice, this is done by only considering a certain number $\chi_\varepsilon \leq \chi$ terms of the Schmidt decomposition. This way, the computational resources needed for numerical simulation no-longer grow exponentially. For this reason, we can thus try to simulate larger systems using the same (limited) computational equipment. Numerical experiments that have already been performed in different groups seem to indicate that this kind of approach is best suited for linear systems (e.g. spin chains) that are submitted to next-neighbours interactions⁹.

Intuitively, the algorithm proposed by Guifré Vidal consists of considering every possible Schmidt decomposition of a chain of particles and then connecting them via some tensors. One-qubit gates are then implemented by considering only the tensor representing the specific gate. Two-qubit next-neighbours gates affect the two corresponding tensors as well as the corresponding Schmidt coefficient. Hence, this construction allows for a considerable reduction of the number of variables needed to describe the

⁸We use i to denote the complex unit, i.e. $i^2 = -1$. As a consequence, we sum over j to avoid confusion.

⁹Recall that a one-qubit gate is just a unitary operation that can be performed onto a single qubit (for example rotation) and does not change the value of any other qubit. Two-qubit next-neighbours gates are unitary operation that involve two qubits that are nearest neighbours on the chain of particles. In the case of a spin chain, for example, it might involve spin number 5 and spin number 6, leaving all other spins unaffected.

system while preserving a good manipulability in terms of one-qubit and two-qubit next-neighbours gates.

So much for the intuition. Mathematically, we consider a one-dimensional spin-chain of n spin- $\frac{1}{2}$ particles. We label them from 1 to n and define the *computational basis* as being the basis states according to state $|0\rangle$ and $|1\rangle$ of each particle. Hence, writing $|i\rangle$ for a given state, i can either be 0 or 1. The algorithm then starts by performing all possible Schmidt decompositions of the linear system: subsystem A contains the first l particles and subsystem B the rest of them. For every partition $A : B$ we then have

$$|\Psi\rangle = \sum_{\alpha=1}^{\chi_A} \lambda_{\alpha} |\Phi_{\alpha}^{[A]}\rangle |\Phi_{\alpha}^{[B]}\rangle. \quad (34)$$

After having performed this Schmidt decomposition for all possible bipartite splittings $A : B$, we can unify our notation by introducing χ as the maximal value of all χ_A . That is

$$\chi = \max_A \chi_A. \quad (35)$$

Now, consider the most general configuration of the state of the complete spin chain,

$$|\Psi\rangle = \sum_{i_1=0}^1 \sum_{i_2=0}^1 \cdots \sum_{i_n=0}^1 c_{i_1, i_2, \dots, i_n} |i_1\rangle |i_2\rangle \cdots |i_n\rangle. \quad (36)$$

This description includes $\mathcal{O}(2^n)$ independent parameters¹⁰ c_{i_1, i_2, \dots, i_n} . The first idea in this algorithm is that it is always possible to write the coefficients c_{i_1, i_2, \dots, i_n} as something like a generalised Schmidt decomposition. This description uses n tensors, $\{\Gamma^{[1]}, \Gamma^{[2]}, \dots, \Gamma^{[n]}\}$, and $n - 1$ vectors, $\{\lambda^{[1]}, \lambda^{[2]}, \dots, \lambda^{[n-1]}\}$ and reads

$$c_{i_1, i_2, \dots, i_n} = \sum_{\alpha_1, \alpha_2, \dots, \alpha_{n-1}} \Gamma_{\alpha_1}^{[1]i_1} \lambda_{\alpha_1}^{[1]} \Gamma_{\alpha_1 \alpha_2}^{[2]i_2} \lambda_{\alpha_2}^{[2]} \cdots \Gamma_{\alpha_{l-1} \alpha_l}^{[l]i_l} \lambda_{\alpha_l}^{[l]} \cdots \Gamma_{\alpha_{n-1}}^{[n]i_n}. \quad (37)$$

The indices i and α take values in $\{0, 1\}$ and $\{1, 2, \dots, \chi\}$, respectively. Hence, the $\mathcal{O}(2^n)$ coefficients, $\{c_{i_1, i_2, \dots, i_n}\}$, are now reexpressed in terms of $\mathcal{O}((2\chi^2 + \chi)n)$ new variables.

This is where the second idea – the key ingredient – of this algorithm comes into play. As we have seen in section 4, the number χ is directly related to the entanglement present in the system. A system that requires a large χ includes strong correlation between (at least some of) its particles. Thus,

¹⁰In fact, there are 2^n complex coefficients c_{i_1, i_2, \dots, i_n} that are, however, not all independent. Normalisation and the freedom to chose any global phase slightly reduce the number of parameters.

fixing (the maximal) value of χ to a rather small value χ_ε corresponds to limiting the (maximal) entanglement of our system. This way, the state described by equations (36) and (37) is approximated by (potentially) far less variables, $\mathcal{O}((2\chi_\varepsilon^2 + \chi_\varepsilon)n)$. For example, the dimension of the Hilbert space of a one-dimensional spin-chain of $n = 50$ particles is $\mathcal{O}(2^n) \approx \mathcal{O}(10^{15})$. In a computer, a real floating point variable is usually described by (at least) four bytes¹¹. Hence, storing all these variables in a computer would require approximately 40'000 GBytes of memory. And that is just for the description of the state! If we restrict ourselves to slightly entangled systems, we can simulate the same system with a lot less variables. Taking $\chi_\varepsilon = 20$, we would need $\mathcal{O}((2\chi_\varepsilon^2 + \chi_\varepsilon)n) = \mathcal{O}(4.1 \cdot 10^4)$ variables. In our example, this would correspond to about 165 kBytes, or as little as 0.000165 GBytes¹²! Furthermore, as this approach scales linearly with the length of the spin-chain, we can even go beyond 50 particles without unreasonably great expenditure.

So what exactly does our approximation consist of? Intuitively, we are approximating a certain state by the first $\chi_\varepsilon < \chi$ (potentially even $\chi_\varepsilon \ll \chi$) terms of its Schmidt decomposition,

$$|\Psi\rangle \stackrel{(34)}{\equiv} \sum_{\alpha=1}^{\chi} \lambda_\alpha |\Phi_\alpha^{[A]}\rangle |\Phi_\alpha^{[B]}\rangle \approx |\psi\rangle = \sum_{\alpha=1}^{\chi_\varepsilon} \lambda_\alpha |\Phi_\alpha^{[A]}\rangle |\Phi_\alpha^{[B]}\rangle. \quad (38)$$

Now consider the fact that the basis vectors $\{|\Phi_\alpha^{[i]}\rangle\}$ are normalised. Hence, for example, if the last $\chi - \chi_\varepsilon$ terms that we neglect all have $\lambda_\alpha = 0$, then the approximated state $|\psi\rangle$ is equal to $|\Psi\rangle$. In fact, since we have ordered $\{\lambda_\alpha\}$ according to their amplitude, i.e. $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{\chi_\varepsilon} \geq \dots \geq \lambda_\chi \geq 0$, we have

$$|\Psi\rangle = |\psi\rangle \Leftrightarrow \lambda_\alpha = 0, \forall \alpha \in \{(\chi_\varepsilon+1), \dots, \chi\} \Leftrightarrow \lambda_{(\chi_\varepsilon+1)} = 0. \quad (39)$$

Refining our argument, recall that the norm of a state is equal to the sum of the square of its schmidt coefficients. Therefore, we will say that our algorithm is a good approximation whenever its norm is close to one. In mathematical terms,

$$|\Psi\rangle \approx |\psi\rangle \Leftrightarrow |\langle \psi | \psi \rangle|^2 \approx 1 \Leftrightarrow \sum_{\alpha=\chi_\varepsilon+1}^{\chi} |\lambda_\alpha|^2 \ll 1 \quad (40)$$

¹¹Note that 4 bytes = 32 bits is used to represent a single-precision floating-point to keep the argument as simple as possible. Usually, however, we implement our algorithms with double-precision complex floating point variables that employ 16 bytes.

¹²In actual implementations, it will be, of course, a lot more because of the double-precision complex variables. However, even if, on top of that, we have to store intermediate results and other variables in our memory, we are within the scope of today's hardware (i.e. < 1 GByte). In fact, we do not even need extremely sophisticated supercomputers to do these computations.

In fact, in real implementations of this loss of norm is often used to quantify (or at least to guess) the accuracy of the simulation¹³.

Of course, this whole algorithm is not of much help as long as we do not know how to compute the $\{\Gamma^{[l]}\}$ tensors and the $\{\lambda^{[l]}\}$ vectors. So here is how to compute them. Start at one end of the one-dimensional system and compute the Schmidt decomposition of the partition $A : B$ for which subsystem A is only the first particle while subsystem B is the rest of the chain. Indexing the basis vectors by the particles instead of the subsystem instead of by their symbol we have

$$|\Psi\rangle = \sum_{\alpha_1} \lambda_{\alpha_1}^{[1]} |\Phi_{\alpha_1}^{[1]}\rangle |\Phi_{\alpha_1}^{[2\dots n]}\rangle. \quad (41)$$

Expanding the left Schmidt vectors in the computational basis we have

$$|\Phi_{\alpha_1}^{[1]}\rangle = \sum_{i_1=0}^1 \Gamma_{\alpha_1}^{[1]i_1} |i_1\rangle, \quad (42)$$

which then leads to

$$|\Psi\rangle = \sum_{\alpha_1, i_1} \Gamma_{\alpha_1}^{[1]i_1} \lambda_{\alpha_1}^{[1]} |i_1\rangle |\Phi_{\alpha_1}^{[2\dots n]}\rangle. \quad (43)$$

Now, we can proceed by repeating the following three steps (start with $l = 2$)

1. Expand each Schmidt vector $|\Phi_{\alpha_1}^{[l\dots n]}\rangle$ in a local basis for qubit l .

$$|\Phi_{\alpha_{l-1}}^{[l\dots n]}\rangle = \sum_{i_l=0}^1 |i_l\rangle |\tau_{\alpha_{l-1}i_l}^{[(l+1)\dots n]}\rangle, \quad (44)$$

where the $\{|\tau_{\alpha_{l-1}i_l}^{[(l+1)\dots n]}\rangle\}$ are not necessarily normalised.

2. Now write each of these $|\tau_{\alpha_{l-1}i_l}^{[(l+1)\dots n]}\rangle$ vectors in terms of *at most* χ Schmidt vectors $\{|\Phi_{\alpha_l}^{[(l+1)\dots n]}\rangle\}_{\alpha_l=1}^{\chi}$ (i.e., the eigenvectors of $\rho^{[(l+1)\dots n]}$) and the corresponding Schmidt coefficients $\lambda_{\alpha_l}^{[l]}$,

$$|\tau_{\alpha_{l-1}i_l}^{[(l+1)\dots n]}\rangle = \sum_{\alpha_l} \Gamma_{\alpha_{l-1}\alpha_l}^{[l]i_l} \lambda_{\alpha_l}^{[l]} |\Phi_{\alpha_l}^{[(l+1)\dots n]}\rangle. \quad (45)$$

¹³To be precise: usually, one has to re-normalise during the computations (for numerical reasons amongst other things). Therefore, the control-parameter that is checked is the *fraction* of the loss in norm.

3. Finally, substitute equation (45) in equation (44) and the whole in equation (43). We then obtain

$$|\Psi\rangle = \sum_{\alpha_1, i_1, \dots, \alpha_l, i_l} \Gamma_{\alpha_1}^{[1]i_1} \lambda_{\alpha_1}^{[1]} \dots \Gamma_{\alpha_{l-1} \alpha_l}^{[l]i_l} \lambda_{\alpha_l}^{[l]} |i_1, \dots, i_l\rangle |\Phi_{\alpha_{(l)}}^{[(l+1)\dots n]}\rangle \quad (46)$$

After the steps are completed for all $l = 2, \dots, n$ we are able to express the state $|\Psi\rangle$ as in equation 37).

6 Computations

As we have seen, the main advantages of the algorithm introduced in section 5 is, that it is very efficient in terms of memory space. On top of that, it is highly efficient in terms of the number of operations required to perform hamiltonian evolutions of the system. In particular, we are often interested in *imaginary time evolution* as well as in *real time evolution*. Denoting the hamiltonian by H , the latter evolution is

$$|\psi_t\rangle = e^{iHt/\hbar} |\psi_0\rangle, \quad (47)$$

where $|\psi_t\rangle$ is the state at time t and $|\psi_0\rangle$ is the initial state at $t = 0$. Imaginary time evolution, as opposed to real time evolution, is used to compute the ground state of a system. In fact, consider the expansion of our state $|\psi\rangle$ into the eigenvectors, $\{|\phi_n\rangle\}$, of the hamiltonian that correspond to energy E_n ,

$$|\psi\rangle = \sum_n c_n |\phi_n\rangle. \quad (48)$$

Introducing an imaginary time $i \cdot t$, the state will evolve as

$$|\psi_{it}\rangle = e^{-Ht/\hbar} |\psi_0\rangle = \sum_n c_n e^{-E_n t/\hbar} |\phi_n\rangle \xrightarrow{n \rightarrow \infty} |\phi_0\rangle, \quad (49)$$

which is the ground state of the system¹⁴.

But how do we apply any kind of operator to our state? First of all, we need to understand that the structure of the decomposition is thus that one-qubit gates and two-qubit next-neighbours gates can easily be implemented. Gates that are not just between next-neighbours require a lot more operations.

One-qubit operations, represented by a matrix U_j^i , are performed by simply updating the corresponding $\Gamma^{[l]}$ tensor. In fact, it can be proven that

$$\Gamma_{\alpha\beta}^{[l]i} = \sum_{j=0}^1 U_j^i \Gamma_{\alpha\beta}^{[l]j} \quad \forall \alpha, \beta = 1, \dots, \chi. \quad (50)$$

¹⁴This evolution is, of course, *not* unitary. Therefore, re-normalisation in the algorithm is central.

The computational cost of such an operation is $\mathcal{O}(\chi^2)$.

Updating the decomposition of a state after application of a two-qubit next-neighbours gate V_{kl}^{ij} can be done with an overall computational cost of $\mathcal{O}(\chi^3)$. It includes $\Gamma^{[l]}$, $\lambda^{[l]}$ and $\Gamma^{[l+1]}$. From a general point of view, it is performed decomposing the spin-chain into four parts: the left side of the two particles involved in the operation, the left particle, the right particle and the right part of the chain. Then, the updated $\Gamma^{[l]}$, $\lambda^{[l]}$ and $\Gamma^{[l+1]}$ are computed by performing a Schmidt decomposition between {left side + left particle} and {right particle + right side}.

Since we are only considering hamiltonians involving next-neighbours interactions, we can also describe evolution of a state in a constructive manner. Generally speaking, a hamiltonian evolution (both real and imaginary) can be performed by decomposing the hamilton operator into two subsets of mutually commuting operators. Mathematically speaking, we can decompose the hamiltonian into

$$H = \sum_{\text{even } l} F^{[l]} + \sum_{\text{odd } l} G^{[l]}, \quad (51)$$

where $F^{[l]}$ and $G^{[l]}$ are two-qubit operators acting on qubits l and $l + 1$. This way, we can be sure that all F operators commute amongst themselves and that the same holds for all G operators. Note, however, that it is possible that $[F, G] \neq 0$. In this case, the evolution is computed iteratively via the Trotter expansion of $\exp(aH)$. For real time evolution, we would then set $a = it/\hbar$ whereas imaginary time evolution would be represented by $a = -t/\hbar$. Often, already second order expansion reproduces very good results. Therefore, this whole approach appears to be very powerful.