UNIVERSITY OF TRENTO

Department of Mathematics

Ph.D. Thesis

Algebraic, geometric and numerical methods for Tensor Network Varieties

Advisor:

Prof. Alessandra Bernardi

Co-advisor:

Prof. Iacopo Carusotto

Candidate:

Claudia De Lazzari

December 7, 2022

To my family.

Abstract

In this thesis, we study tensor network varieties, which are varieties of tensors described by the combinatorial structure of a given graph and two sets of integer weights, called bond and local dimensions, respectively. Tensor network varieties are geometric objects studied in the field of Algebraic Geometry, and they have received much attention in the recent years due to their usefulness in the field of Quantum Physics and other application areas. In the first part of the thesis, we study the dimension of tensor network varieties. We provide a completely general upper bound on their dimension and we give the exact value of the dimension in a particular range of parameters. We refine the upper bound in cases relevant for applications, such as matrix product states and projected entangled pairs states. We then focus on the study of the linear span of uniform matrix product states, which are translation invariant tensor network varieties associated to the cyclic graph. We provide nontrivial linear trace relations which prove the strict containment of the linear span in the ambient space as long as the number of sites is at least quadratic in the bond dimension, improving the state of the art. The outlined results are based on the papers [BDLG22, DLMS22a]. Finally, based on dimensional considerations, we propose a variation of the nonlinear conjugate gradient method used to approximate the ground states of a given Hamiltonian on the variety of matrix product states.

Acknowledgements

First of all, I thank my supervisors Alessandra Bernardi and Iacopo Carusotto for their inspiring commitment and their support in this beautiful research project. I acknowledge the support of GNSAGA of Indam. Moreover, this work has benefited from Q@TN, the joint lab between University of Trento, Fondazione Bruno Kessler, National Institute for Nuclear Physics and National Research Council. I thank all my colleagues of the transdisciplinary program. I would like to thank Matteo Rizzi for essential discussions about the physical interpretation of the mathematical objects of the thesis. I particular acknowledge goes to Fulvio Gesmundo for his support and wide expertise. I acknowledge Mateusz Michałek for inviting me to the Max Planck Institute in Leipzig, which has been a fruitful place where talking about mathematics. A special thank goes to Tim Seynnaeve, colleague and friend. I thank Jarosław Buczyński for his invitation to IMPAM in Warsaw and for the time he dedicated to discuss about my research. I acknowledge IPAM for the opportunity to participate in the TM2021 semester in which I could meet experts in the field of tensor network theory. I thank anyone involved in the Tensor Ph.D Seminar for all comments and suggestions, and for having shortened in some sense the distances imposed by the pandemic. I'm thankful to my colleagues Reynaldo and Pierpaola who shared several adventures with me during these years.

Inside and outside mathematics, there are several people that have been present in Trento and whom I want to thank. My friends Luca, Morena and Giovanni. My office mate Mattia. Mattia, who has always been there for me. My climbing friend Francesco. My flatmate Nicolò. My whole family. Finally, my family of Trento, Emanuele and Giuseppe.

Contents

Introduction

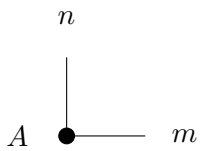
In this thesis, we study tensor network varieties, which are varieties of tensors described by the combinatorial structure of a given graph. The thesis is rooted in the field of algebraic geometry. Varieties of tensors are classical geometric objects studied in Algebraic Geometry, exploiting techniques from Representation Theory, Lie groups and Lie algebras Theory. The thesis is also framed in a transdisciplinary context due to the high impact that tensor network varieties are having in the field of Quantum Physics.

Tensors are algebraic objects that encode multidimensional arrays of numbers. They are present and studied in several research areas, from abstract algebra and algebraic geometry to applied mathematics and physics. Tensor network varieties, in particular, play a major role in quantum many-body physics, where they are used as a variational ansatz class to describe strongly correlated quantum systems whose entanglement structure is encoded in the underlying graph. The original motivation in quantum physics arises from the work of I. Affleck, T. Kennedy, E. H. Lieb and H. Tasaki [AKLT88] where they proved that the exact solution of ground states of the AKLT Hamiltonian admits a matrix product state representation. Matrix product states are tensor network varieties associated to the cyclic (or path) graph. The particular solution of the AKLT model suggested the use of matrix product states and other entanglement structures as variational ansatz classes either for the approximation of ground states or for the simulation of quantum many-body systems. We refer to [Orú14] for a full introductory description of the subject and to [CPGSV21] for a complete overview of theoretical results from the point of view of quantum physics. Tensor network varieties have nowadays a role in several other areas of applied mathematics. Since their original conception in quantum many-body physics, they have found a wide range of applications in different fields, such as numerical tensor calculus [Ose11, Hac19, CLO+16, BSU16], graphical models [Lau96, RS19] applied to phylogenetics [ERSS05, AR08], geometric complexity theory [LQY12, BC92, DMPY12] and machine learning [CLO+16, CCX+18, Ben09].

As a consequence, they have received much attention from the geometric and algebraic community in the recent years. Tensor network varieties were known in mathematics for specific classes of graphs; for example matrix product states with open boundary conditions were known as *tensor trains* [OT09, Ose11, HRS12, LOV15] and more generally, tree tensor networks as tensor admitting a *hierarchical* format [Gra10, Hac12, BBM15,

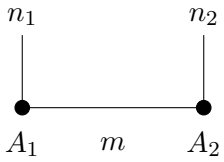
STG+19]. We dare to say that tensor network varieties appeared in mathematics under this unifying name in [LQY12], where a number of basic geometric questions were answered, providing several insights. Methods from differential and complex geometry were introduced in the study of these objects in [HMOV14] and more recently some important developments were achieved using methods from algebraic geometry and representation theory [BBM15, MSV19, GLW18, CLVW20, CGFmcW21, HGS+20].

In tensor network theory, we pictorially represent a matrix $A \in \mathbb{C}^{n\times m}$, i.e. a tensor of order two, as follows:



We have a vertex to which are associated two edges. The edges represent the indices of rows and columns of the matrix A, respectively.

Elements of a tensor network variety are networks of tensors contracted accordingly to the edge structure of an underlying graph. The order of the involved tensors and the contractions along the edges are prescribed by two chosen collections of integer weights, called *bond* and *local* dimensions, respectively. Local and bond dimensions are associated to the vertices and to the edges of the graph, respectively. As an example, we describe the tensor network variety associated to the path graph on two vertices. Consider the path graph on two vertices. Associate to every vertex a local dimension n_k, $k = 1, 2$, and to the unique edge a bond dimension m. Imagine now placing a tensor $A_1 \in \mathbb{C}^{n_1\times m}$ on the first vertex and a tensor $A_2 \in \mathbb{C}^{m\times n_2}$ on the second vertex. Notice that the order of the k-th tensor is prescribed by the local and bond dimensions. Since the underlying graph is the path graph, the two tensors of the network are connected by the shared edge, as in the following picture:



The edge connecting the two tensors pictorially represents a *contraction map* applied to the tensor product $A_1 \otimes A_2$. More precisely, let $\phi : (\mathbb{C}^{n_1} \otimes \mathbb{C}^{m*}) \otimes (\mathbb{C}^m \otimes \mathbb{C}^{n_2}) \to \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2}$ be the contraction map, which contracts $\mathbb{C}^{m*}$ with its dual copy $\mathbb{C}^m$. Let $\{e^{(k)}_{i_k}\}_{i_k=1,\dots,n_k}$ be the canonical basis of $\mathbb{C}^{n_k}$. Then, an element of the tensor network

variety $\phi(A_1 \otimes A_2) \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2}$ is given by the expression

$$\phi(A_1 \otimes A_2) = \sum_{i_1,i_2=1}^{n_1,n_2} \sum_{\alpha=1}^{m} \left(A_1{}_{\alpha}^{i_1} A_2{}_{\alpha}^{i_2}\right) e_{i_1}^{(1)} \otimes e_{i_2}^{(2)} = A_1 A_2.$$

At the level of coordinates, the contraction corresponds to a summation over the shared index of the two tensors, associated to the linking edge, i.e. a summation over the raw index of A_1 and the column index of A_2. The tensor network variety associated to the path graph on two vertices is actually the variety of matrices with rank bounded by the bond dimension m. General tensor network varieties are constructed in an analogous way, once we choose a simple and undirected graph and the collections of local and bond dimensions.

Dimension of tensor network varieties. From an algebraic geometric perspective, the tensor network variety is defined as either the Zariski or Euclidean closure of the image of the contraction map, which is a polynomial map in the entries of the tensors composing the network. Taking the closure is necessary since it has been proved that the set of tensors allowing for a tensor network representation is not always a closed set [LQY12, CLVW20, BLF22]. Given an algebraic variety, or more generally a geometrical object, one of its fundamental properties is the *dimension*. A complete result on the dimension of tensor network varieties is given in [HMOV14] for matrix product states with open boundary conditions. W. Hackbusch suggested [Hac12] the problem of comparing the complexities of different tensor network encodings of the same tensor. In [BBM15] the question is answered when the compared underlying graphs are the perfect binary tree and the train track tree, which correspond to the hierarchical format and matrix product states format, respectively. Also in [YL18], a comparison between tensor network varieties corresponding to different underlying graphs is proposed. In both papers the problem of dimension is addressed; in particular they provide the dimensions of the associated varieties in specific ranges of parameters.

We investigate the problem of the dimension of tensor network varieties in Chapter 3. We introduce tensor network varieties via the language of graph tensors, following [VC17, CVZ19] and we show that the construction is equivalent to the usual one given in terms of the contraction map described above. Then we provide a natural parametrization of an open set of the tensor network variety and we give a completely general *upper bound* on the dimension of any tensor network variety in Theorem 3.0.2. Moreover, we give the *exact* value of the dimension in a particular range of parameters, where the variety is realized as the closure of the orbit of the action of an algebraic group. This result in particular generalizes the previously known results.

Our main result of Chapter 3 is the following:

Let Γ be a simple graph with vertex set $\mathbf{v}(\Gamma)$ of cardinality d and edge set $\mathbf{e}(\Gamma)$. Write $\mathbf{n} = (n_1, \ldots, n_d)$ for the collection of local dimensions, associated to the vertices of the

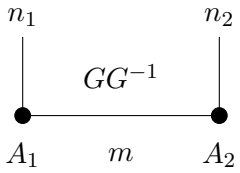
graph, and $\mathbf{m} = (m_e : e \in \mathbf{e}(\Gamma))$ for the collection of bond dimensions, associated to the vertices of the graph.

Theorem 3.0.2. *Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network and let $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ be the corresponding tensor network variety. Then*

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} \leq$$
$$\min\left\{ \sum_{v\in\mathbf{v}(\Gamma)} (n_v \cdot \prod_{e\ni v} m_e) - d + 1 - \sum_{e\in\mathbf{e}(\Gamma)} (m_e^2 - 1) + \dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X),\ \prod_{v\in\mathbf{v}(\Gamma)} n_v \right\}.$$

The upper bound on the dimension is obtained, following the classical Theorem of Dimension of the Fibers [Sha94], by determining a *lower* bound on the dimension of the generic fiber of the parametrization of the variety. When the variety does not fill the ambient space, the bound is determined by three factors. The term $\sum_{v\in\mathbf{v}(\Gamma)}(n_v \cdot \prod_{e\ni v} m_e) - d + 1$ is a parameter count of the domain of the parametrisation. With $\mathcal{G}_{\Gamma,\mathbf{m}}$ we denote the *gauge subgroup*, the term $\sum_{e\in\mathbf{e}(\Gamma)}(m_e^2 - 1)$ is its dimension and $\mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ is the stabilizer of a generic d-tuple of linear maps, under the action of the gauge subgroup. The gauge subgroup is a very important object in our study of the dimension of tensor network varieties. We can prove that the orbit of a generic element of the domain of the parametrization, under the action of the gauge subgroup, is contained in the generic fiber of the map. Consequently, the dimension of the generic fiber is bounded from below by the dimension of the gauge orbit, which is exactly the dimension of the gauge subgroup minus the dimension of the stabilizer, as can be seen in the formula of our Theorem 3.0.2.

We give here an insight into the construction of the gauge subgroup. It is defined as a product of (projectivised) general linear groups, each one associated to an edge of the graph. Consider again the path graph on two vertices. Let $G \in GL_m$ be an invertible matrix. Consider the unique edge of the path graph. We can imagine placing on it the product of the matrix and its inverse, as in the following picture:



The insertion of $GG^{-1} = \mathrm{Id}_m$ leaves the element of the tensor network invariant:

$$\phi(A_1 \otimes A_2) = A_1 A_2 = A_1 G G^{-1} A_2 = \phi(A_1 G \otimes G^{-1} A_2).$$

The action of the gauge subgroup GL_m on the parameter space of the map ϕ is defined precisely as $A_1 \mapsto A_1 G$ and $A_2 \mapsto G^{-1} A_2$, which clearly has the map ϕ constant along its orbits. Again, the construction can be generalized to every given graph.

The role of the gauge subgroup in the theory of tensor network was known [HMOV14] and it is expected that it entirely controls the value of $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$. In fact, it is foreseen that in "most" cases the exact value of the dimension is

$$\min\left\{\sum_{v\in\mathbf{v}(\Gamma)}(n_v\cdot\prod_{e\ni v}m_e)-d+1-\sum_{e\in\mathbf{e}(\Gamma)}(m_e^2-1),\ \prod_{v\in\mathbf{v}(\Gamma)}n_v\right\}, \tag{1}$$

meaning that the generic fiber of the tensor network parametrization contains only the gauge subgroup. We can refine our upper bound in cases relevant for applications, such as matrix product states and projected entangled pairs states (their 2-dimensional generalizations), proving that $\dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ is trivial. However, we observe that for matrix product states with bond dimension two and a low number of sites (in which $\dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ is trivial), the value in (1) provides a *strict* upper bound. This in particular implies that, at least in these cases, the gauge orbit does not fill completely the fiber. We further analyze these cases and we provide a more precise calculation of their dimension. These results are interesting since they involve very relevant tensor network varieties, relatively "small", and because their "unexpected" dimensions were not known. Based on numerical computations, we finally give a conjecture, predicting that the value in (1) is indeed the dimension of the tensor network variety in the case of general matrix product states of bond dimension two, with the only exceptions that we classify. Besides these exceptions, we actually expect the upper bound of Theorem 3.0.2 to give the exact value of the dimension in "most" cases, in a way similar to the Alexander-Hirschowitz Theorem for secant varieties of Veronese varieties [AH95].

Uniform matrix product states. An upper bound analogous to the one of our Theorem 3.0.2 is proposed for uniform matrix product states in [CMS19]. Uniform matrix product states are translation invariant matrix product states, meaning that the same tensor is associated to every vertex of the underlying graph. One can verify that the proposed value coincides with the dimension of the variety for a number of small parameter values; in particular, there are no known exceptions in the translation invariant setting, in contrast with the exceptions that we have found in our general case.

We denote uniform matrix product states by $\mathrm{uMPS}(m,n,d)$, where $m,n\in\mathbb{N}$ are local and bond dimensions, respectively and d is the number of vertices of the underlying cyclic graph. The tensor placed on all the vertices of the graph, denoted by $A\in\mathbb{C}^m\otimes\mathbb{C}^m\otimes\mathbb{C}^n$, is identified with the set of n matrices $(A_0,\dots,A_{n-1})\in(\mathbb{C}^{m\times m})^n$. Let $\{e_i\}_{i=1,\dots,n}$ be the canonical basis of $\mathbb{C}^n$. The variety of uniform matrix product states is defined as either the Zariski or Euclidean closure of the image of the following polynomial map:

$$\begin{aligned}\phi:(\mathbb{C}^{m\times m})^n&\to(\mathbb{C}^n)^{\otimes d}\\ (A_0,\dots,A_{n-1})&\mapsto\sum_{0\le i_1,\dots,i_d\le n-1}\mathrm{Tr}(A_{i_1}\cdots A_{i_d})\ e_{i_1}\otimes\cdots\otimes e_{i_d}.\end{aligned}$$

From a quantum mechanics perspective, uniform matrix product states model translation invariant physical systems of sites placed on a ring. Indeed, an alternative name for

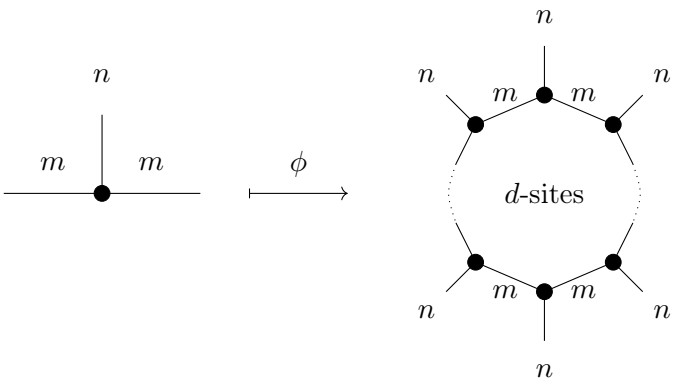


Figure 1: Graphical representation of the map defining uMPS(m, n, d). Each vertex of the cyclic graph on d vertices is associated with a tensor of order $m \times m \times n$. Tensors are contracted along the edges of the graph.

uMPS is *translation invariant matrix product states with periodic boundary conditions* [PGVWC07, CM14]. The name "uniform matrix product states" is sometimes reserved for the thermodynamic limit, where the number d of sites approaches infinity. Our terminology is consistent with [HMOV14, CMS19].

Due to the interest in the field of quantum physics, the geometry of uniform matrix product states has been extensively studied [PGVWC07, HMOV14, CM14, CMS19]. However, several fundamental mathematical problems remain open, even concerning their dimension. The ultimate goal would be to obtain a complete description of the variety, i.e. to find all its defining equations. Critch and Morton gave a complete description of the ideal of the varieties uMPS$(2, 2, 4)$ and uMPS$(2, 2, 5)$ and, in [CM14], several linear equations of uMPS$(2, 2, d)$ are given for d until 12. The generators of the ideal of uMPS$(2, 2, d)$ for $d = 4, 5, 6$ are given in [CMS19]. Nevertheless, phrased in generality, this question is likely to be intractable. Indeed, even just determining which *linear* equations, if any, vanish on the variety is poorly understood. This is precisely the goal of Chapter 4 of the thesis in which we study, via representation theory and linear algebra techniques, the linear span $\langle \text{uMPS}(m, n, d) \rangle$ of the variety of uniform matrix product states, i.e. the smallest linear space containing the variety. In particular, we are interested in determining its dimension. The variety uMPS(m, n, d) is a subset of the space of cyclically invariant tensors, denoted by $\text{Cyc}^d(\mathbb{C}^n) \subset (\mathbb{C}^n)^{\otimes d}$, because of the trace invariance under cyclic permutations of the matrices. Moreover, uMPS$(2, 2, d)$ is a subspace of the space of dihedrally symmetric tensors, denoted by $\text{Dih}^d(\mathbb{C}^2) \subset (\mathbb{C}^2)^{\otimes d}$, [Gre14]. We consider these spaces as natural ambient spaces of the proper variety. As noted in [CMS19], if we fix the local dimension n and the number of sites d and we let the bond dimension m grow, the space uMPS(m, n, d) will eventually fill its entire ambient space. Moreover, for $m = d$, it has been proved that equality holds [CMS19]. On the other hand, it follows from a dimension count [CMS19, NV18] that if $d \gg m$, the inclusion $\langle \text{uMPS}(m, n, d) \rangle \subset \text{Cyc}^d(\mathbb{C}^n)$ is strict.

Our main result of Chapter 4 is the following theorem, based on the Cayley-Hamilton theorem, which prescribes a method to find *nontrivial linear equations* that vanish on uMPS(m, n, d):

Theorem 4.3.6. *Let $A_0, \ldots, A_m, B$ be $m \times m$ matrices. Then for every $\ell \in \mathbb{N}$ it holds that*

$$\sum_{\sigma \in \mathfrak{S}_m, \tau \in C_{m+1}} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \operatorname{Tr}(A_{\tau(0)} B^{\sigma(0)} A_{\tau(1)} B^{\sigma(1)} \cdots A_{\tau(m-1)} B^{\sigma(m-1)} A_{\tau(m)} B^{\ell}) = 0.$$

Here $\mathfrak{S}_m$ is the symmetric group acting on $\{0, 1, \ldots, m-1\}$, and C_{m+1} is the cyclic group acting on $\{0, 1, \ldots, m\}$.

Theorem 4.3.6 actually gives nontrivial trace relations, i.e. linear relations that do not follow either from cyclic permutations or reflections of the factors. As a corollary, we can prove that the inclusion $\langle \text{uMPS}(m, n, d) \rangle \subset \text{Cyc}^d(\mathbb{C}^n)$ is strict already for $d = O(m^2)$, improving the state of the art. More precisely, the linear span of the space of uniform matrix product states is a proper subspace of the space of cyclically invariant tensors under the following conditions:

Corollary 4.3.7. *If $n \geq 3$ and $d \geq \frac{(m+1)(m+2)}{2}$, then* uMPS$(m, n, d)$ *is contained in a proper linear subspace of the space of cyclically invariant tensors.*

For what concerns the characterization of the linear span of the variety we focus in particular on uMPS$(2, 2, d)$, i.e. uniform matrix product states with the first nontrivial range of parameters. Notice that the general linear group GL_2 naturally acts on the space $(\mathbb{C}^2)^{\otimes d}$, leaving the space uMPS$(2, 2, d)$ invariant. Consequently, the space $\langle \text{uMPS}(2, 2, d) \rangle$ can be naturally seen as a representation of GL_2. We undertake a computational study of the linear span and we describe an algorithm that can compute this space, viewed as a GL_2-representation. We exploit further the representation theory perspective in order to speed up the computation of several equations of the variety, until degree 3. Based on the computations done in degree one, we obtain a conjectured formula for the character (and in particular: the dimension) of the linear span $\langle \text{uMPS}(2, 2, d) \rangle$ and we take some first steps towards proving our conjectured character formula, using our Cayley-Hamilton technique. Moreover, using a particular reparametrization of the variety, called the *trace parametrization*, from a simple count of parameters, we show an upper bound on the dimension of $\langle \text{uMPS}(2, 2, d) \rangle$ which asymptotically (for $d \to \infty$) agrees with our conjectured formula and that is close to optimal.

Tensor network variational ansatz. Matrix product states and uniform matrix product states are particularly relevant in quantum physics since they are used to describe quantum spin chains [AKLT88, FNW92, ÖR95]. The original insight is attributed to I. Affleck, T. Kennedy, E. H. Lieb and H. Tasaki [AKLT88], since they proved that ground states of the AKLT model admit an analytic solution coinciding with a matrix product state representation. While in general analytic solutions for ground states are too hard to be found, the AKLT model suggested the use of matrix product states as

variational classes of tensors for the *approximation* of ground states. Indeed, today it is known that the ground states of a local gapped Hamiltonian on 1-dimensional spin chains are well approximated by this class of tensor network varieties [PGVWC07, VMC08]. By the physical interpretation of the model, the edges of the underlying graph of the tensor network variety encode the structure of the entanglement in the quantum state, and the bond dimension associated to every edge is a quantitative measure of the amount of quantum correlation in the wave function. The entanglement theory behind the construction of the matrix product states ansatz has led to the introduction of variational algorithms applied to projected entangled pair states (tensor network variety associated to the two-dimensional lattice graph) [VWPGC06, VC04] and tensor network varieties associated to higher dimensional graphs.

Several algorithms have been designed in the recent years for matrix product states, such as the Density Matrix Renormalization Group (DMRG) [Whi93], the Time-Evolving Block Decimation (TEBD) [Vid04], the Time-Dependent Variational Principle (TDVP) [HCO$^+$11] and the Variational Uniform Matrix Product State (VUMPS) [ZSVF$^+$18]. These methods exploit the "network structure", namely they act on the local tensors that constitute the network. Some of the methods solve a specific problem on a small subset of tensors, leaving the others fixed and then they repeat the procedure on all possible subsets. For example, suppose to have a functional to be minimized on the matrix product states variety. The method starts by selecting a subset of tensors and minimizing on the parameters of these tensors, leaving the other parameters unchanged. Then, this "local" approach is sequentially applied to several other subsets of tensors, until convergence to a good approximation of the minimum is reached. Other methods try to split the problem into several "local" problems, each one involving a single tensor of the network. For example, in the simulation of time evolution, they transform the linear Schrödinger equation into a set of non-linear differential equations, each one involving a single tensor. Essentially, these methods sequentially apply a local technique. Conjugate gradient methods have been introduced in [PVV11, VHCV16, VHV19] in order to approximate ground states of translation invariant systems with periodic boundary conditions. Moreover, based on the TDVP, a variational nonlinear conjugate gradient method has been proposed in [MHO13] and it has been applied to critical quantum field theory. Very recently Riemannian gradient-based optimization has been proved to be a competitive method for optimizing tensor network ansatz [HVDH21].

Since the geometry and global methods are our main tools all along the entire thesis, we consider and decide to test the nonlinear conjugate gradient method (NLCG). We use the NLCG as a global method, applied to all the parameters simultaneously. The NLCG is an adaptation of the conjugate gradient method to nonlinear problems. It performs a conjugate gradient descent to the minimum of a nonlinear functional, invoking in every step a line search routine. The line search is usually approximate since, in general, an analytic solution is too expensive to be found. The global approach would seek to exploit the knowledge on the dimension of the variety, which is strictly smaller than the dimension of the parameter space, at least due to the tensor network gauge invariance.

We implement the variational NLCG on matrix product states and homogeneous matrix product states with open boundary conditions, c.f. Figure 2. The latter are matrix product states constructed via site-independent tensors and a boundary condition, c.f. [NV18], and therefore their dimension is independent on the number of sites. We test the algorithm on the well-known AKLT model, because ground states of the AKLT Hamiltonian admit such a tensor network representation in the finite chain configuration.

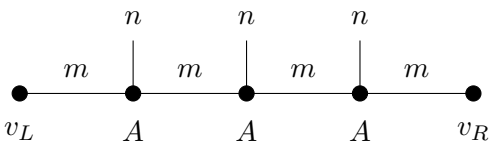


Figure 2: Homogeneous matrix product state with open boundary conditions with bond dimensions m and local dimensions n. The tensor $A \in (\mathbb{C}^{m\times m})^{\times n}$ is placed on each vertex. The boundaries v_L and v_R are vectors.

After verifying that the NLCG properly approximates the minimum of the expectation value functional and the corresponding ground state, we design a variation of the algorithm. The variation of the NLCG we propose modifies the *line search* method, which is the most expensive routine of the NLCG and it is based on a reparametrization of the gradient.

More precisely, we first reduce the number of coordinates of the gradient, computing a proper basis of the parameter space. Indeed, the gradient of the functional naturally belongs to the complementary space of the tangent space to the fiber of the parametrization. The gradient, therefore, admits a representative with a number of coordinates that coincides with the dimension of the variety. Moreover, in the case of matrix product states with open boundary conditions, the fiber is identified with the orbit of the gauge subgroup; therefore the tangent space to the fiber can be computed via the Lie algebra of the gauge subgroup. Fixing the gauge degrees of freedom in the fiber is a well known technique in physics and, in the context of matrix product states, it usually consists in putting the tensors of the network in the so called *canonical form*. We refer to [PGVWC07] for the definition and properties of this representation. After fixing a representative of the fiber of the map, gauge degrees of freedom are left in the representation of vectors of the tangent space to the variety [HMOV14], because, on the other side, the tangent directions to the gauge orbits are in the kernel of the differential of the map. The TDVP [HCO+11] and the VUMPS algorithm [ZSVF+18] use techniques to remove these degrees of freedom. Our approach is analogous.

Our final proposition consists in finding a basis that gives the desired representative of the gradient, for every starting point of the line search. Then, in fixing this basis (and the vector space that it generates) inside the line search routine. This implies that all the gradients that are computed inside the line search are *approximated*, being forced to

belong to a fixed vector space.

We move our preliminary steps towards the comparison with the standard NLCG. The study of this part of the thesis is preliminary. Further work must be done for characterizing completely the performances of our variation and for comparing the runtime of the NLCG (and our variation) with the runtime of the existing sequential algorithms. The preliminary results show that our variation of the line search could be of interest in the case of homogeneous matrix product states, whose dimension is independent of the number of sites. We can notice a gain in runtime to convergence, compared to the standard NLCG, even if not impressive. Moreover, in the homogeneous case, we clearly see that the *global method* preserves the *symmetries* of the tensor network, differently from the majority of sequential methods.

Structure of the thesis. The main contributions of this thesis are given in Chapters 3 and 4. Most of the contents of Chapters 1, 2, 3 and 4 of the thesis are based on two manuscripts. Manuscript [BDLG22], co-authored with Alessandra Bernardi and Fulvio Gesmundo, has been published in Communication in Contemporary Mathematics. Manuscript [DLMS22a], co-authored with Harshit J. Motwani and Tim Seynnaeve, has been submitted.

The thesis is structured as follows:

Chapter 1: Preliminaries. We introduce tensors and classical varieties of tensors. We give preliminary definitions and results on Lie groups and Lie algebra theory and we provide original results on the isotropy group of tensors, needed for the treatment of the following chapters.

Chapter 2: Tensor network varieties. We define tensor network varieties and we give the first basic properties. We introduce the gauge subgroup, which is a group acting on the parameter space of the tensor network variety. It will have a central role in the study of their dimension.

Chapter 3: Dimension. We investigate the dimension of tensor network varieties through algebraic geometric techniques. We provide a completely general upper bound on the dimension of any tensor network variety. A refined upper bound is given in cases relevant for applications such as varieties of matrix product states and projected entangled pairs states. Moreover, we provide a range of parameters in which the upper bound is sharp. On the contrary, we find and analyze small cases of matrix product states in which the bound is not sharp, contrary to what is generally expected. The code [BDLG21] associated to Chapter 3 is implemented in Macaulay2 [GS20] and it is available at `https://fulges.github.io/code/BDG-DimensionTNS.html`.

Chapter 4: Linear span of uniform matrix product states. We study the linear span of uniform matrix product states via representation theory methods and Cayley-Hamilton theorem. We show that the linear span of uniform matrix product states is a strict subspace of its ambient space, as long as the number of sites is at least quadratic

in the bond dimension, improving the state of the art. We give a conjecture on the dimension of the linear span of the variety in the first nontrivial range of parameters. The code [DLMS22b, DLMS22c] associated to Chapter 4 is implemented in Sage [The20] and Macaulay2 [GS20] and it is available at https://github.com/harshitmotwani2015/uMPS/and https://github.com/claudia-dela/uMPS_highest-weight-vectors/.

Chapter 5: **Nonlinear conjugate gradient method on MPS**. We lay the groundwork for the implementation of a matrix product state algorithm for the approximation of ground states. We review the nonlinear conjugate gradient method adapted to the variational approach on matrix product states. Then we investigate the geometry of the fiber of the parametrization in order to propose a variation of the algorithm.

Chapter 6: **Preliminary numerical calculations**. We move our first steps towards the implementation of the methods, the variational nonlinear conjugate gradient method and the variation we designed; and we analyze our preliminary numerical results. The code [DL22] associated to Chapter 6 is implemented in MATLAB [Mat20] and it is available at https://github.com/claudia-dela/NLCG_MPS_open-boundaries/.

Chapter 1

Preliminaries

This chapter is intended to introduce basic notions employed throughout this work. Sections 1.1 contains algebraic and geometrical definitions and results: we refer to [FH13, Hal15] for what concerns representation theory, [Lee13] for basic results on differential geometry and [Zak93] for introducing classical algebraic varieties of tensors. In Section 1.1, Subsection 1.1.2, we focus on the action of a product of linear groups on a tensor space and on the induced Lie algebra action and we prove preliminary results on isotropy Lie algebras of tensors that will be used in Chapter 3. In Section 1.2, we enunciate the postulates of quantum mechanics [CTDL$^+$77] with the twofold role of describing its mathematical formalism and facilitating the non expert readers in the field.

1.1 Tensors

Let $V_1, \dots, V_d$ be finite dimensional complex vector spaces. An element of the tensor product $V_1 \otimes \cdots \otimes V_d$ of the form $v_1 \otimes \cdots \otimes v_d$, with $v_j \in V_j$ for $j = 1, \dots, d$ is called either *decomposable* tensor or rank 1 tensor. Notice that not every element of $V_1 \otimes \cdots \otimes V_d$ can be written as a rank 1 tensor.
If $V_j = \mathbb{C}^{n_j}$ for $j = 1, \dots, d$, with basis $\mathcal{B}_j = \{v_1^j, \dots, v_{n_j}^j\}$, then a basis of the tensor product $V_1 \otimes \cdots \otimes V_d$ associated to the bases $\mathcal{B}_j$, $j = 1, \dots, d$, is given by the rank 1 tensor tensors $\{v_{\mu_1}^1 \otimes \cdots \otimes v_{\mu_d}^d\}$ for $1 \leq \mu_1 \leq n_1, \dots, 1 \leq \mu_d \leq n_d$. In particular every tensor in $V_1 \otimes \cdots \otimes V_d$ can be written as a linear combination of rank 1 tensor tensors.

Let V and W be finite dimensional vector spaces over $\mathbb{C}$. There is a canonical isomorphism $V^* \otimes W \simeq \mathrm{Hom}(V, W)$ given by the linear application $A : V^* \otimes W \to \mathrm{Hom}(V, W)$ acting on rank 1 tensors as follows

$$A(f \otimes v) = f(v)w.$$

The action extends on $V^* \otimes W$ by linearity.

Every tensor can be associated with a set of linear maps called *flattening maps*.

Definition 1.1.1. Given a tensor $T \in V_1 \otimes \cdots \otimes V_d$, for every subset $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, d\}$, T defines a linear map

$$T_I : \bigotimes_{i \in I} V_i^* \to \bigotimes_{i' \notin I} V_{i'}$$

called *flattening map* associated to I. The flattening map is defined on rank 1 tensors $T = v_1 \otimes \cdots \otimes v_d \in V_1 \otimes \cdots \otimes V_d$ as follows

$$\begin{aligned} T_I : V_{i_1}^* \otimes \cdots \otimes V_{i_k}^* &\to V_{i_{k+1}} \otimes \cdots \otimes V_{i_d} \\ (f^{i_1} \otimes \cdots \otimes f^{i_k}) &\mapsto (f^{i_1}(v_{i_1}) \cdots f^{i_k}(v_{i_k}))(v_{i_{k+1}} \otimes \cdots \otimes v_{i_d}), \end{aligned}$$

with $f^j \in V_j^*$, for $j = i_1, \ldots, i_k$. The map extends by linearity. We say that T is *concise* if all the flattening maps $T_i : V_i^* \to \bigotimes_{i' \neq i} V_{i'}$ are injective.

Definition 1.1.2. Given two tensors $T \in V_1 \otimes \cdots \otimes V_d$ and $S \in V_1' \otimes \cdots \otimes V_d'$, the *Kronecker product* of T and S, denoted by $T \boxtimes S$, is the element $T \otimes S$ regarded as a tensor on d factors

$$T \boxtimes S \in (V_1 \otimes V_1') \otimes \cdots \otimes (V_d \otimes V_d').$$

Example 1.1.3. Let $T \in \mathrm{Hom}(V, W) \simeq V^* \otimes W$ and $S \in \mathrm{Hom}(V', W') \simeq V'^* \otimes W'$

$$T : V \to W, \quad S : V' \to W'.$$

Fix the bases $\{v_1, \ldots, v_n\}$,$\{v_1', \ldots, v_m'\}$,$\{w_1, \ldots, w_l\}$ and $\{w_1', \ldots, w_p'\}$ of the vector spaces V, V', W, W' respectively. Denote by $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times l}$ the matrices that represent the linear maps T and S in the given bases, i.e. $A = (a_{ij})$, for $i = 1, \ldots, n$ and $j = 1, \ldots, m$; and $B = (b_{kh})$, for $k = 1, \ldots, l, h = 1, \ldots, p$. Then $A \boxtimes B$ is a matrix that represents the tensor

$$S \boxtimes T : V \otimes V' \to W \otimes W',$$

with respect to the bases $\{v_1 \otimes v_1', v_1 \otimes v_2', \ldots, v_n \otimes v_m'\}$ and $\{w_1 \otimes w_1', w_1 \otimes w_2', \ldots, w_l \otimes w_t'\}$ of $V \otimes V'$ and $W \otimes W'$ respectively, i.e. it is the matrix

$$A \boxtimes B = \begin{pmatrix} a_{11}B & \ldots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \ldots & a_{mn}B \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & \ldots & a_{11}b_{1l} & \ldots & a_{1n}b_{11} & \ldots & a_{1n}b_{1l} \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ a_{11}b_{p1} & \ldots & a_{11}b_{pl} & \ldots & a_{1n}b_{p1} & \ldots & a_{1n}b_{pl} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}b_{11} & \ldots & a_{n1}b_{1l} & \ldots & a_{nn}b_{11} & \ldots & a_{nn}b_{1l} \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ a_{n1}b_{p1} & \ldots & a_{n1}b_{pl} & \ldots & a_{nn}b_{p1} & \ldots & a_{nn}b_{pl} \end{pmatrix}.$$

1.1.1 Geometric perspective

Let V be a complex vector space of dimension $N+1$.

Definition 1.1.4. The projectivisation of V is the space of 1-dimensional subspaces of V. It is defined as

$$\mathbb{P}(V) \simeq \frac{V \setminus \{0\}}{\sim},$$

where, for every $v, w \in V$, $v \sim w$ if and only if $v = \lambda w$, for some $\lambda \in \mathbb{C}$. We denote the equivalence class of $v \in V$ by $[v] \in \mathbb{P}(V)$. The complex dimension of $\mathbb{P}(V)$ is $N = \dim(V) - 1$. Sometimes we denote $\mathbb{P}(V) = \mathbb{P}^N$. If $X \subseteq \mathbb{P}(V)$, the *affine cone* over X is $\widehat{X} := \{v \in V : [v] \in X\}$.

Definition 1.1.5 (Segre embedding)**.** Given $V_1, \ldots, V_d$ complex vector spaces of dimension $\dim(V_j) = n_j + 1$ for $j = 1, \ldots, d$, the Segre map is defined as

$$\begin{aligned} \nu : \mathbb{P}(V_1) \times \cdots \times \mathbb{P}(V_d) &\to \mathbb{P}(V_1 \otimes \cdots \otimes V_d) := \mathbb{P}^N \\ ([v_1], \ldots, [v_d]) &\mapsto [v_1 \otimes \cdots \otimes v_d], \end{aligned}$$

where $N = (n_1 + 1) \cdots (n_d + 1) - 1$. The map is an isomorphism of algebraic varieties between the product of projective spaces and its image, which is called Segre variety and which we denote by

$$\nu(\mathbb{P}(V_1) \times \cdots \times \mathbb{P}(V_d)) := \mathcal{S}_{n_1, \ldots, n_d}.$$

The Segre variety is therefore the variety of rank 1 tensors, up to scalar multiplication.

Definition 1.1.6. Let $X \subseteq \mathbb{P}^N$ be a nondegenerate irreducible projective variety of projective dimension n. The *k-secant variety* of X is the Zariski closure of points of $\mathbb{P}^N$ contained in the linear span of k point of X

$$\sigma_k(X) := \overline{\{p \in \mathbb{P}^N \ : \ p \in \langle x_1, \ldots, x_k \rangle, \ x_1, \ldots, x_k \in X\}}.$$

Definition 1.1.7. Let $X \subseteq \mathbb{P}^N$ be a nondegenerate irreducible projective variety of projective dimension n. Let $T \in X$, then the *X-rank* of T is defined as:

$$\mathrm{rk}_X(T) := \min\{r \in \mathbb{N} \ : \exists \ x_1, \ldots, x_r \in X \text{ s.t } T \in \langle x_1, \ldots, x_r \rangle\}.$$

The *X-border rank* of T is defined as:

$$\mathrm{brk}_X(T) := \min\{r \in \mathbb{N} \ : T \in \sigma_r(X)\}.$$

If $X = \mathcal{S}_{n_1, \ldots, n_d}$, then the k-secant variety of the Segre variety is

$$\sigma_k(X) = \overline{\{T \in \mathbb{P}^N \ : \ T \in \langle x_1, \ldots, x_k \rangle, \ x_1, \ldots, x_k \in X \text{ rank 1 tensors}\}}.$$

Remark 1.1.8. In Definition 1.1.6, the closure is taken in the Zariski topology. Since we are working over the complex field $\mathbb{C}$, the closure can equivalently be taken in the Euclidean topology. Therefore elements of $\sigma_k(X)$ are points of $p \in \mathbb{P}^N$ such that $p \in \langle x_1, \dots, x_k \rangle$, with $x_1, \dots, x_k \in X$ and *limits* of points of this type.

We can give the definitions of rank and border rank of a tensor:

Definition 1.1.9. Given $T \in V_1 \otimes \cdots \otimes V_d$, the rank of T is

$$\mathrm{rk}(T) := \min\{r \in \mathbb{N} \ : T = \sum_{i=1}^{r} v_i^1 \otimes \cdots \otimes v_i^d, \ v_i^j \in V^j\}.$$

Let $v_i^j(t)$ be a curve in V^j, for $j = 1, \dots, d$. Given $T \in V_1 \otimes \cdots \otimes V_d$, the border rank of T is

$$\begin{aligned} \mathrm{brk}(T) &:= \min\{r \in \mathbb{N} \ : T \in \sigma_r(X)\}. \\ &= \min\{r \in \mathbb{N} \ : T = \lim_{t \to 0} \Big(\sum_{i=1}^{r} v_i^1(t) \otimes \cdots \otimes v_i^d(t) \Big), \ v_i^j(t) \in V^j\}. \end{aligned}$$

1.1.2 Isotropy groups of tensors

We briefly recall some useful definitions concerning groups and group actions. Then we move to the study of the isotropy group (and the isotropy algebra) of a tensor. We consider a product of linear groups $G = GL(V_1) \times \cdots \times GL(V_d)$ acting on a tensor space $V_1 \otimes \cdots \otimes V_d$. The isotropy group of a given tensor $T \in V_1 \otimes \cdots \otimes V_d$ is a subgroup of G whose elements stabilize T. The isotropy Lie algebra of T is the subalgebra of $\mathfrak{g}$, the Lie algebra of G, whose elements annihilate the tensor. We prove preliminary results on isotropy Lie algebras of tensors that will be used in Chapter 3.

Definition 1.1.10. Given a group G and a set X, the *action* of G on X is a map

$$\begin{aligned} G \times X &\to X \\ (g, x) &\mapsto g \cdot x, \end{aligned}$$

such that

1. $e \cdot x = x$, for every $x \in X$, with $e \in G$ the identity of the group,
2. $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$ for every $g_1, g_2 \in G$ and $x \in V$.

Definition 1.1.11. The *orbit* of $x \in X$ under the action of G is

$$\mathcal{O}_x := \{g \cdot x : g \in G\} \subset X.$$

The *stabilizer* of an element $x \in X$ under the action of G is the subgroup

$$G_x := \{g \in G : g \cdot x = x\} \leq G.$$

Definition 1.1.12. A Lie group is a smooth manifold G that is also a group in the algebraic sense, with the property that the multiplication map and the inversion map

$$\begin{aligned} G \times G \to G \qquad & \qquad G \to G \\ (g,h) \mapsto gh \qquad & \qquad g \mapsto g^{-1}, \end{aligned}$$

are smooth.

Associated to any Lie group G is a Lie algebra.

Definition 1.1.13. A finite dimensional real or complex *Lie algebra* is a finite dimensional real or complex vector space $\mathfrak{g}$, together with a map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, with the following properties

1. $[\cdot,\cdot]$ is bilinear,
2. $[X,Y] = -[Y,X]$ for every $X, Y \in \mathfrak{g}$,
3. $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ for every $X, Y, Z \in \mathfrak{g}$

The Lie algebra associated to a Lie group G is the set of left-invariant vector fields on G that are determined by their vectors at a single point, e.g. at the identity. Therefore, the Lie algebra associated to a Lie group G can be identified with T_eG, the tangent space to G at the identity $e \in G$.

The Lie algebra of a product of Lie groups is the sum of their Lie algebras and the action of a Lie group naturally induces an action of its Lie algebra.

Example 1.1.14. Let V be a n-dimensional complex vector space, $V \simeq \mathbb{C}^n$. Denote by $GL(V)$ the group of automorphism of V. The linear group $GL(V) \simeq GL_n$ is a Lie group and its Lie algebra, denoted by $\mathfrak{gl}_n$, is the space $V \otimes V^* \simeq \mathrm{End}_n$ of $n \times n$ complex matrices with the commutator of matrices as Lie bracket: $[X,Y] = XY - YX$, for $X, Y \in GL_n$.

Example 1.1.15. Consider $G \subseteq GL(V)$ and its Lie algebra $\mathfrak{g}$. Assume that G acts on $V^{\otimes d}$ as follows

$$g \cdot (v_1 \otimes \cdots \otimes v_d) = (g \cdot v_1) \otimes \cdots \otimes (g \cdot v_d), \quad g \in G, v_1, \ldots, v_d \in V.$$

Then, there is an induced action of the Lie algebra $\mathfrak{g}$ given by

$$X \cdot (v_1 \otimes \cdots \otimes v_d) = (X \cdot v_1) \otimes v_2 \otimes \cdots \otimes v_d + \cdots + v_1 \otimes \cdots \otimes v_{d-1} \otimes (X \cdot v_d), \quad X \in \mathfrak{g},$$

If $g(t) \subset G$ is a curve in G such that $g(0) = \mathrm{Id}_n$, and $g'(0) = X$, the induced action can be computed as the infinitesimal action of the group at the identity

$$\begin{aligned} &\frac{d}{dt}_{|t=0} g(t) \cdot (v_1 \otimes \cdots \otimes v_d) = \frac{d}{dt}_{|t=0} (g(t) \cdot v_1) \otimes \cdots \otimes (g(t) \cdot v_d) \\ &= \left(\frac{d}{dt}_{|t=0} (g(t) \cdot v_1) \right) \otimes v_2 \otimes \cdots \otimes v_d + \cdots + v_1 \otimes \cdots \otimes v_{d-1} \otimes \left(\frac{d}{dt}_{|t=0} (g(t) \cdot v_1) \right) \\ &= (X \cdot v_1) \otimes v_2 \otimes \cdots \otimes v_d + \cdots + v_1 \otimes \cdots \otimes v_{d-1} \otimes (X \cdot v_d). \end{aligned}$$

We are interested exactly in the action of the product of linear groups (and of its Lie algebra) on a tensor space.

Let $V_1, \dots, V_d$ complex vector spaces and consider the natural action of the group $GL(V_1) \times \cdots \times GL(V_d)$ on the tensor product $V_1 \otimes \cdots \otimes V_d$, given by

$$\begin{aligned} GL(V_1) \times \cdots \times GL(V_d) \times V_1 \otimes \cdots \otimes V_d &\to V_1 \otimes \cdots \otimes V_d \\ (g_1, \dots, g_d, v_1 \otimes \cdots \otimes v_d) &\mapsto g_1(v_1) \otimes \cdots \otimes g_d(v_d). \end{aligned}$$

This defines a group homomorphism

$$\begin{aligned} GL(V_1) \times \cdots \times GL(V_d) &\to GL(V_1 \otimes \cdots \otimes V_d) \\ (g_1, \dots, g_d) &\mapsto g_1 \otimes \cdots \otimes g_d \end{aligned}$$

whose kernel is the central subgroup

$$Z_{V_1 \otimes \cdots \otimes V_d} = \{(\lambda_1 \mathrm{Id}_{V_1}, \dots, \lambda_d \mathrm{Id}_{V_d}) : \lambda_1 \cdots \lambda_d = 1\}.$$

Therefore, the group

$$G(V_1, \dots, V_d) := GL(V_1) \times \cdots \times GL(V_d)/Z_{V_1 \otimes \cdots \otimes V_d} \tag{1.1}$$

can be identified naturally with a subgroup of $GL(V_1 \otimes \cdots \otimes V_d)$ acting faithfully on $V_1 \otimes \cdots \otimes V_d$. The elements of $G(V_1, \dots, V_d)$ will be denoted as tensor products $g_1 \otimes \cdots \otimes g_d$ for $g_j \in GL(V_j)$.

The corresponding Lie algebra action defines a Lie algebra homomorphism

$$\begin{aligned} \mathfrak{gl}(V_1) \oplus \cdots \oplus \mathfrak{gl}(V_d) &\to \mathfrak{gl}(V_1 \otimes \cdots \otimes V_d) \\ (X_1, \dots, X_d) &\mapsto X_1 \otimes \mathrm{Id}_{V_2} \otimes \cdots \otimes \mathrm{Id}_{V_d} + \cdots + \mathrm{Id}_{V_1} \otimes \cdots \otimes \mathrm{Id}_{V_{d-1}} \otimes X_d, \end{aligned}$$

whose kernel is the central algebra

$$\mathfrak{z}_{V_1 \otimes \cdots \otimes V_d} = \{(x_1 \mathrm{Id}_{V_1}, \dots, x_d \mathrm{Id}_{V_d}) : x_1 + \cdots + x_d = 0\}.$$

Hence, the Lie algebra $\mathfrak{g}(V_1, \dots, V_d) := \mathfrak{gl}(V_1) \oplus \cdots \oplus \mathfrak{gl}(V_d)/\mathfrak{z}_{V_1 \otimes \cdots \otimes V_d}$ is a subalgebra of $\mathfrak{gl}(V_1 \otimes \cdots \otimes V_d)$ and coincides with the Lie algebra of $G(V_1, \dots, V_d)$. With abuse of notation, denote the elements of $\mathfrak{g}(V_1, \dots, V_d)$ as d-tuples $\mathbf{X} = (X_1, \dots, X_d)$ with $X_j \in \mathfrak{gl}(V_j)$ with the understanding that $\mathbf{X}$ is identified with its image in $\mathfrak{g}(V_1, \dots, V_d)$.

Definition 1.1.16. Let $T \in V_1 \otimes \cdots \otimes V_d$ be a tensor. The isotropy group of T, denoted G_T, is the stabilizer of T under the action of $G(V_1, \dots, V_d)$:

$$G_T = \{g_1 \otimes \cdots \otimes g_d \in G(V_1, \dots, V_d) : g_1 \otimes \cdots \otimes g_d(T) = T\}.$$

The group G_T is algebraic and in general it is union of finitely many connected (irreducible) components. Let G_T° denote the connected component containing the identity: G_T° is normal in G_T (see, e.g., [Ges16, Lemma 2.1]) and $\dim G_T = \dim G_T^\circ$.

The isotropy Lie algebra of T, denoted $\mathfrak{g}_T$, is the Lie algebra of the group G_T, or equivalently the one of G_T°; it is the subalgebra of $\mathfrak{g}(V_1,\dots,V_d)$ which annihilates T under the Lie algebra action induced by $\mathfrak{gl}(V_1)\oplus\cdots\oplus\mathfrak{gl}(V_d)$ [Pro07, Sec. 1.2]

$$\mathfrak{g}_T=\{\mathbf{X}=(X_1,\dots,X_d)\in\mathfrak{g}(V_1,\dots,V_d):\mathbf{X}.T=0\},$$

where $\mathbf{X}.T=\sum_1^d \mathrm{Id}_{V_1}\otimes\cdots\otimes X_k\otimes\cdots\otimes\mathrm{Id}_{V_d}(T)$ denotes the image via the Lie algebra action. We have $\dim\mathfrak{g}_T=\dim G_T^\circ=\dim G_T$.

Remark 1.1.17. The dimension of the orbit-closure of $T\in V_1\otimes\cdots\otimes V_d$ under the action of $G(V_1,\dots,V_d)$ is given by

$$\begin{aligned}\dim(G(V_1,\dots,V_d)\cdot T)&=\dim G(V_1,\dots,V_d)-\dim G_T\\&=\left[\textstyle\sum_{i=1}^d(\dim V_i)^2-d+1\right]-\dim\mathfrak{g}_T.\end{aligned}$$

We prove preliminary results on isotropy Lie algebras of tensors. Lemma 1.1.19 is classical and we recall it here for the reader's convenience. Lemma 1.1.20 concerns the intersection of $\mathfrak{g}_T$ with the subalgebra of $\mathfrak{g}(V_1,\dots,V_d)$ consisting of elements acting only on a subset of the tensor factors.

We first recall an immediate linear algebra fact.

Lemma 1.1.18. *Let V be a vector space and let $A,B_1,\dots,B_N$ be subspaces of V for which there exists a subspace B such that $A\cap B=\{0\}$ and $B_j\subseteq B$ for every $j=1,\dots,N$. Then $\bigcap_j(A\oplus B_j)=A\oplus\bigcap_j B_j$.*

Proof. Since $A\cap B=\{0\}$ and $B_j\subseteq B$ for $j=1,\dots,N$, then for every $j=1,\dots,N$ we have

$$B_j\cap A=(B_j\cap B)\cap A=B_j\cap(B\cap A)=B_j\cap\{0\}=\{0\}$$

and therefore $\bigcap_j B_j\cap A=\{0\}$.

Consider $v\in A\oplus\bigcap_j B_j$. Then $v=a+b$, with $a\in A$ and $b\in\bigcap_j B_j$. In particular $b\in B_j$ for every $j=1,\dots,N$ therefore $v=a+b\in A\oplus B_j$ for every $j=1,\dots,N$. Then $v\in\bigcap_j(A\oplus B_j)$.

Vice versa, if $v\in\bigcap_j(A\oplus B_j)$ then $v\in A\oplus B_j$ for every $j=1,\dots,N$ and we can write $v=a_j+b_j$ for $a_j\in A$ and $b_j\in B_j$ for every $j=1,\dots,N$. Consider $v=a_i+b_i=a_k+b_k$ for $i\neq k\in\{1,\dots,N\}$. If we denote $w=a_i-a_k=b_k-b_i$, then $w\in A\cap(B_i\cup B_k)$ but, since A is disjoint from all the B_j's, in particular $A\cap B_i=\{0\}=A\cap B_k$, we have $w=0$. Therefore $a_i=a_k=:a$, $b_j=b_k=:b$ and $v=a+b\in A\oplus\bigcap_{j=i,k}B_j$. This easily generalizes to $v\in A\oplus\bigcap_j B_j$. □

The following result is classical and follows for instance from [Bri05, Section 1.1].

Lemma 1.1.19. *Let $T \in V_1 \otimes \cdots \otimes V_d$ be a non-concise tensor. Let $V_i' \subseteq V_i$ be subspaces such that $T \in V_1' \otimes \cdots \otimes V_d'$ is concise. Write $\mathfrak{h}_T$ for the isotropy Lie algebra of T in $\mathfrak{g}(V_1', \ldots, V_d')$ (regarded as a subalgebra of $\mathfrak{g}(V_1, \ldots, V_d)$) and $\mathfrak{g}_T$ for the isotropy Lie algebra of T in $\mathfrak{g}(V_1, \ldots, V_d)$. Then*

$$\mathfrak{g}_T = \mathfrak{h}_T \oplus \mathfrak{p}$$

where $\mathfrak{p} \subseteq \mathfrak{g}(V_1, \ldots, V_d)$ is the Lie algebra which annihilates the subspace $V_1' \otimes \cdots \otimes V_d'$, that is the algebra generated by $\bigoplus_{i=1}^d (V_i'^{\perp} \otimes V_i) \subseteq \mathfrak{gl}(V_1) \oplus \cdots \oplus \mathfrak{gl}(V_d)$.

Lemma 1.1.20. *Let $T \in V_1 \otimes \cdots \otimes V_d$. For $I \subseteq \{1, \ldots, d\}$, let $F_T := T_{I^c} : \bigotimes_{j \in I^c} V_j^* \to \bigotimes_{i \in I} V_i$ be the flattening map of T corresponding to the subset I. Then*

$$\mathfrak{g}_T \cap \mathfrak{g}(V_i : i \in I) = \bigcap_{S \in \operatorname{Im} F_T} \mathfrak{g}_S. \tag{1.2}$$

In particular, if T is concise, $\mathfrak{g}_T \cap \mathfrak{gl}(V_j) = 0$ for every j.

Proof. Let $k = |I|$; up to reordering the factors, assume $I = \{1, \ldots, k\}$.

Given $\mathbf{X} \in \mathfrak{g}(V_1, \ldots, V_d)$, write $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ with $\mathbf{X}_1 = (X_1, \ldots, X_k)$ and $\mathbf{X}_2 = (X_{k+1}, \ldots, X_d)$. Let $\mathbf{X}.T$ be the image of T via the action of $\mathbf{X}$ and let $F_{\mathbf{X}.T}$ be the corresponding flattening map. By Leibniz's rule, given an element $S' \in V_{k+1}^* \otimes \cdots \otimes V_d^*$, $F_{\mathbf{X}.T}$ is characterized by the expression

$$F_{\mathbf{X}.T}(S') = F_T(\mathbf{X}_2.S') + \mathbf{X}_1.F_T(S'),$$

where $\mathbf{X}_2$ acts on $V_{k+1}^* \otimes \cdots \otimes V_d^*$, $\mathbf{X}_1$ acts on $V_1 \otimes \cdots \otimes V_k$.

Now, let $\mathbf{X} \in \mathfrak{g}_T \cap (\mathfrak{gl}(V_1) \oplus \cdots \oplus \mathfrak{gl}(V_k))$. Hence, $\mathbf{X} = (\mathbf{X}_1, \mathbf{0})$ and $\mathbf{X}.T = 0$. Therefore $0 = F_{\mathbf{X}.T}(S') = \mathbf{X}_1.F_T(S')$, showing $\mathbf{X}_1 \in \mathfrak{g}_S$ for every $S \in \operatorname{Im} F_T$.

Conversely, let $\mathbf{X}_1 \in \bigcap_{S \in \operatorname{Im} F_T} \mathfrak{g}_S$. Let $S_1, \ldots, S_N \in \operatorname{Im} F_T$ be a set of generators and write $T = \sum_{i=1}^N S_i \otimes P_i$ for some $P_i \in V_{k+1} \otimes \cdots \otimes V_d$. Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{0})$. Then

$$\mathbf{X}.T = \sum_{i=1}^N (\mathbf{X}_1.S_i) \otimes P_i + \sum_{i=1}^N S_i \otimes \mathbf{0}.P_i = \sum_{i=1}^N (\mathbf{X}_1.S_i) \otimes P_i = 0$$

showing $\mathbf{X} \in \mathfrak{g}_T$. This concludes the proof of Equation (1.2).

The last claim follows by taking $I = \{j\}$: if T is concise, then F_T is surjective and therefore $\bigcap_{S \in \operatorname{Im} F_T} \mathfrak{g}_S = \bigcap_{v \in V_j} \mathfrak{g}_v = 0$. □

By linearity the intersection in Lemma 1.1.20 can be restricted to a basis of the image of the flattening map $\operatorname{Im} F_T$, as it is clear from the proof.

Additional results on isotropy groups. In this paragraph, we prove a generalization of a result of [CGL+20]. In [CGL+20, Thm. 4.1(iii)], it is proved that the Kronecker product of tensors which have a 0-dimensional isotropy group, has a 0-dimensional isotropy group. We generalize the result proving that this holds even in the case that only one factor of the Kronecker product has a 0-dimensional isotropy group. The isotropy group is defined in Definition 1.1.16.

Given two spaces V, W, there is a natural embedding $GL(V) \to GL(V \otimes W)$ defined by $g \mapsto g \otimes \mathrm{Id}_W$; correspondingly the Lie algebra $\mathfrak{gl}(V)$ can be regarded as a subalgebra of $\mathfrak{gl}(V \otimes W)$. In particular, if $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is a subalgebra, then $\mathfrak{g}$ is naturally identified with a subalgebra of $\mathfrak{gl}(V \otimes W)$.

Proposition 1.1.21. *Let $T \in V_1 \otimes \cdots \otimes V_d$ and $S \in W_1 \otimes \cdots \otimes W_d$ be concise tensors. Assume $\mathfrak{g}_T = \{0\} \subseteq \mathfrak{g}(V_1, \ldots, V_d)$. Then*

$$\mathfrak{g}_{T \boxtimes S} = \mathfrak{g}_S$$

regarded as a subalgebra of $\mathfrak{g}(V_1 \otimes W_1, \ldots, V_d \otimes W_d)$.

Proof. The inclusion

$$\mathfrak{g}_S \subseteq \mathfrak{g}_{T \boxtimes S}$$

is immediate from the definition of Kronecker product 1.1.2.

Let $\mathbf{X} \in \mathfrak{g}_{T \boxtimes S}$. Write $\mathbf{X} = (X_1, \ldots, X_d)$ with $X_k \in \mathfrak{gl}(V_k \otimes W_k)$. Our goal is to show that $X_k = \mathrm{Id}_{V_k} \otimes Z_k$ for some $Z_k \in \mathfrak{gl}(W_k)$ with $\mathbf{Z} := (Z_1, \ldots, Z_d) \in \mathfrak{g}_S$.

For every $p = 1, \ldots, d$, fix bases $\{v_j^p : j = 1, \ldots, \dim V_p\}$ of V_p and similarly for W_p. Write

$$T = \sum T^{i_1, \ldots, i_d} v_{i_1}^1 \otimes \cdots \otimes v_{i_d}^d,$$
$$S = \sum S^{j_1, \ldots, j_d} w_{j_1}^1 \otimes \cdots \otimes w_{j_d}^d.$$

For $k = 1, \ldots, d$, write $(x_k)_{i'j'}^{ij}$ for the entries of X_k with respect to the basis $v_i^k \otimes w_j^k$. By Leibniz's rule, the condition $\mathbf{X}.(T \boxtimes S) = 0$ is equivalent to

$$\sum_{k=1}^{d} (x_k)_{i'_k j'_k}^{i_k j_k} T^{i_1, \ldots, i'_k, \ldots, i_d} S^{j_1, \ldots, j'_k, \ldots, j_d} = 0 \quad \text{for every } i_1, \ldots, i_d, j_1, \ldots, j_d, \tag{1.3}$$

where we use the summation convention that repeated upper and lower indices are to be summed over their range.

For every $j_1, \ldots, j_d$, and for every $k = 1, \ldots, d$, define $Y_k(j_1, \ldots, j_d) \in \mathfrak{gl}(V_k)$ by

$$(y_k(j_1, \ldots, j_d))_{i'_k}^{i_k} = (x_k)_{i'_k j'_k}^{i_k j_k} S^{j_1, \ldots, j'_k, \ldots, j_d}.$$

Regard $\mathbf{Y}(j_1, \ldots, j_d) = (Y_1(j_1, \ldots, j_d), \ldots, Y_d(j_1, \ldots, j_d))$ as an element of $\mathfrak{g}(V_1, \ldots, V_d)$. From (1.3), we deduce that $\mathbf{Y}(j_1, \ldots, j_d)$ satisfies $\mathbf{Y}.T = 0$ and therefore $\mathbf{Y} \in \mathfrak{g}_T$. From

the hypothesis $\mathfrak{g}_T = \{0\}$ and therefore, for every k, there exists $\lambda_k(j_1, \dots, j_d)$ such that $Y_k(j_1, \dots, j_d) = \lambda_k(j_1, \dots, j_d)\mathrm{Id}_{V_k}$ and $\sum_k \lambda_k(j_1, \dots, j_d) = 0$.

Since $Y_k(j_1, \dots, j_d)$ is a multiple of the identity, we have

$$0 = (y_k(j_1, \dots, j_d))^{i_k}_{i'_k} = (x_k)^{i_k j_k}_{i'_k j'_k} S^{j_1, \dots, j'_k, \dots, j_d} \quad \text{for } i_k \neq i'_k,$$
$$0 = (y_k(j_1, \dots, j_d))^{i_k}_{i_k} - (y_k(j_1, \dots, j_d))^1_1 = [(x_k)^{i_k j_k}_{i_k j'_k} - (x_k)^{1 j_k}_{1 j'_k}] S^{j_1, \dots, j'_k, \dots, j_d}.$$

In other words, if $i_k \neq i'_k$, setting $Z_k(i_k, i'_k) \in \mathfrak{gl}(W_k)$ to be defined by $(z_k(i_k, i'_k))^{j_k}_{j'_k} = (x_k)^{i_k j_k}_{i'_k j'_k}$, we have $Z_k(i_k, i'_k).S = 0$. This means that $Z_k(i_k, i'_k) \in \mathfrak{g}_S \cap \mathfrak{gl}(W_k)$: since S is concise, Lemma 1.1.20 implies $Z_k(i_k, i'_k) = 0$. This shows that $(x_k)^{i_k j_k}_{i'_k j'_k} = 0$ whenever $i_k \neq i'_k$. Similarly, if $i_k \geq 2$, setting $(z_k(i_k))^{j_k}_{j'_k} = (x_k)^{i_k j_k}_{i_k j'_k} - (x_k)^{1 j_k}_{1 j'_k}$, we have $Z_k(i_k).S = 0$, hence $Z_k(i_k) = 0$ and therefore $(x_k)^{i_k j_k}_{i_k j'_k} = (x_k)^{1 j_k}_{1 j'_k}$ for every i_k.

We deduce that $X_k = \mathrm{Id}_{V_k} \otimes Z_k$ for some $Z_k \in \mathfrak{gl}(W_k)$. Now, let $\mathbf{Z} = (Z_1, \dots, Z_k)$. We conclude

$$0 = \mathbf{X}.(T \boxtimes S) = \mathbf{Z}.(T \boxtimes S) = T \boxtimes \mathbf{Z}.S$$

and therefore $\mathbf{Z} \in \mathfrak{g}_S$. This concludes the proof. □

1.2 Quantum Physics

We present the six postulates of Quantum Mechanics, mainly following [CTDL+77], with the aim to introduce notations and definitions useful to the reader. The postulates give the foundation of the physical theory and describe its mathematical formulation. In Subsection 1.2.2 we briefly define spin systems, highlighting the connection to the representation theory of the Lie algebra of SL_2 given in Chapter, 1, Section 1.1.2.

1.2.1 Postulates of Quantum Mechanics

In Quantum Mechanics, a physical system is described by three entities: states, observables and the dynamics or law of time evolution.

States. Each isolated physical system is associated to a complex Hilbert space $\mathcal{H}$, that in this thesis will always be a finite dimensional complex vector space, $\mathcal{H} \simeq \mathbb{C}^n$, endowed with a positive definite Hermitian inner product, i.e.

$$h : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$
$$(x, y) \mapsto h(x, y),$$

such that

1. $h(x, x) > 0$ for every $x \neq 0$.

2. $h(\alpha x + \beta y, z) = \alpha h(x, z) + \beta h(y, z)$, for every $\alpha, \beta \in \mathbb{C}$ and $x, y, z \in \mathcal{H}$,

3. $h(x, \alpha y + \beta z) = \overline{\alpha} h(x, y) + \overline{\beta} h(y, z)$, for every $\alpha, \beta \in \mathbb{C}$ and $x, y, z \in \mathcal{H}$.

and that satisfies $h(x, y) = \overline{h(y, x)}$, for every $x, y \in \mathcal{H}$.

The standard Hermitian inner product on $\mathbb{C}^n$ is defined as $h(x, y) := y^\dagger x$, where $\dagger$ denotes the conjugate transpose. The inner product allows the identification between $\mathcal{H}$ and its dual space $\mathcal{H}^*$ via the map

$$\begin{aligned} R : \mathcal{H} &\to \mathcal{H}^* \\ y &\mapsto \left(y^\dagger : x \mapsto y^\dagger x \right). \end{aligned}$$

The norm of $x \in \mathcal{H}$ is defined as $|x| := \sqrt{x^\dagger x}$.

Given a linear operator $A \in \mathrm{End}(\mathcal{H})$, the *adjoint* operator of A is $A^\dagger \in \mathrm{End}(\mathcal{H})$; it satisfies $h(A^\dagger x, y) = h(x, Ay)$. An operator A is called Hermitian if $A = A^\dagger$.

Postulate (1-th)**.** The *state* of a quantum physical system is represented, at a fixed time t_0, by a projective point $[v(t_0)] \in \mathbb{P}(\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space, called the *state space.*

Two vectors $v, w \in \mathcal{H}$ represent the same state if and only if their projective classes are equal. Therefore a quantum state $v \in \mathcal{H}$ can be identified with its equivalence class $[v] \in \mathbb{P}(\mathcal{H})$. The linear combination of states is called *superposition.*

Remark 1.2.1. In physics, vectors are denoted by the Dirac notation *ket* $|v\rangle \in \mathcal{H}$ and linear funcionals by the *bra* $\langle f| \in \mathcal{H}^*$.

Composite systems. The state space of a system that includes several quantum subsystems is a Hilbert space and it is the tensor product of the state spaces associated to the subsystems.

Given $d \geq 1$ Hilbert spaces $(\mathcal{H}_i, h_i)$, $i = 1, \ldots, d$, the Hermitian scalar product on $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d$ is given by the unique map $h : (\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d) \times (\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d) \to \mathbb{C}$ which verifies

$$h(v_1 \otimes \cdots \otimes v_d, u_1 \otimes \cdots \otimes u_d) := h_1(v_1, u_1) \cdots h_d(v_d, u_d),$$

for all $v_i, u_i \in \mathcal{H}_i$, $i = 1, \ldots, d$.

The tensor product of operators acts on the tensor product of vector spaces as follows:

Proposition 1.2.2. *Let $\mathcal{H}_1, \ldots, \mathcal{H}_d$ be finite dimensional complex vector spaces. Let $A_i \in \mathrm{End}(\mathcal{H}_i)$, for $i = 1, \ldots, d$ be linear operators. Then there exists a unique linear operator in $\mathrm{End}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d)$, denoted by $A_1 \otimes \cdots \otimes A_d$, such that*

$$A_1 \otimes \cdots \otimes A_d(v_1 \otimes \cdots \otimes v_d) = (A_1 v_1) \otimes \cdots \otimes (A_d v_d),$$

for all $v_i \in \mathcal{H}_i$, $i = 1, \dots, d$.

If $A_i, B_i \in \mathrm{End}(\mathcal{H}_i)$, *for* $i = 1, \dots, d$, *then*

$$(A_1 \otimes \cdots \otimes A_d)(B_1 \otimes \cdots \otimes B_d) = (A_1 B_1) \otimes \cdots \otimes (A_d B_d).$$

In physics, operators acting on single factors or some of the factors of the tensor product are called *local operators*. They naturally extend to operators acting on the tensor product of spaces:

Definition 1.2.3 (Extension of local operators)**.** Let $\mathcal{H}_i$, $i = 1, \dots, d$, be Hilbert spaces. Consider $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d$ and let $A_i \in \mathrm{End}(\mathcal{H}_i)$, for $i = 1, \dots, d$. For every $i = 1, \dots, d$, associate to A_i a linear operator $\widetilde{A}_i \in \mathrm{End}(\mathcal{H})$

$$\widetilde{A}_i := \mathrm{Id}_{\mathcal{H}_1} \otimes \cdots \otimes \mathrm{Id}_{\mathcal{H}_{i-1}} \otimes A_i \otimes \mathrm{Id}_{\mathcal{H}_{i+1}} \otimes \cdots \otimes \mathrm{Id}_{\mathcal{H}_d},$$

where $\mathrm{Id}_{\mathcal{H}_i}$ is the identity on $\mathcal{H}_i$, defined by

$$\widetilde{A}_i(v_1 \otimes \cdots \otimes v_d) := v_1 \otimes \cdots \otimes A_i(v_i) \otimes \cdots \otimes v_d.$$

The linear operator $\widetilde{A}_i$, defined on the whole Hilbert space $\mathcal{H}$, is called the *extension* of A_i on $\mathcal{H}$, c.f. [CTDL+77].

Fix $1 \leq i < j \leq d$. Let $A \in \mathrm{End}(\mathcal{H}_i)$ and $B \in \mathrm{End}(\mathcal{H}_j)$. We write

$$A_i B_j := \mathrm{Id}_{\mathcal{H}_1} \otimes \cdots \otimes \mathrm{Id}_{\mathcal{H}_{i-1}} \otimes A \otimes \mathrm{Id}_{\mathcal{H}_{i+1}} \otimes \cdots \otimes \mathrm{Id}_{\mathcal{H}_{j-1}} \otimes B \otimes \mathrm{Id}_{\mathcal{H}_{j+1}} \otimes \cdots \otimes \mathrm{Id}_{\mathcal{H}_d}.$$

Observables. Physical observables are physical quantities that can be measured. They are represented by *Hermitian* operators on $\mathcal{H}$, i.e. $H \in \mathrm{End}(\mathcal{H})$ such that $H = H^\dagger := \overline{H}^t$. Since H is Hermitian, its eigenvalues are real.

Postulate (2-nd)**.** Every measurable physical quantity $\mathcal{A}$ is described by a Hermitian operator A acting on the state space $\mathcal{H}$. This operator is an observable.

In the finite dimensional case the *spectrum* of A, denoted by $\sigma(A)$, is discrete. Since A is Hermitian, its eigenvectors form a basis of $\mathcal{H}$. Given $\lambda \in \sigma(A)$, the number of independent eigenvectors associated to the eigenvalue λ is usually known as *geometrical multiplicity*; in physics, it is called *degeneration* and it is denoted by $d(\lambda)$. An eigenvalue $\lambda \in \sigma(A)$ with $d(\lambda) = 1$ is said to be non-degenerate.

Postulate (3-rd)**.** The only possible result of the measurement of a physical quantity $\mathcal{A}$ is one of the eigenvalues of the corresponding observable A of $\mathcal{A}$.

Consider a system whose state is characterized, at a given time, by $v \in \mathcal{H}$.

Definition 1.2.4. Let $A \in \mathrm{End}(\mathcal{H})$ be an observable, i.e. a hermitian operator associated to a physical quantity $\mathcal{A}$. The *expectation value* of A in the state $v \in \mathcal{H}$ is defined as

$$\langle A \rangle_v := \frac{v^\dagger A v}{v^\dagger v}.$$

Remark 1.2.5. Notice that the expectation value of any observable does not depend on the choice of the representative $v \in \mathcal{H}$, since $\langle A \rangle_v = \langle A \rangle_w$, for every $w \in [v]$.

Let $\mathcal{H} = \mathbb{C}^n$ and let $A \in \mathrm{End}(\mathcal{H})$ be an observable. Suppose $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ with all eigenvalues λ_i, $i = 1, \dots, n$ distinct and non-degenerate, i.e. $d(\lambda_i) = 1$ for every $i = 1, \dots, n$. Then there is a unique (modulo scalar multiplication) eigenvector associated to each eigenvalue: $Av_i = \lambda_i v_i$, for every $i = 1, \dots, n$, with all v_i, $i = 1, \dots, n$ distinct and linearly independent. We can assume $v_i^\dagger v_i = 1$ for every $i = 1, \dots, n$. Let $v \in \mathcal{H}$ be a normalized vector, $v^\dagger v = 1$. It can be written in the eigenvectors basis as $v = \sum_{i=1}^n c_i v_i$, with $c_i \in \mathbb{C}$, $i = 1, \dots, n$. Then

$$\langle A \rangle_v = v^\dagger A v = v^\dagger \left(\sum_i \lambda_i v_i v_i^\dagger \right) v = \sum_i \lambda_i (v^\dagger v_i)(v_i^\dagger v) = \sum_i \lambda_i |v^\dagger v_i|^2 = \sum_{i=1}^n \lambda_i |c_i|^2.$$

Postulate (4-th, discrete non-degenerate spectrum)**.** When a physical quantity $\mathcal{A}$, associated to the observable A, is measured on a system in a *normalized* state $v \in \mathcal{H}$, the probability $\mathcal{P}(\lambda_i)$ of obtaining the *non-degenerate* eigenvalue λ_i of the corresponding observable A is

$$\mathcal{P}(\lambda_i) = |v^\dagger v_i|^2,$$

where v_i is a normalized eigenvector associated to λ_i.

Assume λ_k is degenerate, of degeneration $d(\lambda_k) = g_k$. The eigenspace associated to λ_k is denoted by E_{λ_k}, with dimension g_k. Then there exist $\{v_k^j\}_{j=1,\dots,g_k}$ independent vectors associated to λ_k: $Av_k^j = \lambda_i v_k^j$, and they give a basis of E_{λ_k}.

Postulate (4-th, discrete spectrum)**.** When the physical quantity $\mathcal{A}$ is measured on a system in a *normalized* state $v \in \mathcal{H}$, the probability $\mathcal{P}(\lambda_i)$ of obtaining the eigenvalue λ_i of the corresponding observable A is

$$\mathcal{P}(\lambda_i) = \sum_{j=1}^{g_i} |v^\dagger v_i^j|^2,$$

where g_i is the degeneration of λ_i and $\{v_i^j\}_{j=1,\dots,g_i}$ is an orthonormal set of vectors which form a basis of the eigenspace E_{λ_i} associated to λ_i of A.

Given $w \in \mathcal{H}$ such that $w^\dagger w = 1$, define $P = ww^\dagger \in End(\mathcal{H})$ the one dimensional projector on the subspace generated by $w \in \mathcal{H}$. It is an observable, its expectation value on $v \in \mathcal{H}$ is

$$v^\dagger (ww^\dagger) v = (v^\dagger w)(w^\dagger v) = |v^\dagger w|^2,$$

and it is called *transition probability* from v to w.

Denote the projector on E_{λ_i} by P_i, given by $P_i = \sum_{j=1}^{g_i} v_i^{j\dagger} v_i^j$.

Postulate (5-th)**.** If the measurement of a physical quantity $\mathcal{A}$ on the system in a state $v \in \mathcal{H}$ gives the result λ_i, the state of the system immediately after the measurement is the normalized projection

$$\frac{P_i(v)}{\sqrt{v^\dagger P_i v}},$$

of v onto the eigenspace associated to λ_i.

Recall that, for any $A \in \mathrm{End}(\mathcal{H})$ hermitian, the projector operators have the following properties: $P_\alpha^2 = P_\alpha$, for every $\alpha \in \sigma(A)$, $P_\alpha P_\beta = 0$ when $\alpha \neq \beta \in \sigma(A)$ and $\sum_{\alpha\in\sigma(A)} P_\alpha = \mathrm{Id}_{\mathcal{H}}$.

Postulate (6-th)**.** The time evolution of a state vector $v(t) \in \mathcal{H}$ is governed by the Schrödinger equation

$$i\hbar \frac{d}{dt} v(t) = H(t)v(t), \tag{1.4}$$

where $\hbar$ is the reduced Plank constant and where $H(t)$ is the observable associated to the total energy of the system, called Hamiltonian.

Definition 1.2.6 (Ground state)**.** Denote λ_0 the lowest eigenvalue of the Hamiltonian H, i.e. $\lambda_0 = \min\{\lambda \in \mathbb{R} \ : \ \lambda \in \sigma(H)\}$. If unique (modulo scalar multiplication), the eigenvector associated to λ_0 is called the *ground state* of H. If λ_0 has degeneration $d(\lambda_0)$ strictly greater than 1, then the associated eigenspace E_{λ_0} is $d(\lambda_0)$-dimensional and vectors in E_{λ_0} are called *degenerate* ground states.

In the thesis, we restrict our attention mainly to ground states of gapped local Hamiltonian.

Definition 1.2.7. Let $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d$, with $d \in \mathbb{N}$, and let $H : \mathcal{H} \to \mathcal{H}$ be the Hamiltonian of a physical quantum system. The gap of the Hamiltonian is defined as $\Delta := \lambda_1 - \lambda_0$, where λ_0 and λ_1 are the two smallest distinct eigenvalues of H. We say that H is gapless [Mov17] if for any constant $\epsilon > 0$ there exists a $d > 0$ such that $\Delta \leq \epsilon$. Otherwise, we say that H is gapped.

Definition 1.2.8. Let $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d$, with $d \in \mathbb{N}$, and let $H : \mathcal{H} \to \mathcal{H}$ be the Hamiltonian of a physical quantum system. The Hamiltonian is said to be *local* if

$$H = \sum_{i,j} h_{i,j} = \sum_{i,j} \mathrm{Id}^{\otimes(d-2)} \otimes h_{i,j},$$

with $h : \mathcal{H}_i \otimes \mathcal{H}_j \to \mathcal{H}_i \otimes \mathcal{H}_j$ local operator acting on sites $0 < i < j < d$. We are interested in nearest-neighbor Hamiltonians $H : (\mathbb{C}^n)^{\otimes d} \to (\mathbb{C}^n)^{\otimes d}$ of the form

$$H = \sum_{j=1}^{d-1} h_j = \sum_j \mathrm{Id}^{j-1} \otimes h \otimes \mathrm{Id}^{d-j-1},$$

with $h : \mathbb{C}^n \otimes \mathbb{C}^n \to \mathbb{C}^n \otimes \mathbb{C}^n$ local operator acting on two nearby sites. In words, the physical system associated with a site interacts only with its neighbors.

1.2.2 Spin representation

Definition 1.2.9. A *representation* of a group G on a vector space V is a *linear action*, i.e. a group homomorphism

$$\begin{aligned} \rho : G &\to GL(V) \\ g &\mapsto \rho(g) : v \mapsto \rho(g)(v). \end{aligned}$$

We denote $\rho(g)(v)$ with either $g \cdot v$ or $g(v)$. We will call representation either the vector space V or the group homomorphism ρ. A representation is called *irreducible* if there is no non-trivial subspace $W \subsetneqq V$, such that $g \cdot w \subset W$ for every $g \in G$ and $w \in W$.

If $\mathfrak{g}$ is a Lie algebra, then a complex representation of $\mathfrak{g}$ is a Lie algebra homomorphism $\pi : \mathfrak{g} \to \mathfrak{gl}_n$. The notion of irreducibility is defined analogously for representations of Lie algebras. Moreover, every definition holds for the real case: if $V = \mathbb{R}^n$ we have the notions of real (matrix) Lie group and real (matrix) Lie algebra.

Remark 1.2.10. Every representation of a matrix Lie group gives rise to a representation of the associated Lie algebra. The vice versa holds in the case of simply-connected Lie groups, therefore there is a one-to-one correspondence between the representations of a simply-connected matrix Lie group and its Lie algebra.

Irreducible representation of $\mathfrak{sl}_2$. The special unitary group, denoted by $SU_n \subset SL_n$, is the group of unitary matrices with determinant one, i.e. $U^\dagger U = UU^\dagger = \mathrm{Id}_n$ and $\det U = 1$ for every $U \in SU_n$. The group SU_2 is a relevant group in physics since its representations describe *spins*. The matrix Lie group SU_2 is simply-connected and its representations are found studying the representations of it Lie algebra $\mathfrak{su}_2$, which is a *real* Lie algebra. Moreover, the irreducible complex representations of $\mathfrak{su}_2$, are in one-to-one correspondence with the irreducible representations of its complexification $\mathfrak{su}_n + i\mathfrak{su}_n \simeq \mathfrak{sl}_n$, which is the Lie algebra of SL_2 [Hal15]. We introduce the representation theory of $\mathfrak{sl}_2$ and we give the definition of *spin* in terms of its irreducible representations, c.f. Remark 1.2.16.

Denote by SL_2 the complex special linear group, the group of complex 2×2 invertible matrices. Its Lie algebra, denoted by $\mathfrak{sl}_2$, is the simple Lie algebra of 2×2 trace zero complex matrices. It is $\mathbb{C}$-span by the three operators

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

which satisfy the following commutation relations

$$[H, X] = 2X, \quad [H, Y] = -2Y \quad [X, Y] = H.$$

We recall here classical results about the complete classification of the irreducible representation of $\mathfrak{sl}_2$, we refer to [FH13, Hal15].

Proposition 1.2.11. *For each integer $m \geq 0$, there is an irreducible representation $V^{(m)}$ of $\mathfrak{sl}_2$ with dimension $m+1$.*

Removing the hypothesis of irreducibility we have the following result:

Theorem 1.2.12. *Suppose π is any finite-dimensional, complex-linear representation of $\mathfrak{sl}_2$ acting on a space V. Then, we have the following results:*

1. *Every eigenvalue of $\pi(H)$ is an integer.*
2. *If v is a nonzero element of V such that $\pi(X)v = 0$ and $\pi(H)v = \lambda v$, then there is a non-negative integer m such that $\lambda = m$. Furthermore, the vectors $v, \pi(Y)v, \ldots, \pi(Y)^m v$ are linearly independent and their span is an irreducible invariant subspace of dimension $m+1$.*

Proposition 1.2.13. *If π is an irreducible representation of $\mathfrak{sl}_2$ with dimension $m+1$, then π is equivalent to the m-th symmetric power of the standard representation $\mathbb{C}^2$.*

Proof from [FH13]. The trivial one-dimensional representation $\mathbb{C}$ of $\mathfrak{sl}_2$ is $V^{(0)}$. Consider the standard representation of $\mathfrak{sl}_2$ on $\mathbb{C}^2$. If $\{x, y\}$ is the standard basis of $\mathbb{C}^2$, then $H(x) = x$ and $H(y) = -y$. Therefore $V = \mathbb{C} \cdot x \oplus \mathbb{C} \cdot y = V_{-1} \oplus V_1 = V^{(1)}$.

A basis of the 2-nd symmetric power $\mathrm{Sym}^2(\mathbb{C}^2)$ is given by $\{x^2, xy, y^2\}$ and the action is

$$\begin{aligned} H(x \cdot x) &= xH(x) + H(x)x = 2x \cdot x, \\ H(x \cdot y) &= xH(y) + H(y)x = 0, \\ H(y \cdot y) &= -2y \cdot y, \end{aligned}$$

so that $\mathrm{Sym}^2(\mathbb{C}^2) = V_{-2} \oplus V_0 \oplus V_2 = V^{(2)}$.

In general, a basis of the m-th symmetric power of $\mathbb{C}^2$, $\mathrm{Sym}^m(\mathbb{C}^2)$, is given by $\{x^m, x^{m-1}y, \ldots, y^m\}$ and the action is

$$\begin{aligned} H(x^{m-k}y^k) &= (m-k)H(x) \cdot x^{m-k-1}y^k + kH(y) \cdot x^{m-k}y^{k-1} \\ &= (m-2k) \cdot x^{m-k}y^k. \end{aligned}$$

The eigenvalues of H on $\mathrm{Sym}^m(\mathbb{C}^2)$ are $-m, -m+2, \ldots, m-2, m$, each one of multiplicity 1; therefore $\mathrm{Sym}^m(\mathbb{C}^2)$ is irreducible and isomorphic to $V^{(m)}$. □

Definition 1.2.14. If $\mathfrak{g}$ is a Lie algebra and π_1 and π_2 are representations of $\mathfrak{g}$ acting on spaces V_1 and V_2, then the tensor product of π_1 and π_2 is a representation of $\mathfrak{g}$ acting on $V_1 \otimes V_2$ defined by

$$\pi_1 \otimes \pi_2(X) = \pi_1(X) \otimes \mathrm{Id} + \mathrm{Id} \otimes \pi_2(X),$$

for every $X \in \mathfrak{g}$.

Suppose π_1 and π_2 are irreducible representations of $\mathfrak{sl}_2$ acting on $V^{(m)}$ and $V^{(n)}$. In general $V^{(m)} \otimes V^{(n)}$ will not be an irreducible representation but it can be decomposed into a direct sum of irreducible invariant subspaces.

Theorem 1.2.15. *For any non-negative integer k, let $V^{(k)}$ denote the irreducible representation of $\mathfrak{sl}_2$ of dimension $k+1$. For two non-negative integers m and n, consider $V^{(m)} \otimes V^{(n)}$ as a representation of $\mathfrak{sl}_2$. Assume $m \geq n$. Then*

$$V^{(m)} \otimes V^{(n)} \simeq V^{(m+n)} \oplus V^{(m+n-2)} \oplus \cdots \oplus V^{(m-n+2)} \oplus V^{(m-n)},$$

where $\simeq$ denotes an equivalence of $\mathfrak{sl}_2$ representations.

Remark 1.2.16. In the physics literature, this decomposition is referred to as either the *Clebsch-Gordan* theory or as the *addition of angular momentum*. The irreducible representations $V^{(m)}$ have been labeled here with the integer m; physicists use to label them by the *spin* s:

$$s = \frac{m}{2},$$

that can have therefore rational values : $0, \frac{1}{2}, 1, \frac{3}{2}, ...$

In particular with *spin* s we refer to the $\mathfrak{sl}_2$ representation

$$V^{(m)} \simeq \mathbb{C}^{m+1} = \mathbb{C}^{2s+1}.$$

Example 1.2.17. Consider the first case $W = V^{(1)} \otimes V^{(1)}$, with $V^{(1)} = \mathbb{C}^2$ the standard representation, with basis $\{e_1, e_2\}$. The basis of W is given by $\{e_{ij} := e_i \otimes e_j \ : \ i,j = 1,2\}$. Since e_1 and e_2 are eigenvectors of H with eigenvalue $+1$ and -1 respectively, then $\{e_{ij}\}_{i,j=1,2}$ are eigenvectors of eigenvalues $\{2,0,0,-2\}$. Since e_{11} is associated to the largest eigenvalue 2, then $X(e_{11}) = X(e_1) \otimes e_1 + e_1 \otimes X(e_1) = 0$; on the other hand $Y(e_{11}) = e_{12} + e_{21}, Y^2(e_{11}) = 2e_{22}$ and $\langle e_{11}, Y(e_{11}) = e_{12} + e_{21}, Y^2(e_{11}) = 2e_{22} \rangle = V^{(2)}$ is the 3-dimensional irreducible representation. The space $\langle e_{12} - e_{21} \rangle = V^{(0)}$ is also invariant under the action of $\mathfrak{sl}_2$ and it is the orthogonal complement of $V^{(2)}$ inside $\mathbb{C}^2 \otimes \mathbb{C}^2$. We have the following decomposition as $\mathfrak{sl}_2$ representations

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \simeq V^{(0)} \oplus V^{(2)}.$$

Spins: $\mathfrak{sl}_2$ representations in physics. In physics, the representation $V^{(m)}$ is called the spin s representation, for $s = \frac{m}{2}$, c.f. Remark 1.2.16.

If $\mathcal{H} = \mathbb{C}^2$ then any Hermitian matrix in End_2 is an observable. The space of such matrices is spanned by the Pauli matrices

$$\sigma^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma^2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Therefore the algebra of observables is isomorphic to $\mathfrak{sl}_2$ and $\mathcal{H}$ is the 2-dimensional $\mathfrak{sl}_2$-representation, that is $\mathcal{H} = V^{(1)}$.

If $\mathcal{H} = \mathbb{C}^3$ then it is the 3-dimensional $\mathfrak{sl}_2$-representation, isomorphic to $V^{(2)}$ and it is spanned by the so called spin 1 operators

$$S^1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, S^2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, S^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{1.5}$$

Let $\epsilon^{\alpha\beta\gamma}$ be the Levi-Civita antisymmetric tensor, i.e.

$$\epsilon^{\alpha\beta\gamma} := \begin{cases} 1 & \text{if } \operatorname{sgn}(\alpha,\beta,\gamma) \equiv 0 \mod 2 \\ -1 & \text{if } \operatorname{sgn}(\alpha,\beta,\gamma) \equiv 1 \mod 2 \\ 0 & \text{if } \alpha = \beta \text{ or } \alpha = \gamma \text{ or } \beta = \gamma. \end{cases}$$

Both bases satisfy the commutation relations of $\mathfrak{sl}_2$

$$[\sigma^\alpha, \sigma^\beta] = i\epsilon^{ijk}\sigma^k, \quad [S^\alpha, S^\beta] = i\epsilon^{ijk}S^k.$$

Consider $d \in \mathbb{N}$ spin s particles. The state space of the j-particle is $\mathcal{H}_j = V^{(m)} = V^{(2s)} \simeq \mathbb{C}^{2s+1}$ and the state space of the composite system consisting of all the particles is therefore $\mathcal{H} = \bigotimes_{j=1}^d \mathcal{H}_j = \bigotimes_{j=1}^d \mathbb{C}^{2s+1}$. We call *chain* the composite system of particles if we place the particles on a line.

Chapter 2

Tensor network varieties

Tensor network varieties are varieties of tensors described by the combinatorial structure of an undirected simple graph. In this chapter, we define the tensor network variety via the language of *graph tensors* , following [VC17, CVZ19]. We exhibit the basic properties of tensor network varieties and we show the equivalence between two different construction of them: their constructions via the graph tensor and via *contractions* of tensors associated to graphs. Finally, we give the definition of tensor network variety associated to hypergraphs and we show that the Secant variety of the Segre variety is isomorphic to a specific tensor network variety associated to a hypergraph.

2.1 Graph tensor and tensor network variety

Definition 2.1.1. An undirected simple graph Γ is a pair $\Gamma = (\mathbf{v}(\Gamma), \mathbf{e}(\Gamma))$ where $\mathbf{v}(\Gamma)$ is the set of vertices, and $\mathbf{e}(\Gamma)$ is a non-empty set of pairs of vertices, called edges. In addition, the graph has no loops on the vertices and no multiple edges between two vertices. More precisely, for every $e \in \mathbf{e}(\Gamma)$, then $e = \{i, j\}$, for $i, j \in \mathbf{v}(\Gamma)$, $i \neq j$; and if it exists $e = \{i, j\} \in \mathbf{e}(\Gamma)$ connecting the vertices i and j, then it is unique. If $e = \{i, j\} \in \mathbf{e}(\Gamma)$, with the notation $e \ni i$ we mean that the edge $e \in \mathbf{e}(\Gamma)$ is incident to the vertex $i \in \mathbf{v}(\Gamma)$.

Let $\Gamma = (\mathbf{v}(\Gamma), \mathbf{e}(\Gamma))$ be an undirected simple graph, with vertex set $\mathbf{v}(\Gamma) = \{1, \dots, d\}$ and edge set $\mathbf{e}(\Gamma) = \{e_1, \dots, e_R\}$. A collection of *bond dimensions* is a set of weights $\mathbf{m} = (m_e : e \in \mathbf{e}(\Gamma))$ on the edges of Γ. Given a collection of bond dimensions $\mathbf{m}$, define the graph tensor associated to Γ as follows.

For an edge $e = \{i_1, i_2\}$, denote by $v_0^{(i)}$, for $i \in \mathbf{v}(\Gamma) \setminus \{i_1, i_2\}$, a generator of $\mathbb{C}^1$ and, for $p = 1, 2$, $\{v_j^{(i_p)} : j = 1, \dots, m_e\}$ two bases of a copy of $\mathbb{C}^{m_e}$. The *unit tensor* defined on

the edge e and bond dimension m_e, is

$$\mathbf{u}_{(e)}(m_e) = \sum_{j=1}^{m_e} v_j^{(i_1)} \otimes v_j^{(i_2)} \otimes \bigotimes_{\substack{i=1,\dots,d \\ i\neq i_1,i_2}} v_0^{(i)} \in \mathbb{C}^{m_e} \otimes \mathbb{C}^{m_e} \otimes \mathbb{C}^1 \otimes \cdots \otimes \mathbb{C}^1,$$

with the superscripts indicating the ordering of the tensor factors. The graph tensor associated to a graph Γ with bond dimensions $\mathbf{m}$ is

$$T(\Gamma, \mathbf{m}) = \boxtimes_{e\in\mathbf{e}(\Gamma)} \mathbf{u}_{(e)}(m_e), \tag{2.1}$$

where $\boxtimes$ denotes the Kronecker product defined in Definition 1.1.2. This is a tensor of order d whose i-th factor has a local structure

$$W_i = \bigotimes_{e\ni i} \mathbb{C}^{m_e}. \tag{2.2}$$

Remark 2.1.2. In coordinates, we may describe the graph tensor $T(\Gamma, \mathbf{m})$ as the tensor product of identity matrices $\mathrm{Id}_{m_e} \in \mathbb{C}^{m_e} \otimes \mathbb{C}^{m_e}$ for $e \in \mathbf{e}(\Gamma)$ laying on the edges of the graph; this product is regarded as a tensor of order d where the i-th factor is the product of the spaces $\mathbb{C}^{m_e}$ incident to vertex i. Note that from this point of view one of the two copies of $\mathbb{C}^{m_e}$ is identified with its dual space $\mathbb{C}^{m_e *}$, see Figure 2.1.

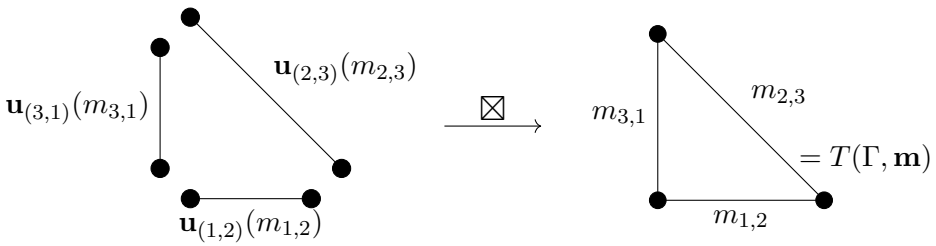


Figure 2.1: Pictorial representation of the construction of the graph tensor $T(\Gamma, \mathbf{m})$ on the triangular graph: $T(\Gamma, \mathbf{m})$ is the tensor product of the three identity matrices $\mathbf{u}_e(m_e)$; regarded as a tensor on three factors.

Remark 2.1.3. Let Γ and Γ' be two graphs on the same set of vertices and with $\mathbf{e}(\Gamma) = \mathbf{e}(\Gamma') \cup \{e\}$. In other words, Γ' is the graph obtained from Γ after removing the edge e. Let $\mathbf{m}$ be a collection of bond dimensions on Γ and let $\mathbf{m}'$ be the collection $\mathbf{m}$ restricted to Γ'. It is clear from the definitions that if $m_e = 1$ then $T(\Gamma, \mathbf{m}) = T(\Gamma', \mathbf{m}')$ because in this case $\mathbf{u}_{(e)}(m_e)$ is a decomposable tensor hence $T \boxtimes \mathbf{u}_{(e)}(m_e) = T$ for every tensor T.

Remark 2.1.3 guarantees that up to modifying the underlying graph, one can always assume $m_e \geq 2$.

Let $n_i \in \mathbb{N}$ be integers associated to the vertices of Γ and let $V_i = \mathbb{C}^{n_i}$. Write $\mathbf{n} = (n_i)_{i=1,\dots,d}$ for the collection of dimensions of the vector spaces V_i. We call the elements of $\mathbf{n}$ *local dimensions* associated to Γ; in physics, they are usually called *physical dimensions*. A triple $(\Gamma, \mathbf{m}, \mathbf{n})$ consisting of a simple graph, a collection of bond dimensions and a collection of local dimensions is a *tensor network*; see Figure 2.2.

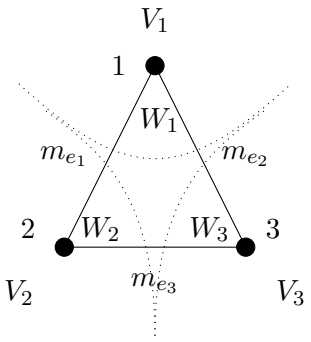


Figure 2.2: The tensor network data pictorial representation. The graph $\Gamma := C_3$ is the cyclic graph with 3 vertices, we call it the triangle graph. The vector spaces associated to the vertices $i = 1, 2, 3$ are $W_i = \bigotimes_{e \ni i} \mathbb{C}^{m_e}$, associated to the collection of bond dimensions $m_e \in \mathbf{m}$, $e \in \mathbf{e}(\Gamma)$ incident to the vertex $i \in \mathbf{v}(\Gamma)$; and V_i, associated to the physical dimensions $n_i \in \mathbf{n}$, i.e. $\dim(V_i) = n_i$.

A tensor network naturally provides the following algebraic variety.

Definition 2.1.4. The *tensor network variety* in $V_1 \otimes \cdots \otimes V_d$ associated to the tensor network $(\Gamma, \mathbf{m}, \mathbf{n})$ is

$$\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \overline{\left\{ T \in V_1 \otimes \cdots \otimes V_d : T = (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}), X_j \in \mathrm{Hom}(W_j, V_j) \right\}},$$

where the closure can be taken equivalently either in the Euclidean or the Zariski topology since we are working over the complex numbers.

The definition of $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ provides a natural parametrization of a Zariski open dense subset, given by the image of the following map

$$\begin{aligned} \widehat{\Phi} : \mathrm{Hom}(W_1, V_1) \oplus \cdots \oplus \mathrm{Hom}(W_d, V_d) &\to V_1 \otimes \cdots \otimes V_d, \\ (X_1, \dots, X_d) &\mapsto (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}). \end{aligned}$$

Let $\mathcal{TNS}^{\Gamma\circ}_{\mathbf{m},\mathbf{n}}$ be the image of the map $\widehat{\Phi}$. The map $\widehat{\Phi}$ factors as follows:

$$\begin{array}{ccc} \bigoplus_{i=1}^{d} \operatorname{Hom}(W_i, V_i) & \xrightarrow{\ \mu\ } & \operatorname{Hom}(W_1 \otimes \cdots \otimes W_d, V_1 \otimes \cdots \otimes V_d) \\ & \searrow^{\widehat{\Phi}} & \downarrow{\scriptstyle \overline{\Phi}} \\ & & V_1 \otimes \cdots \otimes V_d \end{array}$$

where μ is the d-linear map defined as $\mu(X_1, \ldots, X_d) = X_1 \otimes \cdots \otimes X_d$. We denote the image of μ by

$$\operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d) := \operatorname{Im}(\mu); \tag{2.3}$$

If we denote $N_v := \dim W_v$ for $v \in \mathbf{v}(\Gamma)$, then

$$\begin{aligned} \mathbb{P}(\operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)) &= \nu(\mathbb{P}(\operatorname{Hom}(W_1, V_1)) \times \cdots \times \mathbb{P}(\operatorname{Hom}(W_d, V_d))) \\ &= \mathcal{S}_{N_1 n_1 - 1, \ldots, N_d n_d - 1}, \end{aligned}$$

where ν is the Segre embedding of $\mathbb{P}(\operatorname{Hom}(W_1, V_1)) \times \cdots \times \mathbb{P}(\operatorname{Hom}(W_d, V_d))$ in the space $\bigotimes_{i=1}^{d} \operatorname{Hom}(W_i, V_i) \simeq \operatorname{Hom}(W_1 \otimes \cdots \otimes W_d, V_1 \otimes \cdots \otimes V_d)$ and $\mathcal{S}_{N_1 n_1 - 1, \ldots, N_d n_d - 1}$ is the Segre variety; c.f. Definition 1.1.5. Therefore the image of μ is the cone over the Segre variety, i.e. $\widehat{\mathcal{S}}_{N_1 n_1 - 1, \ldots, N_d n_d - 1} = \operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$, and its affine dimension is

$$\begin{aligned} \dim(\operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)) &= \sum_{i=1}^{d} (\dim(\operatorname{Hom}(W_i, V_i)) - 1) + 1 \\ &= \sum_{i=1}^{d} \dim(\operatorname{Hom}(W_i, V_i)) - d + 1 \\ &= \sum_{v \in \mathbf{v}(\Gamma)} N_v n_v - d + 1. \end{aligned}$$

The map

$$\overline{\Phi} : \operatorname{Hom}(W_1 \otimes \cdots \otimes W_d, V_1 \otimes \cdots \otimes V_d) \to V_1 \otimes \cdots \otimes V_d$$

is the evaluation at the graph tensor, therefore the restriction of $\overline{\Phi}$ to the subvariety $\operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$ provides a parametrization of $\mathcal{TNS}^{\Gamma\circ}_{\mathbf{m},\mathbf{n}}$; denote this restriction by

$$\begin{aligned} \Phi : \operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d) &\to V_1 \otimes \cdots \otimes V_d \\ (X_1 \otimes \cdots \otimes X_d) &\mapsto (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}). \end{aligned} \tag{2.4}$$

The set $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ is an irreducible algebraic variety [YL18]. Indeed, the Segre variety $\mathcal{S}_{N_1 n_1 - 1, \ldots, N_d n_d - 1}$ is irreducible and then its affine cone $\widehat{\mathcal{S}}_{N_1 n_1 - 1, \ldots, N_d n_d - 1}$ is irreducible. Since the map Φ is a morphism of varieties, then $\operatorname{Im}(\Phi)$ is a constructible set and $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ is irreducible. It is known that if the graph Γ is a tree, i.e. a graph without loops, then the closure in the definition of $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ is not needed, but if Γ contains cycles then there are examples where it is necessary to take the closure of the set $\mathcal{TNS}^{\Gamma\circ}_{\mathbf{m},\mathbf{n}}$ in order to get an algebraic variety, c.f. [LQY12, CLVW20, BLF22].

Remark 2.1.5. If $\mathbf{m}$ and $\mathbf{m}'$ are two collections of bond dimensions on Γ such that $m'_e \leq m_e$ for every edge $e \in \mathbf{e}(\Gamma)$, then $\mathcal{TNS}^{\Gamma}_{\mathbf{m}',\mathbf{n}} \subseteq \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$.

Indeed, assume $T \in \mathcal{TNS}^{\Gamma\circ}_{\mathbf{m}',\mathbf{n}}$, then there exist $X'_i \in \mathrm{Hom}(W'_i, V_i)$, for $i = 1, \dots, d$, such that $T = (X'_1 \otimes \cdots \otimes X'_d) \cdot T(\Gamma, \mathbf{m}')$. Since $m'_e \leq m_e$ for every edge $e \in \mathbf{e}(\Gamma)$, $W_i = \bigotimes_{e\in\mathbf{e}(\Gamma)} \mathbb{C}^{m'_e} \subseteq \bigotimes_{e\in\mathbf{e}(\Gamma)} \mathbb{C}^{m_e}$. In general, if W' is a vector subspace of W and $X : W' \to V$ a linear map, then X can be extended linearly to W. Denote $N' = \dim(W') \leq N = \dim(W)$ and assume that $\{w'_1, \dots, w'_{N'}\}$ is a basis of W'. We extend the basis of W' to a basis $\{w'_1, \dots, w'_{N'}, w_1, \dots, w_{N-N'}\}$ of W. Define $X : W \to V$ such that $X(w'_j) = X'(w'_j)$ for $1 \leq j \leq N'$ and $X(w_j) = 0$ for $1 \leq j \leq N - N'$. For every $i = 1, \dots, d$, $X'_i : W'_i \to V_i$ can be extended to $X_i : W_i \to V_i$ and clearly, by construction, $T = (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m})$, that is $T \in \mathcal{TNS}^{\Gamma\circ}_{\mathbf{m},\mathbf{n}}$.

Notation 2.1.6. Let Γ be an undirected simple graph and let $\mathbf{e}(\Gamma) = \{e_1, \dots, e_R\}$ be the set of edges. Define $m_j := m_{e_j}$ for $j = 1, \dots, R$, so that $\mathbf{m} = (m_1, \dots, m_R)$ is the set of bond dimensions associated to the edges of Γ. For the local structure of the graph tensor (2.2) and for Remark 2.1.2, every edge is associated to $\mathbb{C}^{m_j *} \otimes \mathbb{C}^{m_j}$. Denote by α_j the formal index associated to the vector space $\mathbb{C}^{m_j}$ (and $\mathbb{C}^{m_j *}$) for $j = 1, \dots, R$, i.e. $\alpha_j \in \{1, \dots, m_j\}$ and define the set of all indices associated to the (edges of the) underlining graph: $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_R\}$. Consider a vertex $i \in \mathbf{v}(\Gamma)$ and assume $e_1, \dots, e_k \ni i$, i.e. $e_1, \dots, e_k$ are the edges incident to vertex $i \in \mathbf{v}(\Gamma)$. We denote by $\boldsymbol{\alpha}_{|_i} := \{\alpha_1, \dots, \alpha_k\} = \{\alpha_j : e_j \ni i\}$ the corresponding set of indices.

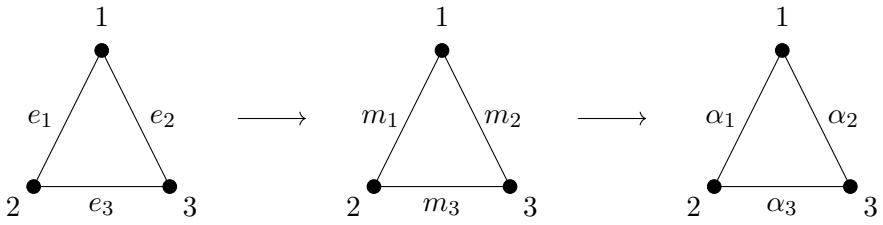


Figure 2.3: Pictorial example of Remark 2.1.6. Let Γ be the triangle graph with $\mathbf{e}(\Gamma) = \{e_1, e_2, e_3\}$. The set of indices associated to Γ is $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \alpha_3\}$, $\alpha_i = 1, \dots, m_i$, with m_i bond dimension associated to edge e_i. Then $\boldsymbol{\alpha}_{|_1} = \{\alpha_1, \alpha_2\}$, $\boldsymbol{\alpha}_{|_2} = \{\alpha_1, \alpha_3\}$, $\boldsymbol{\alpha}_{|_3} = \{\alpha_2, \alpha_3\}$. Every α_j for $j = 1, 2, 3$ belongs to two sets $\boldsymbol{\alpha}_{|_i}$,$\boldsymbol{\alpha}_{|_k}$, with $i \neq k$.

Let $i \in \mathbf{v}(\Gamma)$ be a vertex of Γ and assume $e_1, \dots, e_k \ni i$, i.e. $e_1, \dots, e_k$ are the edges incident to vertex $i \in \mathbf{v}(\Gamma)$. By Notation 2.1.6 we denote the corresponding set of indices by $\boldsymbol{\alpha}_{|_i} := \{\alpha_1, \dots, \alpha_k\} = \{\alpha_j : e_j \ni i\}$ and we define

$$e_{\boldsymbol{\alpha}_{|_i}} := e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \in \bigotimes_{j=1}^{k} \mathbb{C}^{m_j} \simeq W_i.$$

With this notation, we can fix a basis $\{e_{\boldsymbol{\alpha}_{|_i}}\}$ of $W_i = \bigotimes_{e\ni i} \mathbb{C}^{m_e}$, for every vertex

$i = 1, \dots, d$. The graph tensor can be written as

$$T(\Gamma, \mathbf{m}) = \sum_{\alpha \in \boldsymbol{\alpha}} \bigotimes_{i=1}^{d} e_{\boldsymbol{\alpha}_{|_i}} \in \bigotimes_{i=1}^{d} W_i.$$

Remark 2.1.7. For construction, every index $\alpha_j \in \{\boldsymbol{\alpha}_{|_i}\}$ is such that $\alpha_j \in \{\boldsymbol{\alpha}_{|_k}\}$, for one and only one $k \in \{1, \dots, d\}$, $k \neq i$.

We fix the basis $\{v_j\}_{j=1}^{n_i}$ of V_i and we denote by $\{e^*_{\boldsymbol{\alpha}_{|_i}}\}$ the basis of W_i^*. Then $X_i \in \mathrm{Hom}(W_i, V_i) \simeq W_i^* \otimes V_i$ is written in coordinates as

$$X_i = \sum_{j=1}^{n_i} \sum_{\alpha \in \boldsymbol{\alpha}_{|_i}} (X_i)^j_{\boldsymbol{\alpha}_{|_i}} e^*_{\boldsymbol{\alpha}_{|_i}} \otimes v_j.$$

The tensor $T = (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m})$ in these coordinates is

$$\begin{aligned}
T &= (X_1 \otimes \cdots \otimes X_d) \cdot \left(\sum_{\alpha \in \boldsymbol{\alpha}} \bigotimes_{i=1}^{d} e_{\boldsymbol{\alpha}_{|_i}} \right) \\
&= \left(\sum_{j_1, \dots, j_d = 1}^{n_1, \dots, n_d} \sum_{\alpha \in \boldsymbol{\alpha}_{|_1}} \cdots \sum_{\alpha \in \boldsymbol{\alpha}_{|_d}} (X_1)^{j_1}_{\boldsymbol{\alpha}_{|_1}} \dots (X_d)^{j_d}_{\boldsymbol{\alpha}_{|_d}} e^*_{\boldsymbol{\alpha}_{|_1}} \otimes \cdots \otimes e^*_{\boldsymbol{\alpha}_{|_d}} \otimes v_{j_1 \dots j_d} \right) \cdot \sum_{\alpha \in \boldsymbol{\alpha}} \bigotimes_{i=1}^{d} e_{\boldsymbol{\alpha}_{|_i}} \\
&= \sum_{j_1, \dots, j_d = 1}^{n_1, \dots, n_d} \left(\sum_{\alpha \in \boldsymbol{\alpha}} \sum_{\alpha \in \boldsymbol{\alpha}_{|_1}} \cdots \sum_{\alpha \in \boldsymbol{\alpha}_{|_d}} (X_1)^{j_1}_{\boldsymbol{\alpha}_{|_1}} \dots (X_d)^{j_d}_{\boldsymbol{\alpha}_{|_d}} e^*_{\boldsymbol{\alpha}_{|_1}}(e_{\boldsymbol{\alpha}_{|_1}}) \otimes \cdots \otimes e^*_{\boldsymbol{\alpha}_{|_d}}(e_{\boldsymbol{\alpha}_{|_d}}) \right) \otimes v_{j_1 \dots j_d} \\
&= \sum_{j_1, \dots, j_d = 1}^{n_1, \dots, n_d} \left(\sum_{\alpha \in \boldsymbol{\alpha}} (X_1)^{j_1}_{\boldsymbol{\alpha}_{|_1}} \dots (X_d)^{j_d}_{\boldsymbol{\alpha}_{|_d}} \right) v_{j_1} \otimes \cdots \otimes v_{j_d}, \qquad (2.5)
\end{aligned}$$

where the internal sum is taken accordingly to the given rule of Remark 2.1.7.

2.1.1 Another construction of the tensor network variety

By Definition 2.1.4, an element of $\mathcal{TNS}^{\Gamma\circ}_{\mathbf{m},\mathbf{n}}$ is given by the application of a multilinear map on a specific tensor, i.e. the graph tensor, given in Equation (2.1). The most common construction of the tensor network variety consists instead in building an element of $\mathcal{TNS}^{\Gamma\circ}_{\mathbf{m},\mathbf{n}}$ as the contraction of a collection of tensors in a way prescribed by the given graph, see for example [Orú14, LQY12]. More precisely, let $\Gamma = (\mathbf{v}(\Gamma), \mathbf{e}(\Gamma))$ be an undirected simple graph. To every vertex $v \in \mathbf{v}(\Gamma)$ we associate a tensor $X_v \in \mathbb{C}^{n_v} \otimes \bigotimes_{e \ni v} \mathbb{C}^{m_e}$, with $\mathbf{n} = (n_v : v \in \mathbf{v}(\Gamma))$ local dimensions and $\mathbf{m} = (m_e : e \in \mathbf{e}(\Gamma))$ bond dimensions. We get a collection of tensors, one for each vertex. We consider the tensor product of all these tensors

$$\bigotimes_{v \in \mathbf{v}(\Gamma)} X_v \in \bigotimes_{v \in \mathbf{v}(\Gamma)} \mathbb{C}^{n_v} \otimes \bigotimes_{e \in \mathbf{e}(\Gamma)} \left(\mathbb{C}^{m_e} \otimes \mathbb{C}^{m_e} \right).$$

Then a contraction map contracts $\mathbb{C}^{m_e}$ with its dual $\mathbb{C}^{m_e *}$, for every $e \in \mathbf{e}(\Gamma)$, (after the identification of one copy of $\mathbb{C}^{m_e}$ with its dual $\mathbb{C}^{m_e *}$). Finally, an element of the tensor network variety is defined as the image of the application of the contraction map on the element $\bigotimes_{v \in \mathbf{v}(\Gamma)} X_v$. We describe the second construction in detail since it will be used in the thesis and we show that the two definitions of tensor network varieties are equivalent.

Firstly, we recall the notion of contraction map between isomorphic vector spaces.

Definition 2.1.8. Let V be a vector space over $\mathbb{C}$. The contraction map between V and the dual space $V^* := \mathrm{Hom}(V, \mathbb{C})$ is the linear map

$$\begin{aligned} C : V^* \otimes V &\to \mathbb{C} \\ f \otimes w &\mapsto f(w), \end{aligned}$$

given by the bilinear form

$$\begin{aligned} V^* \times V &\to \mathbb{C} \\ (f, w) &\mapsto f(w). \end{aligned}$$

Let $\{v_i\}_{i=1}^n$ be a basis of V, then $f = \sum_{i=1}^n f_i v_i^* \in V^*$ and $w = \sum_{i=1}^n w_i v_i \in V$, for $f_i, w_i \in \mathbb{C}$, $i = 1, \dots, n$. By linearity

$$\begin{aligned} C(f \otimes w) = C(\sum_{i,j=1}^n f_i v_i^* \otimes w_j v_j) &= \sum_{i,j=1}^n f_i w_j \; C(v_i^* \otimes v_j) \\ = \sum_{i,j=1}^n f_i w_j \cdot v_i^*(v_j) &= \sum_{i,j=1}^n \delta_{i,j} \cdot f_i w_j = \sum_{i=1}^n f_i w_i. \end{aligned}$$

Definition 2.1.9. Given two tensor spaces $W = W_1 \otimes \cdots \otimes W_m$ and $V = V_1 \otimes \cdots \otimes V_n$ such that $W_1 \simeq V_1^*$, the *contraction map* between V_1^* and W_1 is $\mathcal{C} : W \otimes V \to W_2 \otimes \cdots \otimes W_n \otimes V_2 \otimes \cdots \otimes V_n$, given by the bilinear map defined on rank 1 tensors as follows

$$\begin{aligned} W \times V &\to W_2 \otimes \cdots \otimes W_n \otimes V_2 \otimes \cdots \otimes V_n \\ (w_1 \otimes \cdots \otimes w_m, v_1 \otimes \cdots \otimes v_n) &\mapsto (w_1(v_1))(w_2 \otimes \cdots \otimes w_m \otimes v_2 \otimes \cdots \otimes v_n). \end{aligned}$$

Fix a basis of V_1 and the dual basis of V_1^* and let $T_{i_1 \dots i_n}$ and $S_{j_1 \dots j_m}$ be the coordinates of tensors $T \in V$ and $S \in W$, respectively. Then their contraction is $\mathcal{C}(S \otimes T) = C \in W_2 \otimes \cdots \otimes W_n \otimes V_2 \otimes \cdots \otimes V_n$ with coordinates given by

$$\begin{aligned} C_{j_2 \dots j_m i_2 \dots i_n} &= \sum_{k,l=1}^{\dim(V_1)} \delta_{kl} \; S_{l j_2 \dots j_m} T_{k i_2 \dots i_n} \\ &= \sum_{k=1}^{\dim(V_1)} S_{k j_2 \dots j_m} T_{k i_2 \dots i_n}. \end{aligned} \tag{2.6}$$

Contraction map. Given the data of the tensor network $(\Gamma, \mathbf{m}, \mathbf{n})$, we have defined for every vertex $i \in \mathbf{v}(\Gamma)$ the vector space

$$W_i = \bigotimes_{e \ni i} U_{m_e}.$$

with either $U_{m_e} = \mathbb{C}^{m_e}$ or the dual copy $U_{m_e} = \mathbb{C}^{m_e *}$.

By the isomorphism $\mathrm{Hom}(W_i, V_i) \simeq W_i^* \otimes V_i$, the element $X_i \in \mathrm{Hom}(W_i, V_i)$ can be seen as a matrix $X_i \in W_i^* \otimes V_i$, for every $i = 1, \dots, d$.

Fix an edge $e' \in \Gamma$. For construction, see Remark 2.1.7, we can select two vertices $j, k \in \mathbf{v}(\Gamma)$, $j \neq k$, such that $\mathbb{C}^{m_{e'}}$ is a factor of W_j and $(\mathbb{C}^{m_{e'}})^*$ is a factor of W_k. We highlight this two factors of the domain of the map:

$$\begin{aligned}
\bigotimes_{i=1}^{d}(W_i^* \otimes V_i) &= (W_1^* \otimes V_1) \otimes \cdots \otimes (W_j^* \otimes V_j) \otimes \cdots \otimes (W_k^* \otimes V_k) \otimes \cdots \otimes (W_1^* \otimes V_1) \\
&\simeq (W_1^* \otimes V_1) \otimes \cdots \otimes \Big(\bigotimes_{e \ni j} U_e \otimes V_j \Big) \otimes \cdots \otimes \Big(\bigotimes_{e \ni k} U_e \otimes V_k \Big) \otimes \cdots \otimes (W_1^* \otimes V_1) \\
&\simeq (\mathbb{C}^{m_{e'}} \otimes \bigotimes_{\substack{e \ni j \\ e \neq e'}} U_e \otimes V_j) \otimes \Big(\mathbb{C}^{m_{e'}*} \otimes \bigotimes_{\substack{e \ni k \\ e \neq e'}} U_e \otimes V_k \Big) \otimes \bigotimes_{\substack{i=1 \\ i \neq j,k}}^{d} (W_i^* \otimes V_i) \\
&\simeq (\mathbb{C}^{m_{e'}*} \otimes \mathbb{C}^{m_{e'}}) \otimes \Big(\bigotimes_{\substack{e \ni j \\ e \neq e'}} U_e \otimes V_j \Big) \otimes \Big(\bigotimes_{\substack{e \ni k \\ e \neq e'}} U_e \otimes V_k \Big) \otimes \bigotimes_{\substack{i=1 \\ i \neq j,k}}^{d} (W_i^* \otimes V_i)
\end{aligned}$$

This can be done for every edge $e' \in \mathbf{e}(\Gamma)$ of the graph leading to

$$\bigotimes_{i=1}^{d}(W_i^* \otimes V_i) \simeq \bigotimes_{e \in \mathbf{e}(\Gamma)} (\mathbb{C}^{m_e *} \otimes \mathbb{C}^{m_e}) \otimes \bigotimes_{i=1}^{d} V_i.$$

Finally, we consider the *contraction map*:

$$\overline{\phi} : \bigotimes_{i=1}^{d}(W_i^* \otimes V_i) \simeq \bigotimes_{e \in \mathbf{e}(\Gamma)} (\mathbb{C}^{m_e *} \otimes \mathbb{C}^{m_e}) \otimes \bigotimes_{i=1}^{d} V_i \to V_1 \otimes \cdots \otimes V_d, \tag{2.7}$$

which, for every $e \in \mathbf{e}(\Gamma)$, contracts the factors $\mathbb{C}^{m_e *} \otimes \mathbb{C}^{m_e}$, as prescribed in Definition 2.1.9. The image of contraction map $\overline{\phi}$ will be written in coordinates in the proof of the following proposition.

Proposition 2.1.10. *The map $\overline{\phi}$ is well defined and the definition of the tensor network variety associated to $(\Gamma, \mathbf{m}, \mathbf{n})$ given in Definition 2.1.4 is equivalently given by*

$$\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \overline{\Big\{ T \in V_1 \otimes \cdots \otimes V_d : T = \overline{\phi}(X_1 \otimes \cdots \otimes X_d), X_i \in W_i^* \otimes V_i \Big\}},$$

where the closure can be taken equivalently either in the Euclidean or the Zariski topology.

Proof. The contraction map given in Equation (2.7)

$$\overline{\phi} : \bigotimes_{i=1}^{d} (W_i^* \otimes V_i) \simeq \bigotimes_{e \in \mathbf{e}(\Gamma)} (\mathbb{C}^{m_e *} \otimes \mathbb{C}^{m_e}) \otimes \bigotimes_{i=1}^{d} V_i \to V_1 \otimes \cdots \otimes V_d,$$

contracts, for every $e \in \mathbf{e}(\Gamma)$, the factors $\mathbb{C}^{m_e *} \otimes \mathbb{C}^{m_e}$, as prescribed in Definition 2.1.9. The map is well defined since for every vertex $v \in \mathbf{v}(\Gamma)$ and every edge incident to v, $e \ni v$, there exists one and only one $w \in \mathbf{e}(\Gamma)$, $w \neq v$, $w \ni e$ such that the vector space $\mathbb{C}^{m_e}$ (factor of W_v) admits one isomorphic dual copy $\mathbb{C}^{m_e *}$ (factor of W_w).

Consider the restriction of $\overline{\phi}$ to $\mathrm{Im}\,(\mu) \subset \mathrm{Hom}(W_1 \otimes \cdots \otimes W_d, V_1 \otimes \cdots \otimes V_d) \simeq (W_1^* \otimes V_1) \otimes \cdots \otimes (W_d^* \otimes V_d)$, c.f. Equation (2.3), which is given by:

$$\begin{aligned} \phi : \mathrm{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d) &\to V_1 \otimes \cdots \otimes V_d \\ (X_1 \otimes \cdots \otimes X_d) &\mapsto \phi(X_1 \otimes \cdots \otimes X_d). \end{aligned}$$

with $X_j \in W_j^* \otimes V_j$ for every $j = 1, \ldots, d$. Clearly we have that

$$\mathrm{Im}\,(\phi) = \left\{ T \in V_1 \otimes \cdots \otimes V_d : T = \overline{\phi}(X_1 \otimes \cdots \otimes X_d), X_i \in W_i^* \otimes V_i \right\}.$$

The maps ϕ and the parametrization of the variety Φ of Equation (2.4) have the same domain and the same codomain. If we prove that $\mathrm{Im}\,(\phi) = \mathrm{Im}\,(\Phi)$, then also their closure will coincide and we can conclude.

Let $T \in \mathrm{Im}\,(\phi)$, $T = \phi(X_1 \otimes \cdots \otimes X_d)$, for some $X_i \in W_i^* \otimes V_i$. We fix the basis $\{v_j\}_{j=1}^{n_i}$ of V_i and we denote by $\{e^*_{\boldsymbol{\alpha}_{|_i}}\}$ the basis of W_i^*; see Notation 2.1.6. Then $X_i \in W_i^* \otimes V_i$ is written in coordinates as

$$X_i = \sum_{j=1}^{n_i} \sum_{\alpha \in \boldsymbol{\alpha}_{|_i}} (X_i)^j_{\boldsymbol{\alpha}_{|_i}} e^*_{\boldsymbol{\alpha}_{|_i}} \otimes v_j.$$

Following the indices contraction given in Definition 2.1.9, Equation (2.6), it is now straightforward to see that $T = \phi(X_1 \otimes \cdots \otimes X_d)$ is

$$T = \sum_{j_1, \ldots, j_d = 1}^{n_1, \ldots, n_d} \left(\sum_{\alpha \in \boldsymbol{\alpha}} (X_1)^{j_1}_{\boldsymbol{\alpha}_{|_1}} \ldots (X_d)^{j_d}_{\boldsymbol{\alpha}_{|_d}} \right) v_{j_1} \otimes \cdots \otimes v_{j_d},$$

where the sum is taken over all indices of $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_R\}$, accordingly to Remark 2.1.7.

Finally, by the canonical isomorphisms $\iota_j : W_j^* \otimes V_j \to \mathrm{Hom}(W_j, V_j)$ we have that, if $T \in \mathrm{Im}\,(\phi)$, $T = \phi(X_1 \otimes \cdots \otimes X_d)$, for $X_j \in W_j^* \otimes V_j$, $j = 1, \ldots, d$, then $T \in \mathrm{Im}\,(\Phi)$, $T = (X'_1 \otimes \cdots \otimes X'_d) \cdot T(\Gamma, \mathbf{m})$, for $X'_j \simeq \iota_j(X_j) \in \mathrm{Hom}(W_j, V_j)$, with $j = 1, \ldots, d$, c.f. Equation (2.5).

Vice versa, if $T \in \text{Im }(\Phi)$, $T = \Phi(X_1 \otimes \cdots \otimes X_d) = (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m})$, for $X_j \in \text{Hom}(W_j, V_j)$, $j = 1, \ldots, d$, then $T \in \text{Im }(\phi)$, $T = \phi(X_1' \otimes \cdots \otimes X_d')$, for $X_j' \simeq \iota_j^{-1}(X_j) \in W_j^* \otimes V_j$, $j = 1, \ldots, d$.

Therefore Im $(\phi) =$ Im (Φ), and this concludes the proof. □

2.1.2 Matrix product states

We explicitly compute the image of the maps Φ and ϕ for the tensor network variety associated to the cyclic graph and known as matrix product state. From the viewpoint of the second construction presented, the matrix product state variety is determined by the contraction of tensors of order 3 (usually denoted by the letter A instead of X).

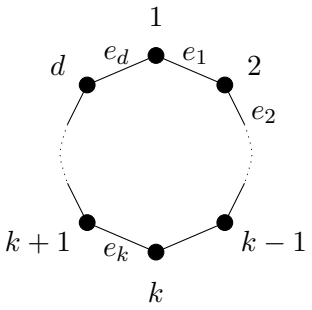


Figure 2.4: The cyclic graph C_d, with d vertices $\mathbf{v}(\Gamma) = \{1, \ldots, d\}$ and d edges $\mathbf{e} = \{e_1, \ldots, e_d\}$, such that $e_i = \{i, i+1\}$, $i = 1, \ldots, d$.

Consider the cyclic graph Γ with d vertices $\mathbf{v}(\Gamma) = \{1, \ldots, d\}$ and d edges $\mathbf{e} = \{e_1, \ldots, e_d\}$. Every edge $e_i \in \mathbf{e}(\Gamma)$ is such that $e_i = \{i, i+1\}$, i.e. it connects the vertices i and $(i+1)$, with vertices $d+1$ and 1 identified, see Figure 2.4. For every edge $e_i \in \mathbf{e}(\Gamma)$, fix $m_i \in \mathbb{N}$.

Fix the bases $\{e_{(i)}^j : j = 1, \ldots, m_i\}$ and $\{e_j^{(i+1)} : j = 1, \ldots, m_i\}$ of $\mathbb{C}^{m_i *}$ and $\mathbb{C}^{m_i}$ respectively and $v_0^{(k)}$, a generator of $\mathbb{C}^1$ for $k \neq i, i+1$. We define the unit tensor on $e_i \in \mathbf{e}(\Gamma)$, as

$$\mathbf{u}_{e_i}(m_i) = \sum_{j=1}^{m_i} e_{(i)}^j \otimes e_j^{(i+1)} \otimes \Big(\bigotimes_{k \neq i, i+1} v_0^{(k)} \Big) \in \mathbb{C}^{m_i *} \otimes \mathbb{C}^{m_i} \otimes \Big(\bigotimes_{k \neq i, i+1} \mathbb{C}^1 \Big).$$

In this case, we explicitly compute the graph tensor

$$\begin{aligned} T(\Gamma, \mathbf{m}) &= \boxtimes_{i=1}^d \mathbf{u}_{(e_i)}(m_i) \\ &= \sum_{\substack{j_i = 1 \ldots, m_i \\ i = 1 \ldots, d}} (e_{j_d}^{(1)} \otimes e_{(1)}^{j_1}) \otimes (e_{j_1}^{(2)} \otimes e_{(2)}^{j_2}) \otimes \cdots \otimes (e_{j_{d-1}}^{(d)} \otimes e_{(d)}^{j_d}) \in \bigotimes_{i=1}^d W_i. \end{aligned}$$

The local structure on the vertex $k \in \mathbf{v}(\Gamma)$ is given by

$$W_k = \mathbb{C}^{m_{k-1}} \otimes \mathbb{C}^{m_k *},$$

with the convention $\mathbb{C}^{m_0} = \mathbb{C}^{m_d}$; we refer to Figure 2.5 for a pictorial representation.

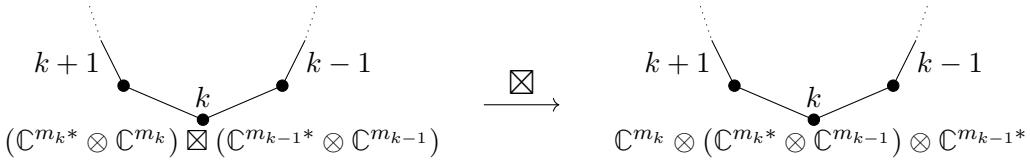


Figure 2.5: Pictorial representation of the spaces associated to the edges of the graph (left) and then associated to the vertices of the graph (right), after the Kronecker product.

We fix d complex vector spaces V_i of dimension n_i, for $i = 1, \dots d$, and the respective canonical bases $\{v_j^{(i)} : i = 1, \dots, d, j = 1, \dots n_i\}$.

We identify $X_j' = \iota_j^{-1}(X_j) \simeq X_j$ via the isomorphism $\iota_j : W_j^* \otimes V_j \to \mathrm{Hom}(W_j, V_j)$ and we show $\Phi(X_1 \otimes \cdots \otimes X_d) = (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}) = \phi(X_1 \otimes \cdots \otimes X_d)$.

In what follows, we use the Einstein convention, i.e. there is implicit summation over an index variable appearing twice. The elements $X_i \in \mathrm{Hom}(W_i, V_i) \simeq W_i^* \otimes V_i \simeq \mathbb{C}^{m_{i-1}*} \otimes \mathbb{C}^{m_i} \otimes V_i$ can be written in coordinates as

$$X_i = (x_i)_\beta^{\alpha,\gamma} e_{(i)}^\beta \otimes e_\alpha^{(i)} \otimes v_\gamma^{(i)}.$$

We recall that $e_\alpha^{(i)}(e_{(i+1)}^\beta) = \delta(i)_\alpha^\beta = e_{(i+1)}^\beta(e_\alpha^{(i)})$. It is straightforward to see that an element in the image of Φ is given by

$$\begin{aligned}(X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}) &= \left[(x_1)_{\alpha_d}^{\alpha_1\gamma_1}(x_2)_{\alpha_1}^{\alpha_2\gamma_2} \dots (x_d)_{\alpha_{d-1}}^{\alpha_d\gamma_d}\right] v_{\gamma_1}^{(1)} \otimes \cdots \otimes v_{\gamma_d}^{(d)} \\ &= \mathrm{Tr}(X_1^{\gamma_1} \cdots X_d^{\gamma_d}) v_{\gamma_1}^{(1)} \otimes \cdots \otimes v_{\gamma_d}^{(d)}.\end{aligned}$$

Now, the domain of the contraction map $\overline{\phi}$ can be reshaped into

$$\bigotimes_{k=1}^{d} (W_i^* \otimes V_i) \simeq \bigotimes_{k=1}^{d} (\mathbb{C}^{m_{i-1}*} \otimes \mathbb{C}^{m_i} \otimes V_i) \simeq (\mathbb{C}^{m_1 *} \otimes \mathbb{C}^{m_1}) \otimes V_1 \otimes \cdots \otimes (\mathbb{C}^{m_d *} \otimes \mathbb{C}^{m_d}) \otimes V_d.$$

The map $\overline{\phi}$ and its restriction ϕ are defined by contracting the spaces $\mathbb{C}^{m_i *} \otimes \mathbb{C}^{m_i}$ for every $i = 1, \dots, d$

$$\begin{aligned}\phi : (\mathbb{C}^{m_1 *} \otimes \mathbb{C}^{m_1}) \otimes V_1 \otimes \cdots \otimes (\mathbb{C}^{m_d *} \otimes \mathbb{C}^{m_d}) \otimes V_d &\to V_1 \otimes \cdots \otimes V_d \\ X_1 \otimes \cdots \otimes X_d &\mapsto \phi(X_1 \otimes \cdots \otimes X_d).\end{aligned}$$

Each contraction generates a Kronecker delta $\delta(i)_{\alpha_i}^{\beta_i}$, as shown in Equation (2.6). In coordinates we get

$$\phi\left(\bigotimes_{k=1}^{d}(x_i)_{\beta_{i-1}}^{\alpha_i,\gamma_i}e_{(i)}^{\beta_{i-1}}\otimes e_{\alpha_i}^{(i)}\otimes v_{\gamma_i}^{(i)}\right)=\left[(x_1)_{\alpha_1}^{\alpha_d,\gamma_1}(x_2)_{\alpha_2}^{\alpha_1,\gamma_2}\dots(x_d)_{\alpha_d}^{\alpha_{d-1},\gamma_d}\right]v_{\gamma_1}^{(1)}\otimes\dots\otimes v_{\gamma_d}^{(d)}$$
$$=\mathrm{Tr}\left(X_1^{\gamma_1}\cdots X_d^{\gamma_d}\right)v_{\gamma_1}^{(1)}\otimes\dots\otimes v_{\gamma_d}^{(d)}.$$

2.1.3 Hypergraph and secant varieties of Segre variety

In this thesis, we always restrict to simple graphs but the theory generalizes to more general notions of graphs; we refer to [CLVW20, CGMZ21] for the definitions and the basics in the general setting. In this subsection, we state the definition of tensor network variety associated to a hypergraph and we show that secant varieties of the Segre variety, c.f. 1.1.1, can be seen as tensor network varieties. The proof of the equivalence between secant varieties of the Segre variety and a tensor network variety associated to particular tensor networks is based on a characterization of tensors of bounded border rank given in [Lan17, CGMZ21].

Definition 2.1.11. An undirected hypergraph $\mathcal{H}$ is a pair $\mathcal{H}=(\mathbf{v}(\mathcal{H}),E(\mathcal{H}))$ where $\mathbf{v}(\mathcal{H})$ is the set of vertices, and $E(\mathcal{H})$ is a set of non-empty subsets of $\mathbf{v}(\mathcal{H})$ called hyperedges. Therefore, $E(\mathcal{H})$ is a subset of $\mathcal{P}(\mathbf{v}(\mathcal{H}))\setminus\{\emptyset\}$, where $\mathcal{P}(\mathbf{v}(\mathcal{H}))$ denotes the power set of $\mathbf{v}(\mathcal{H})$, i.e. the set of all the subsets of $\mathbf{v}(\mathcal{H})$.

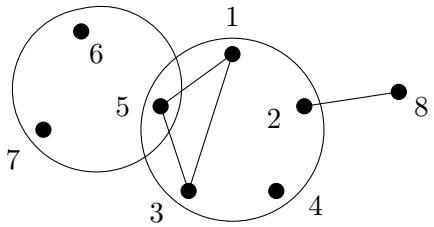


Figure 2.6: The hypergraph $\mathcal{H}$, with $\mathbf{v}(\mathcal{H})=\{1,\dots,8\}$ and $E(\mathcal{H})=\{I_1,I_2,I_3,I_4\}$, with $I_1=\{1,2,3,4,5\}$, $I_2=\{5,6,7\}$, $I_3=\{1,3,5\}$ and (simple edge) $I_4=\{2,8\}$.

Let $\mathcal{H}=(\mathbf{v}(\mathcal{H}),E(\mathcal{H}))$ be a *hypergraph*, with vertex set $\mathbf{v}(\mathcal{H})=\{1,\dots,d\}$ and edge set $E(\mathcal{H})$; see Figure 2.6 for a pictorial representation. For every hyperedge $I\in E(\mathcal{H})$, let $m_I\in\mathbb{N}$ be an integer weight.

For every hyperedge $I=\{i_1,\dots,i_p\}\in E(\mathcal{H})$, define the tensor

$$\mathbf{u}_{(I)}(m_I):=\sum_{j=1}^{m_I}e_j^{(i_1)}\otimes\dots\otimes e_j^{(i_p)}\otimes\bigotimes_{i'\notin I}e_0^{(i')}\in\bigotimes_{i\in I}\mathbb{C}^{n_I}\otimes\bigotimes_{i'\notin I}\mathbb{C}^1,$$

where $e_1^{(i)},\dots,e_{m_I}^{(i)}$ is a basis of $\mathbb{C}^{n_I}$ for every $i\in I$, and $e_0^{(i')}$ is a fixed basis element of $\mathbb{C}^1$ for $i'\notin I$.

Definition 2.1.12. The *graph tensor* associated to the hypergraph $\mathcal{H} = (\mathbf{v}(\mathcal{H}), E(\mathcal{H}))$ with weights $\mathbf{m} = (m_I : I \in E(\mathcal{H}))$ is defined as the Kronecker product

$$T(\Gamma, \mathbf{m}) := \boxtimes_{I \in E(\mathcal{H})} \mathbf{u}_{(I)}(m_I).$$

In this way $T(\mathcal{H}, \mathbf{m})$ is a d-tensor in $W_1 \otimes \cdots \otimes W_d$ whose j-th factor has local structure $W_j = \left(\bigotimes_{I \ni j} \mathbb{C}^{m_I}\right) \otimes \left(\bigotimes_{I \not\ni j} \mathbb{C}^1\right)$ and has dimension $\dim W_j = \prod_{I \ni j} n_I$.

Let $n_i \in \mathbb{N}$ be integers associated to the vertices of $\mathcal{H}$ and let $V_i = \mathbb{C}^{n_i}$. Write $\mathbf{n} = (n_i)_{i=1,\dots,d}$ for the d-uple of dimensions of the vector spaces V_i. Definition 2.1.4 can be restated in the case in which $\mathcal{H}$ is a hypergraph.

Definition 2.1.13. The *tensor network variety* in $V_1 \otimes \cdots \otimes V_d$ associated to the tensor network $(\mathcal{H}, \mathbf{m}, \mathbf{n})$ is defined as

$$\mathcal{TNS}^{\mathcal{H}}_{\mathbf{m},\mathbf{n}} = \overline{\left\{ T \in V_1 \otimes \cdots \otimes V_d : T = (X_1 \otimes \cdots \otimes X_d) \cdot T(\mathcal{H}, \mathbf{m}), X_j \in \mathrm{Hom}(W_j, V_j) \right\}},$$

where the closure can be taken equivalently either in Euclidean or Zariski topology.

Remark 2.1.14. The variety $\mathcal{TNS}^{\mathcal{H}}_{\mathbf{m},\mathbf{n}}$ is invariant under the action of $G = GL(V_1) \times \cdots \times GL(V_d)$, defined by

$$(g_1, \dots, g_d) \cdot ((X_1 \otimes \cdots \otimes X_d) \cdot T(\mathcal{H}, \mathbf{m})) = ((g_1 X_1) \otimes \cdots \otimes (g_d X_d)) \cdot T(\mathcal{H}, \mathbf{m}),$$

with $(g_1, \dots, g_d) \in G$.

We recall the definition of *degeneration* of a tensor $T \in V_1 \otimes \cdots \otimes V_d$ under the action of $G = \bigtimes_{i=1}^d GL(V_i)$, c.f. [Lan17, CGMZ21].

Definition 2.1.15. Given two tensors $T, S \in V_1 \otimes \cdots \otimes V_d$, we say that T is a degeneration of S, and write $T \trianglelefteq S$, if

$$T \in \overline{G \cdot S}$$

that is, T belongs to the closure (equivalently either in Zariski or Euclidean topology) of the G-orbit of S.

Definition 2.1.16. Let $V_i = \mathbb{C}^r$, for every $i = 1, \dots, d$. The *d-tensor* or *rank-r unit tensor* is defined as

$$u_d(r) = \sum_{j=1}^{r} e_j^{(1)} \otimes \cdots \otimes e_j^{(d)} \in (\mathbb{C}^r)^{\otimes d}, \tag{2.8}$$

where $\{e_1^{(i)}, \dots, e_r^{(i)}\}$ is a fixed basis of V_i.

We have the following characterization of tensors of bounded border rank in terms of degeneration of the unit tensor, c.f. [Lan17, Section 3.3.1]. Given $T \in (\mathbb{C}^r)^{\otimes d}$, then

$$\mathrm{brk}(T) \leq r \quad \text{iff} \quad T \trianglelefteq u_d(r).$$

This equivalence leads directly to the following Proposition.

Proposition 2.1.17. *Let $V = V_1 \otimes \cdots \otimes V_d$, such that $\dim(V_i) = n_i + 1$ and let $\mathcal{S}_{n_1,\ldots,n_d} \subseteq \mathbb{P}V$ be the Segre variety. Let $r \in \mathbb{N}$ be such that $r \leq \min\{n_i + 1 \ : \ i = 1, \ldots, d\}$. Then*

$$\sigma_r(\mathcal{S}_{n_1,\ldots,n_d}) = \mathcal{TNS}^{\mathcal{H}_d}_{r,\mathbf{n}},$$

where $\mathcal{H}_d = (\mathbf{v}(\mathcal{H}_d), E(\mathcal{H}_d))$ is the hypergraph with d vertices $\mathbf{v}(\mathcal{H}_d) = \{1, \ldots, d\}$ and $E(\mathcal{H}_d) = \{I\}$, one single hyperedge $I = \{1, \ldots, d\}$ of associated weight $r \in \mathbb{N}$.

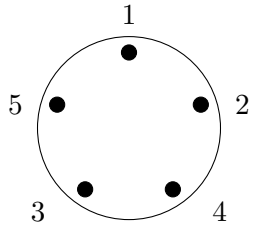


Figure 2.7: Example of a hypergraph with only one hyperedge: the graph is $\mathcal{H}_5$, with $\mathbf{v}(\mathcal{H}_5) = \{1, 2, 3, 4, 5\}$ and $E(\mathcal{H}_5) = \{I\}$, with only one hyperedge $I = \{1, 2, 3, 4, 5\}$ containing all the vertices.

Proof. We simply notice that $T(\mathcal{H}_d, r) = u_d(r)$. Indeed, since there is only one hyperedge $I = \{1, \ldots, d\}$, of weight $r \in \mathbb{N}$, connecting all the vertices of $\mathcal{H}_d$, the graph tensor coincides with the unit tensor defined on hyperedge I:

$$T(\mathcal{H}_d, r) = u_I(m_I) = \sum_{j=1}^{r} e_j^{(1)} \otimes \cdots \otimes e_j^{(d)} \in (\mathbb{C}^r)^{\otimes d} \subseteq \bigotimes_{i=1}^{d} (\mathbb{C}^{n_i}),$$

which is exactly $u_d(r)$, c.f. (2.8). By construction

$$\begin{aligned} \mathcal{TNS}(\mathcal{H}_d, r, \mathbf{n}) &= \overline{(GL_r \times \cdots \times GL_r) \cdot T(\mathcal{H}_d, r)} \\ &= \overline{(GL_r \times \cdots \times GL_r) \cdot u_d(r)}. \end{aligned}$$

Therefore, given $T \in (\mathbb{C}^r)^{\otimes d}$, then $T \in \mathcal{TNS}(\mathcal{H}_d, r, \mathbf{n})$ if and only if $T \trianglelefteq u_d(r)$ if and only if $\mathrm{brk}(T) \leq r$ if and only if $T \in \sigma_r(\mathcal{S}_{n_1,\ldots,n_d} \mathcal{G}(X))$. □

2.2 Gauge subgroup

From now on the graph we consider is always an undirected simple graph, see Definition 2.1.1. In this section, we define the gauge subgroup, which is a group acting on the vector spaces associated to the edges of the graph. This means that the gauge subgroup acts on both the graph tensor and the factor W_v of X_v, for every $v \in \mathbf{v}(\Gamma)$, i.e. it acts on both the graph tensor and the domain of the Parametrization (2.4). The gauge subgroup will be a fundamental object in Chapter 3 because the study of its action on the domain of the parametrization will lead to determine an upper bound on the dimension of the

tensor network variety and an exact value of the dimension in a particular range of the parameters.

Let Γ be a graph and $\mathbf{m} = (m_e : e \in \mathbf{e}(\Gamma))$ be a collection of bond dimensions. Let $T = T(\Gamma, \mathbf{m}) \in W_1 \otimes \cdots \otimes W_d$ be the associated graph tensor. Fix an edge $e = \{i_1, i_2\} \in \mathbf{e}(\Gamma)$: by definition of $T(\Gamma, \mathbf{m})$ there exist vector spaces U_e, W'_{i_1}, W'_{i_2} such that $W_{i_1} = U_e \otimes W'_{i_1}$, and $W_{i_2} = U_e^* \otimes W'_{i_2}$ where $\dim U_e = m_e$ and the tensor product structure depends on the local structure at the vertices i_1 and i_2, see Figure 2.8. The group $GL(U_e) \times GL(U_e^*)$ acts on the factor $U_e \otimes U_e^*$ of $W_{i_1} \otimes W_{i_2}$ with kernel the central subgroup $Z_e = \{(\lambda \mathrm{Id}_{U_e}, \lambda^{-1} \mathrm{Id}_{U_e^*}) : \lambda \in \mathbb{C}^*\}$.

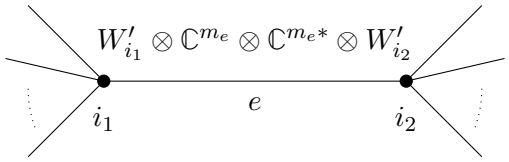


Figure 2.8: Fix an edge $e \in \mathbf{e}(\Gamma)$ and m_e bond dimension. The vector space U_e is either $\mathbb{C}^{m_e}$ or $\mathbb{C}^{m_e *}$.

This defines a homomorphism

$$\Psi_e : (GL(U_e) \times GL(U_e^*))/Z_e \to GL(W_{k_1} \otimes W_{k_2}) \to G(W_k : k \in \mathbf{v}(\Gamma)).$$

We recall that the group $G(W_k : k \in \mathbf{v}(\Gamma))$ is a subgroup of $GL(\bigotimes_{k \in \mathbf{v}(\Gamma)} W_k)$ acting faithfully on $\bigotimes_{k \in \mathbf{v}(\Gamma)} W_k$; c.f. Equation (1.1), where the group $G(W_k : k \in \mathbf{v}(\Gamma))$ is defined.

As e varies among the edges of Γ, the images of the different Ψ_e's commute and therefore they induce a homomorphism

$$\Psi : \bigtimes_{e \in \mathbf{e}(\Gamma)} (GL(U_e) \times GL(U_e^*))/Z_e \to G(W_k : k \in \mathbf{v}(\Gamma)), \tag{2.9}$$

which turns out to be injective. Regrouping the factors, we can write

$$\mathrm{Im}\,(\Psi) = \left[\bigtimes_{v \in \mathbf{v}(\Gamma)} H_v \right] / \left[\bigtimes_{e \in \mathbf{e}(\Gamma)} Z_e \right]$$

where $H_v = \bigtimes_{v \ni e} GL_{m_e}$; here GL_{m_e} is $GL(U_e)$ or $GL(U_e^*)$ depending on whether U_e or U_e^* is the tensor factor appearing in W_v. With abuse of notation, we will denote by H_v the quotient $\left\langle H_v, \left[\bigtimes_{e \in \mathbf{e}(\Gamma)} Z_e\right] \right\rangle / \left[\bigtimes_{e \in \mathbf{e}(\Gamma)} Z_e\right] \subseteq \mathrm{Im}\,(\Psi)$ as well, where for subgroups H, K, one denotes by $\langle H, K \rangle$ the subgroup generated by H and K.

Let $GL_{m_e}^{\Delta} \subseteq GL(U_e) \times GL(U_e^*)$ be the subgroup lying "diagonally", that is $GL_{m_e}^{\Delta} = \{(A, A^{-1^T}) \in GL(U_e) \times GL(U_e^*) : A \in GL(U_e)\}$; its image under the homomorphism Ψ_e is a copy of $PGL_{m_e} \subseteq G(W_1, \ldots, W_d)$ called *gauge subgroup* on the edge e.

The image of the homomorphism Ψ restricted to $\bigtimes_e GL_{m_e}^{\Delta}$ is a subgroup

$$\mathcal{G}_{\Gamma,\mathbf{m}} \simeq \bigtimes_{e\in\mathbf{e}(\Gamma)} PGL_{m_e} \subseteq G(W_k : k \in \mathbf{v}(\Gamma))$$

called *gauge subgroup* of Γ with bond dimensions $\mathbf{m}$. Denote by $\mathbf{g}_{\Gamma,\mathbf{m}}$ the Lie algebra of the gauge subgroup $\mathcal{G}_{\Gamma,\mathbf{m}}$ of Γ.

Definition 2.2.1. Consider the vertex $v \in \mathbf{v}(\Gamma)$. Regard $X_v \in \mathrm{Hom}(W_v, V_v)$ as a tensor in $W_v^* \otimes V_v = \left(\bigotimes_{e\ni v} U_e\right) \otimes V_v$, where $U_e = \mathbb{C}^{m_e}$ or $U_e = \mathbb{C}^{m_e *}$.

Suppose $\{e \in \mathbf{e}(\Gamma) \,:\, e \ni v\} = \{e_1, \ldots, e_k\}$, with associated bond dimensions $(m_{e_j})_{j=1,\ldots,k}$. Given $g_v = (g_{m_{e_1}}, \ldots, g_{m_{e_k}}) \in \bigtimes_{e\ni v} PGL_{m_e} \subseteq \mathcal{G}_{\Gamma,\mathbf{m}}$, define

$$g_v \cdot X_v = X_v \left(g_{m_{e_1}} \otimes \cdots \otimes g_{m_{e_k}}\right) \in \bigotimes_{e\ni v} U_e \otimes V_v.$$

Let $g \in \mathcal{G}_{\Gamma,\mathbf{m}}$ and $X = X_1 \otimes \cdots \otimes X_d \in \mathrm{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$, then the *gauge action* is given by

$$g \cdot X = (g_{v_1} \cdot X_1) \otimes \cdots \otimes (g_{v_d} \cdot X_d).$$

The action of $\mathcal{G}_{\Gamma,\mathbf{m}}$ is trivial on $\bigotimes_{v\in\mathbf{v}(\Gamma)} V_v$, and, for every $v \in \mathbf{v}(\Gamma)$, it is the multilinear multiplication of g_v on the factor W_v of X_v.

Example 2.2.2. Let Γ be the path graph with two vertices $\mathbf{v}(\Gamma) = \{1, 2\}$ and a single edge $e = \{1, 2\}$. Fix $\mathbf{m} = (m)$ and $\mathbf{n} = (n_1, n_2)$. Then the graph tensor is $T(\Gamma, \mathbf{m}, \mathbf{n}) = \sum_{j=1}^{m} e_j^{(1)} \otimes e_j^{(2)} \in \mathbb{C}^m \otimes \mathbb{C}^m$. Given $X = X_1 \otimes X_1 \in \mathrm{Hom}(\mathbb{C}^m, \mathbb{C}^m; V_1, V_2)$, then

$$\begin{aligned}
\Phi(X_1 \otimes X_2) &= (X_1 \otimes X_2) \cdot \sum_{j=1}^{m} e_j^{(1)} \otimes e_j^{(2)} = \sum_{j=1}^{m} X_1(e_j^{(1)}) \otimes X_2(e_j^{(2)}) \\
&= \sum_{j=1}^{m} X_1(e_j^{(1)})(X_2(e_j^{(2)}))^t = \sum_{j=1}^{m} X_1(e_j^{(1)})(e_j^{(2)})^t X_2^t \\
&= X_1(\mathrm{Id}_m) X_2^t = X_1 X_2^t.
\end{aligned}$$

Given $A \in PGL_m = \mathcal{G}_{\Gamma,\mathbf{m}}$, the action given in Definition 2.2.1 is

$$A \cdot X = (X_1 A^{-1}) \otimes (X_2 A^t).$$

Notice that $\Phi(X_1 A^{-1}) \otimes (X_2 A^t) = X_1 A^{-1} A X_2^t = X_1 X_2^t = \Phi(X_1 \otimes X_2)$.

2.3 Isotropy group of the graph tensor

The main result of this section states that the identity component of the isotropy group of the graph tensor $G_{T(\Gamma,\mathbf{m})}$ coincides with the gauge subgroup. Equivalently, the graph

tensor is stabilized only by the gauge subgroup. In order to claim this, we prove that the isotropy Lie algebra of the graph tensor $T(\Gamma, \mathbf{m})$ coincides with Lie algebra of the gauge subgroup. We refer to Chapter 1, Section 1.1.2, for definitions and results on the isotropy group of tensors under the action of products of linear groups. The result generalizes the known result for the iterated matrix multiplication tensor [dG78, Ges16], that is the graph tensor associated to the cycle graph. We prove a more general form of this fact in Theorem 2.3.3; the result on graph tensors will follow via an inductive argument in Corollary 2.3.4.

The following result is immediate from the definitions:

Lemma 2.3.1. *The gauge subgroup on the edge $e \in \mathbf{e}(\Gamma)$, $PGL_{m_e} \subseteq G(W_1, \ldots, W_d)$, stabilizes $T(\Gamma, \mathbf{m})$.*

Proof. Let $e = \{i_1, i_2\}$, so that PGL_{m_e} only acts on the copy of $U_e \otimes U_e^* \subseteq W_{i_1} \otimes W_{i_2}$. In fact, because of the structure of $T(\Gamma, \mathbf{m})$, PGL_{m_e} only acts on the Kronecker factor $\mathbf{u}_e = \mathrm{Id}_{m_e}^{(i_1,i_2)} \otimes \left(\bigotimes_{j \neq i_1, i_2} v_0^{(j)}\right) \in U_e \otimes U_e^* \otimes \bigotimes_{j \neq i_1, i_2} \mathbb{C}^1$.

For $A \in PGL_{m_e}$, we have $A \cdot \mathbf{u}_e = (A^{-1}\mathrm{Id}_{m_e}^{(i_1,i_2)} A) \otimes \left(\bigotimes_{j \neq i_1, i_2} v_0^{(j)}\right) = \mathbf{u}_e$. Therefore PGL_{m_e} stabilizes $\mathbf{u}_e$. □

Corollary 2.3.2. *The graph tensor $T(\Gamma, \mathbf{m})$ is stabilized by $\mathcal{G}_{\Gamma,\mathbf{m}}$.*

Theorem 2.3.3. *Let $T' \in \mathbb{C}^1 \otimes \bigotimes_{j=1}^d W_j'$ be a concise tensor of order $d+1$. Let $\Sigma = (\mathbf{v}(\Sigma), \mathbf{e}(\Sigma))$ be a the graph on $d+1$ vertices $\mathbf{v}(\Sigma) = \{0, \ldots, d\}$ with edge set $\mathbf{e}(\Sigma) = \{e_1, \ldots, e_k\}$, where $e_j = \{0, j\}$. Let $\mathbf{m} = (m_j : j = 1, \ldots, k)$ be a set of bond dimensions on Σ. Let*

$$S := T(\Sigma, \mathbf{m}) \in \mathbb{C}^{m_1 \cdots m_k} \otimes \mathbb{C}^{m_1} \otimes \cdots \otimes \mathbb{C}^{m_k} \otimes \mathbb{C}^1 \otimes \cdots \otimes \mathbb{C}^1$$

be the associated graph tensor. Let $T = S \boxtimes T' \in V_0 \otimes \cdots \otimes V_d$. Then

$$\mathfrak{g}_T = \mathfrak{h}_{T'} + \mathbf{g}_{\Sigma,\mathbf{m}} \subseteq \mathfrak{g}(V_0, \ldots, V_d)$$

where $\mathbf{g}_{\Sigma,\mathbf{m}}$ is the Lie algebra of the gauge subgroup $\mathcal{G}_{\Sigma,\mathbf{m}}$ of Σ and $\mathfrak{h}_{T'}$ is the isotropy Lie algebra of T' in $\mathfrak{g}(\mathbb{C}^1, W_1', \ldots, W_d')$.

A pictorial representation of the tensor T of Theorem 2.3.3 is given in Figure 2.9.

Proof. The inclusion $\mathfrak{h}_{T'} + \mathbf{g}_{\Sigma,\mathbf{m}} \subseteq \mathfrak{g}_T$ is immediate.

For $j = 1, \ldots, k$, write $V_j = U_j \otimes W_j'$ where $U_j = \mathbb{C}^{m_j}$. Write $V_0 = \mathbb{C}^1 \otimes U_1^* \otimes \cdots \otimes U_k^*$. For $j = 1, \ldots, k$, let $\{u_{i_j}^j : i_j = 1, \ldots, m_j\}$ be a basis of the U_j; let $\{u_{i_1,\ldots,i_k}^0 : i_j = 1, \ldots, m_j\}$ be the basis of $V_0 \simeq U_1^* \otimes \cdots \otimes U_k^*$ dual to the induced basis $\{u_{i_1}^1 \otimes \cdots \otimes u_{i_k}^k : i_j = 1, \ldots, m_j\}$ of $U_1 \otimes \cdots \otimes U_k$. Therefore

$$S = \sum_{i_1,\ldots,i_k} u_{i_1,\ldots,i_k}^{(0)} \otimes u_{i_1}^{(1)} \otimes \cdots \otimes u_{i_k}^{(k)} \otimes u_0^{k+1} \otimes \cdots \otimes u_0^d$$

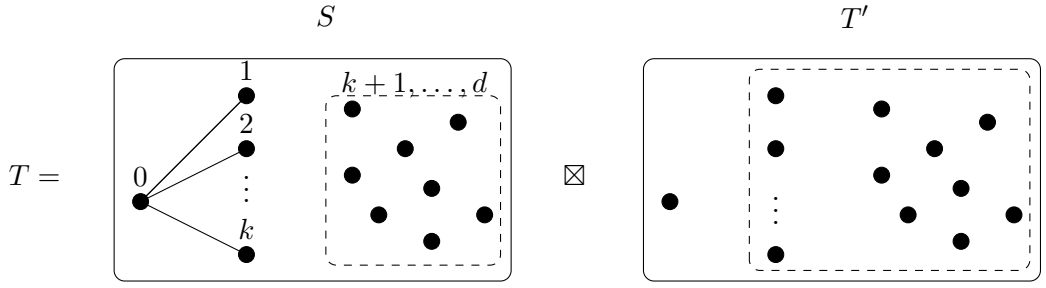


Figure 2.9: The tensor T in Theorem 2.3.3: the Kronecker product of the tensor S, associated to the star graph $\mathbf{e}(\Sigma)$ centered at vertex 0, and a tensor T' whose 0-th factor is 1-dimensional.

where for $j = k+1, \dots, d$, u_0^j is a generator of the corresponding $\mathbb{C}^1$ factor.

Let $\mathbf{X} = (X_0, \dots, X_d) \in \mathfrak{g}(V_0, \dots, V_d)$. Suppose $\mathbf{X} \in \mathfrak{g}_T$, that is $\mathbf{X}.T = 0$. By Leibniz's rule $\mathbf{X}.T = \sum_{j=0}^{d} X_j.T = 0$.

Write $X_0 = ((x^0)^{i'_1,\dots,i'_k}_{i_1,\dots,i_k})$ in the chosen basis: we have

$$X_0.T = (X_0.S) \boxtimes T' = \left[\sum_{\substack{i_1,\dots,i_k \\ i'_1,\dots,i'_k}} (x^0)^{i'_1,\dots,i'_k}_{i_1,\dots,i_k} u^0_{i'_1,\dots,i'_k} \otimes u^1_{i_1} \otimes \cdots \otimes u^k_{i_k} \right] \boxtimes T'.$$

For $j = 1, \dots, k$, write $X_j \in \mathfrak{gl}(V_j)$ as $X_j = \sum \Delta_j^{(\rho)} \boxtimes \Theta_j^{(\rho)}$ where $\Delta_j^{(\rho)} = ((\delta^{\rho,j})^{i_j}_{i'_j}) \in \mathfrak{gl}(U_j)$ and $\Theta_j^{(\rho)} \in \mathfrak{gl}(W'_j)$; then

$$\begin{aligned} X_j.T &= \sum_{\rho} \left[\Delta_j^{(\rho)}.S \right] \boxtimes \left[\Theta_j^{(\rho)}.T' \right] = \\ &= \sum_{\rho} \left[\sum_{i_1,\dots,i_k,i'_j} u^0_{i_1,\dots,i_k} \otimes u^1_{i_1} \otimes \cdots \otimes (\delta^{\rho,j})^{i_j}_{i'_j} u^j_{i_j} \otimes \cdots \otimes u^k_{i_k} \right] \boxtimes \left[\Theta_j^{(\rho)}.T' \right]. \end{aligned}$$

If $j > k$, then $V_j = \mathbb{C}^1 \otimes W'_j$ and we have $X_j.T = S \boxtimes X_j.T'$.

For indices $i_1^*, \dots, i_k^*, \widetilde{\imath}_1, \dots, \widetilde{\imath}_k$, write $\mathbf{X}.T = u^0_{i_1^*,\dots,i_k^*} \otimes u^1_{\widetilde{\imath}_1} \otimes \cdots \otimes u^k_{\widetilde{\imath}_k} \boxtimes T^{i_1^*,\dots,i_k^*}_{\widetilde{\imath}_1,\dots,\widetilde{\imath}_k}$ for tensors $T^{i_1^*,\dots,i_k^*}_{\widetilde{\imath}_1,\dots,\widetilde{\imath}_k} \in W'_1 \otimes \cdots \otimes W'_d$. Since $u^0_{i_1^*,\dots,i_k^*} \otimes u^1_{\widetilde{\imath}_1} \otimes \cdots \otimes u^k_{\widetilde{\imath}_k}$ are linearly independent, the condition $\mathbf{X}.T = 0$ is equivalent to $T^{i_1^*,\dots,i_k^*}_{\widetilde{\imath}_1,\dots,\widetilde{\imath}_k} = 0$ for every $i_1^*, \dots, i_k^*, \widetilde{\imath}_1, \dots, \widetilde{\imath}_k$.

Note that if $(i_1^*, \dots, i_k^*)$ and $(\widetilde{\imath}_1, \dots, \widetilde{\imath}_k)$ differ in at least two entries, then $T^{i_1^*,\dots,i_k^*}_{\widetilde{\imath}_1,\dots,\widetilde{\imath}_k}$ only depends on X_0: indeed, the summands $X_j.T$ for $j \neq 0$ only give rise to terms where $(i_1^*, \dots, i_k^*)$ and $(\widetilde{\imath}_1, \dots, \widetilde{\imath}_k)$ differ in at most one entry. Write $X_0 = X'_0 + X''_0$ where X'_0 is the component where $(i_1^*, \dots, i_k^*)$ and $(\widetilde{\imath}_1, \dots, \widetilde{\imath}_k)$ differ in at least two entries and X''_0

is the complementary component. In particular, X_0' is the only component of $\mathbf{X}$ which contributes to $T^{i_1^*,\dots,i_k^*}_{\widetilde{\imath}_1,\dots,\widetilde{\imath}_k}$ when the two sets of indices differ in at least two entries. By linearity, the discussion above shows $X_0'.T = 0$. Since T is concise, Lemma 1.1.20 implies that $X_0' = 0$. This shows

$$X_0 = Y_1 \otimes \mathrm{id}_{U_2^* \otimes \cdots \otimes U_k^*} + \cdots + \mathrm{id}_{U_1^* \otimes \cdots \otimes U_{k-1}^*} \otimes Y_k$$

with $Y_j \in \mathfrak{gl}(U_j^*)$. Hence, we may renormalize $\mathbf{X}$ using $\mathbf{g}_{\Gamma,\mathbf{m}}$ and obtain $X_0 = 0$. In particular, we reduced the analysis to $\mathbf{X} \in \mathfrak{g}(V_j : j \neq 0)$.

Consider $\mathbf{X} \in \mathfrak{g}_T \cap \mathfrak{g}(V_j : j \neq 0)$. By Lemma 1.1.20, we have

$$\mathfrak{g}_T \cap \mathfrak{g}(V_j : j \neq 0) = \bigcap_{R \in \mathrm{Im}\ (\mathrm{Flat}(T))} \mathfrak{g}_R, \tag{2.10}$$

where $\mathrm{Flat}(T) : V_0^* \to V_1 \otimes \cdots \otimes V_d$ is the 0-th flattening map. For indices $(i_1, \dots, i_k)$, write $T'(i_1, \dots, i_k) = \mathrm{Flat}(T)(u^{(0)}_{i_1,\dots,i_k}) = u^{(1)}_{i_1} \otimes \cdots \otimes u^{(k)}_{i_k} \boxtimes T'$. The intersection in (2.10) can be reduced to a set of generators of Im (Flat(T)); therefore we obtain

$$\mathfrak{g}_T \cap \mathfrak{g}(V_j : j \neq 0) = \bigcap_{i_1,\dots,i_k} \mathfrak{g}_{T'(i_1,\dots,i_k)}.$$

Since $T'(i_1, \dots, i_k)$ is not concise in $V_1 \otimes \cdots \otimes V_d$, we have $\mathfrak{g}_{T'(i_1,\dots,i_k)} = \mathfrak{h}_{T'(i_1,\dots,i_k)} \oplus \mathfrak{p}_{i_1,\dots,i_k}$, where $\mathfrak{h}_{T'(i_1,\dots,i_k)}$ is the annihilator of $T'(i_1, \dots, i_k)$ in $\mathfrak{gl}(\langle u^1_{i_1}\rangle \otimes W_1') \oplus \cdots \oplus \mathfrak{gl}(\langle u^k_{i_k}\rangle \otimes W_k') \oplus \mathfrak{gl}(V_{k+1}) \oplus \cdots \oplus \mathfrak{gl}(V_d)$ and $\mathfrak{p}_{i_1,\dots,i_k}$ is the parabolic subspace which annihilates $(u^1_{i_1} \otimes W_1') \otimes \cdots \otimes (u^k_{i_k} \otimes W_k') \otimes V_{k+1} \otimes \cdots \otimes V_d$, that is

$$\mathfrak{p}_{i_1,\dots,i_k} = \left[(\langle u^1_{i_1}\rangle^\perp \otimes W_1'^*) \otimes (U_1 \otimes W_1')\right] \oplus \cdots \oplus \left[(\langle u^k_{i_k}\rangle^\perp \otimes W_k'^*) \otimes (U_k \otimes W_k')\right].$$

Since $T'(i_1, \dots, i_k) = u^1_{i_1} \otimes \cdots \otimes u^k_{i_k} \boxtimes T'$, we have

$$\mathfrak{h}_{T'(i_1,\dots,i_k)} = \mathrm{Id}_{\langle u^1_{i_1}\rangle \otimes \cdots \otimes \langle u^k_{i_k}\rangle} \otimes \mathfrak{g}_{T'},$$

regarded as a subalgebra acting on the subspace $u^1_{i_1} \otimes \cdots \otimes u^k_{i_k} \otimes W_1' \otimes \cdots \otimes W_k' \otimes V_{k+1} \otimes \cdots \otimes V_d$.

Observe that, as a subspace of $\mathrm{End}(V_1 \otimes \cdots \otimes V_d)$, we have

$$\mathfrak{g}_{T'(i_1,\dots,i_k)} = \left[\mathrm{Id}_{\langle u^1_{i_1}\rangle \otimes \cdots \otimes \langle u^k_{i_k}\rangle} \otimes \mathfrak{g}_{T'}\right] \oplus \mathfrak{p}_{i_1,\dots,i_k} = \left[\mathrm{Id}_{U_1 \otimes \cdots \otimes U_k} \otimes \mathfrak{g}_{T'}\right] \oplus \mathfrak{p}_{i_1,\dots,i_k}.$$

This follows directly from Leibniz rule and the fact that, for every $i_1, \dots, i_k$, $\mathrm{Id}_{U_1 \otimes \cdots \otimes U_k} = \mathrm{Id}_{\langle u^1_{i_1} \otimes \cdots \otimes u^k_{i_k}\rangle} + P_{i_1,\dots,i_k}$ where $P_{i_1,\dots,i_k} \in \mathfrak{p}_{i_1,\dots,i_k}$. We deduce

$$\mathfrak{g}_T \cap \mathfrak{g}(V_j : j \neq 0) = \bigcap_{i_1,\dots,i_k} \left[(\mathrm{Id}_{U_1 \otimes \cdots \otimes U_k} \otimes \mathfrak{g}_{T'}) \oplus \mathfrak{p}_{i_1,\dots,i_k}\right]$$

and by Lemma 1.1.18, we have $\mathfrak{g}_T \cap \mathfrak{g}(V_j : j \neq 0) = (\mathrm{Id}_{U_1 \otimes \cdots \otimes U_k} \otimes \mathfrak{g}_{T'}) \oplus \bigcap_{i_1,\ldots,i_k} \mathfrak{p}_{i_1,\ldots,i_k} = \mathrm{Id}_{U_1 \otimes \cdots \otimes U_k} \otimes \mathfrak{g}_{T'}$ because $\bigcap_{i_1,\ldots,i_k} \mathfrak{p}_{i_1,\ldots,i_k} = 0$.

This concludes the proof, as we showed

$$\mathfrak{g}_T = \mathfrak{g}_T + \mathbf{g}_{\Sigma,\mathbf{m}} = \mathfrak{g}_T \cap \mathfrak{g}(V_j : j \neq 0) + \mathbf{g}_{\Sigma,\mathbf{m}} = \mathfrak{g}_{T'} + \mathbf{g}_{\Sigma,\mathbf{m}}. \qquad \square$$

Applying Theorem 2.3.3 to graph tensors, we deduce the following result:

Corollary 2.3.4. *Let $\Gamma = (\mathbf{v}(\Gamma), \mathbf{e}(\Gamma))$ be a graph with d vertices and let $\mathbf{m} = (m_e : e \in \mathbf{e}(\Gamma))$ be a set of bond dimensions on Γ. Let $T := T(\Gamma, \mathbf{m}) \in \bigotimes_{j=1}^{d} W_j$ be the associated graph tensor. Then the isotropy Lie algebra of T coincides with Lie algebra of the gauge subgroup of Γ; in symbols*

$$\mathfrak{g}_T = \mathbf{g}_{\Gamma,\mathbf{m}}.$$

Proof. We proceed by induction on the number of vertices d. If $d = 1$, the statement is clear as T is a single vector, with trivial isotropy Lie algebra.

Suppose Γ is a graph with $d+1$ vertices and write $\mathbf{v}(\Gamma) = \{0, \ldots, d\}$. Let Σ be the subgraph of Γ given by the edges incident to the vertex 0. In other words $\mathbf{v}(\Sigma) = \{0, \ldots, d\}$, $\mathbf{e}(\Sigma) = \{e \in \mathbf{e}(\Gamma) : 0 \in e\}$. Let Γ' be the graph with $\mathbf{v}(\Gamma') = \{0, \ldots, d\}$ and $\mathbf{e}(\Gamma') = \mathbf{e}(\Gamma) \setminus \mathbf{e}(\Sigma)$ and let $\mathbf{m}'', \mathbf{m}'$ be the corresponding subsets of the collection of bond dimensions $\mathbf{m}$. Write $S = T(\Sigma, \mathbf{m}'')$ and $T' = T(\Gamma', \mathbf{m}')$; then

$$T = S \boxtimes T'.$$

By the induction hypothesis, $\mathfrak{g}_{T'} = \mathbf{g}_{\Gamma',\mathbf{m}'}$ and $\mathfrak{g}_S = \mathbf{g}_{\Sigma,\mathbf{m}''}$. By Theorem 2.3.3

$$\mathfrak{g}_T = \mathfrak{g}_{T'} + \mathbf{g}_{\Sigma,\mathbf{m}''} = \mathbf{g}_{\Gamma',\mathbf{m}'} + \mathbf{g}_{\Sigma,\mathbf{m}''} = \mathbf{g}_{\Gamma,\mathbf{m}},$$

and this concludes the proof. $\square$

Conclusion. We have introduced the definition and some properties of the tensor network variety. We have defined the gauge subgroup $\mathcal{G}_{\Gamma,\mathbf{m}}$ and proved that it coincides with the stabilizer of the graph tensor under the action of $G(W_1, \ldots, W_d)$, i.e. $G_{T(\Gamma,\mathbf{m})} = \mathcal{G}_{\Gamma,\mathbf{m}}$. This result is essential for determining the exact value of the dimension of the tensor network variety in a particular range of parameters, Corollary 3.5.2 in Chapter 3, Section 3.5.

Chapter 3

Dimension

In this chapter, we address the problem of determining the dimension of tensor network varieties. The results of this chapter have appeared in slightly altered form in the paper [BDLG22], co-authored with Alessandra Bernardi and Fulvio Gesmundo, that has been published in Communication in Contemporary Mathematics. The dimension of the variety provides a measure of how large the set of tensors allowing a certain tensor network representation is, which in turn gives a measure of the expressiveness of the tensor network class. We provide a completely general *upper bound* in Theorem 3.0.2. The result is based on a lower bound on the dimension of the generic fiber of the parametrization of the variety. The lower bound is given in Theorem 3.3.1, and it descends from the fact that the generic fiber of the parametrization contains the orbit of a generic element of the domain of the map, under the action of the gauge subgroup. We illustrate how to refine the upper bound in cases relevant for applications in Corollary 3.4.5 and Corollary 3.4.6. In Corollary 3.5.2, we give the exact value of the dimension of the tensor network variety in a particular range of parameters, where it can be realized as the closure of the orbit of the graph tensor under the action of $G(W_1, \dots, W_d)$. Therefore, in this case, the dimension is completely controlled by the dimension of the stabilizer of the graph tensor that, by Corollary 2.3.4 (previous chapter), coincides with the gauge subgroup. Finally, in Section 3.6 we further analyze some cases arising from small values of the parameters, and we provide a more precise calculation of their dimension.

The *dimension* of an irreducible algebraic variety is defined as the dimension of its tangent space at a smooth point. We refer to [Sha94, Ch. 3] for the basic properties of dimension. Recall that $\mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d) := \mathrm{Im}\,(\mu)$, where

$$\begin{aligned} \mu : \bigoplus_{i=1}^{d} \mathrm{Hom}(W_i, V_i) &\to \mathrm{Hom}(W_1 \otimes \cdots \otimes W_d, V_1 \otimes \cdots \otimes V_d) \\ (X_1, \dots, X_d) &\mapsto X_1 \otimes \cdots \otimes X_d. \end{aligned}$$

The parametrization of $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ is therefore given by

$$\begin{aligned}\Phi : \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d) &\to V_1 \otimes \cdots \otimes V_d \\ (X_1 \otimes \cdots \otimes X_d) &\mapsto (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}).\end{aligned}$$

The Theorem of Dimension of the Fibers [Sha94, Thm. 1.25] provides the following expression for the affine dimension of the tensor network variety:

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \dim\left[\mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)\right] - \dim \Phi^{-1}(T) \tag{3.1}$$

where T is a generic tensor in the image of Φ.

We recall that the affine dimension of $\mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)$ is

$$\dim \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d) = \sum_{1=1}^{d} \dim(\mathrm{Hom}(W_i, V_i)) - d + 1.$$

Therefore we want to determine the dimension of the fiber $\dim \Phi^{-1}(T)$ for a generic $T \in \mathrm{Im}\,(\Phi)$. We focus on determining lower bounds for $\dim \Phi^{-1}(T)$, which via Equation (3.1) provide upper bounds for $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$.

We first give the following definitions, which determine the ranges of parameters of the tensor network, c.f. [LQY12].

Definition 3.0.1. Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network. A vertex $v \in \mathbf{v}(\Gamma)$ is called

- *subcritical* if $\prod_{e \ni v} m_e \geq n_v$; *strictly subcritical* if the inequality is strict;
- *supercritical* if $\prod_{e \ni v} m_e \leq n_v$; *strictly supercritical* if the inequality is strict;
- *critical* if v is both subcritical and supercritical.

The tensor network $(\Gamma, \mathbf{m}, \mathbf{n})$ is called [strictly] subcritical (resp. supercritical) if all its vertices are [strictly] subcritical (resp. supercritical).

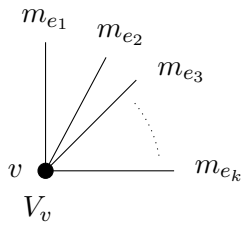


Figure 3.1: A vertex $v \in \mathbf{v}(\Gamma)$, with k incoming edges $\{e_1, \dots, e_k\}$ with associated bond dimensions m_{e_i}, for $i = 1, \dots, k$. Then $\dim W_v = \prod_{e \ni v} m_e$.

The main theorem of the chapter is the following.

Theorem 3.0.2. *Let* $(\Gamma, \mathbf{m}, \mathbf{n})$ *be a tensor network and let* $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ *be the corresponding tensor network variety. Then*

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} \leq$$
$$\min\left\{\sum_{v\in\mathbf{v}(\Gamma)}(n_v\cdot N_v)-d+1-\sum_{e\in\mathbf{e}(\Gamma)}(m_e^2-1)+\dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X),\ \prod_{v\in\mathbf{v}(\Gamma)} n_v\right\},$$

where $N_v=\prod_{e\ni v} m_e$, $X = X_1\otimes\cdots\otimes X_d$ *with* $X_v\in \mathrm{Hom}(W_v, V_v)$ *generic and* $\mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ *is the stabilizer of* X *under the action of the gauge subgroup.*

In the statement of the theorem, $\mathcal{G}_{\Gamma,\mathbf{m}}$ is the gauge subgroup associated to the tensor network, defined in Chapter 2, Section 2.2, and the term $\sum_{e\in\mathbf{e}(\Gamma)}(m_e^2-1)$ is its dimension. The term $\mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ is the stabilizer under the action of the gauge subgroup of a generic d-tuple of linear maps.

The proof of Theorem 3.0.2 consists of different steps.

First, in Section 3.1 we determine a reduction that allows us to assume that the bond dimensions associated to the edges incident to a fixed vertex are *balanced*, in a way made precise in Lemma 3.1.1.

Then, in Section 3.2 we provide a second reduction, proving that the tensor network variety built from a tensor network having strictly supercritical vertices can be realized via a vector bundle construction as a natural extension of the tensor network variety where the strictly supercritical vertices are reduced to be critical.

Finally, in Section 3.3 we provide an upper bound for $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ in the subcritical range that concludes the proof of Theorem 3.0.2. Moreover, we provide the exact value of the dimension $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ in the critical and supercritical ranges in Section 3.5.

3.1 Reduction of bond dimension

We observed in Remark 2.1.3 that we may always assume bond dimensions at least 2. Here, we show that if they are "too unbalanced", then they can be reduced without affecting the dimension of the tensor network variety.

We say that a tensor network $(\Gamma, \mathbf{m}, \mathbf{n})$ has *overabundant bond dimension* if there exist a vertex $v\in\mathbf{v}(\Gamma)$ and an edge $e\in\mathbf{e}(\Gamma)$ incident to v such that

$$m_e > n_v \prod_{e'\ni v, e'\neq e} m_{e'}. \tag{3.2}$$

The following result shows that overabundant bond dimensions do not contribute to the dimension of the tensor network variety.

Lemma 3.1.1. *Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network. Fix $v \in \mathbf{v}(\Gamma)$, let k be the degree of the vertex v and $\{e_1, \dots, e_k\}$ be the edges incident to v; assume $m_{e_1} \le \cdots \le m_{e_k}$. If (3.2) holds for v, that is*

$$m_{e_k} > n_v \cdot m_{e_1} \cdots m_{e_{k-1}},$$

then

$$\mathcal{TNS}_{\mathbf{m},\mathbf{n}} = \mathcal{TNS}_{\overline{\mathbf{m}},\mathbf{n}}$$

where $\overline{\mathbf{m}}$ is defined by $\overline{m}_e = m_e$ if $e \neq e_k$ and $\overline{m}_{e_k} = n_v \cdot m_1 \cdots m_{e_{k-1}}$.

Proof. Let $T \in \mathcal{TNS}_{\mathbf{m},\mathbf{n}} \subseteq V_1 \otimes \cdots \otimes V_d$ be a generic element and let $(X_1, \dots, X_d) \in \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)$ be an element such that $(X_1, \dots, X_d) \cdot T(\Gamma, \mathbf{m}) = T$.

Suppose $v = d$ and $e_j = \{d, j\}$ for $j = 1, \dots, k$. Write $U_j = \mathbb{C}^{m_j}$; let $W_d = U_1^* \otimes \cdots \otimes U_k^*$, so that, for $j = 1, \dots, k$, we have $W_j = U_j \otimes W_j'$ where W_j' depends on the other edges incident to the vertex j.

Regard X_d as a tensor in $W_d^* \otimes V_d = U_1 \otimes \cdots \otimes U_k \otimes V_d$. Since $m_{e_k} > m_{e_1} \cdots m_{e_{k-1}} \cdot n_v$, X_d is not concise on the factor U_k: let $\overline{U}_k \subseteq U_k$ with $\dim \overline{U}_k = m_{e_1} \cdots m_{e_{k-1}} \cdot n_v$ be a subspace such that $X_d \in U_1 \otimes \cdots \otimes U_{k-1} \otimes \overline{U}_k \otimes V_d$. Correspondingly, let $\overline{U}_k^* = U_k^* / \overline{U}_k^{\perp}$. Note that $T(\Gamma, \overline{\mathbf{m}})$ coincides with the image of $T(\Gamma, \mathbf{m})$ via the projection $U_k^* \to \overline{U}_k^*$ on the d-th factor.

Now, define $\overline{W}_d = U_1^* \otimes \cdots \otimes U_{k-1}^* \otimes \overline{U}_k^*$ and $\overline{W}_k = W_k' \otimes \overline{U}_k$. Let $\overline{X}_d = X_d$ be the linear map regarded as an element of $\mathrm{Hom}(\overline{W}_d, V_d)$. Moreover, the space $\mathrm{Hom}(W_k, V_k) = (W_k' \otimes U_k)^* \otimes V_k = W_k'^* \otimes U_k^* \otimes V_k$ naturally projects onto $W_k'^* \otimes \overline{U}_k^* \otimes V_k = \mathrm{Hom}(\overline{W}_k, V_k)$: let $\overline{X}_k$ be the image of X_k under this projection.

Now, one can verify that

$$T = (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}) = (\overline{X}_1 \otimes \cdots \otimes \overline{X}_d) \cdot T(\Gamma, \overline{\mathbf{m}})$$

where $\overline{X}_v = X_v$ if $v \neq k, d$. □

3.2 Reduction for supercritical vertices

The tensor network variety built from a tensor network having strictly supercritical vertices can be realized as an extension of the tensor network variety where the strictly supercritical vertices are reduced to be critical. The reduction of this section appeared already in [LQY12]. We include it here for completeness.

For a vector space V with $\dim V = n$ and an integer $k \le n$, let $\mathbf{G}(k, V)$ be the Grassmannian of k-dimensional linear subspaces of V. Recall that $\dim \mathbf{G}(k, V) = k(n-k)$. The variety $\mathbf{G}(k, V)$ has a *tautological bundle*

$$\sigma : \mathcal{S} \to \mathbf{G}(k, V);$$

the fiber of $\mathcal{S}$ over a point $[E] \in \mathbf{G}(k, V)$ is the plane E itself: $\mathcal{S}_{[E]} = E$.

Proposition 3.2.1. *Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network. Suppose that the vertex $d \in \mathbf{v}(\Gamma)$ is supercritical and write $N = \dim W_d = \prod_{e \ni d} m_e$. Let $\mathbf{n}' = (n'_v : v \in \mathbf{v}(\Gamma))$ be the d-tuple of local dimensions defined by $n'_v = n_v$ if $v \neq d$ and $n'_d = N$.*

Then

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = N(n_d - N) + \dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}'}.$$

Proof. Let $\mathcal{S}_d^{V_1 \otimes \cdots \otimes V_{d-1}}$ be the vector bundle over the Grassmannian $\mathbf{G}(N, V_d)$ whose fiber over a plane $[E]$ is $V_1 \otimes \cdots \otimes V_{d-1} \otimes E$; this is the tautological bundle augmented by the trivial bundle with constant fiber $V_1 \otimes \cdots \otimes V_{d-1}$. Consider the diagram

$$\begin{array}{ccc} & \mathcal{S}_d^{V_1 \otimes \cdots \otimes V_{d-1}} & \\ {\scriptstyle \sigma}\swarrow & & \searrow{\scriptstyle \pi} \\ \mathbf{G}(N, V_d) & & V_1 \otimes \cdots \otimes V_d \end{array}$$

where the second projection π maps an element of the bundle to its fiber component: $([E], T) \mapsto T$. By conciseness, this projection is generically one-to-one.

Consider the subbundle of $\mathcal{S}_d^{V_1 \otimes \cdots \otimes V_{d-1}}$ whose fiber at $[E]$ is $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}'}$ where the d-th factor is identified with E. Let $\underline{\mathcal{TNS}}^{\Gamma}_{\mathbf{m},\mathcal{S}}$ be the total space of this subbundle. We have

$$\dim \underline{\mathcal{TNS}}^{\Gamma}_{\mathbf{m},\mathcal{S}} = \dim \mathbf{G}(N, V_d) + \dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}'} = N(n_d - N) + \dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}'}.$$

The projection π is generically one-to-one and maps $\underline{\mathcal{TNS}}^{\Gamma}_{\mathbf{m},\mathcal{S}}$ surjectively onto $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$. Therefore $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \dim \underline{\mathcal{TNS}}^{\Gamma}_{\mathbf{m},\mathcal{S}}$ and this concludes the proof. □

Iteratively applying Proposition 3.2.1, one can reduce all strictly supercritical vertices to critical vertices.

Theorem 3.2.2. *Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network. For every $v \in \mathbf{v}(\Gamma)$ let $N_v = \prod_{e \ni v} m_e$. Let $\mathbf{n}'$ be the set of local dimensions defined by $n'_v = \min\{N_v, n_v\}$. Then*

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \sum_{v \in \mathbf{v}(\Gamma)} n'_v(n_v - n'_v) + \dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}'}.$$

Note that the tensor network $(\Gamma, \mathbf{m}, \mathbf{n}')$ appearing in Theorem 3.2.2 is, by definition, subcritical.

It remains to understand $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ in the subcritical range.

3.3 Subcritical range and general case

We provide an upper bound for $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ when the tensor network $(\Gamma, \mathbf{m}, \mathbf{n})$ is subcritical. The upper bound is obtained, via Equation (3.1), by determining a lower bound on the dimension of the generic fiber of Φ.

Theorem 3.3.1. *Let* $(\Gamma, \mathbf{m}, \mathbf{n})$ *be a subcritical tensor network. Then the dimension of the generic fiber of the map* Φ *is bounded from below by the dimension of the* $\mathcal{G}_{\Gamma,\mathbf{m}}$*-orbit of a generic element of* $\mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)$.

Proof. Consider $T \in \mathrm{Im}\ (\Phi)$, $T = (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m})$, with $X_1 \otimes \cdots \otimes X_d \in \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)$ a generic element. The fiber of $\Phi : \mathrm{Hom}(W_1, \dots, W_d; V_1 \otimes \cdots \otimes V_d) \to \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ over T is

$$\Phi^{-1}(T) = \{Y_1 \otimes \cdots \otimes Y_d \in \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d) : (Y_1 \otimes \cdots \otimes Y_d) \cdot T(\Gamma, \mathbf{m}) = T\}$$

Since every vertex is subcritical, for every j, a generic element of $\mathrm{Hom}(W_j, V_j)$ is surjective. Let $Y_1 \otimes \cdots \otimes Y_d \in \Phi^{-1}(T)$. By conciseness, Y_j has the same image as X_j, therefore Y_j is surjective as well, and there exists $g \in GL(W_j)$ such that $Y_j = X_j g_j$.

For $X = X_1 \otimes \cdots \otimes X_d$, and $g = g_1 \otimes \cdots \otimes g_d \in G(W_1, \dots, W_d)$, write $g.X = X_1 g_1 \otimes \cdots \otimes X_d g_d$. In particular, if $g \in \mathcal{G}_{\Gamma,\mathbf{m}}$ then

$$\begin{aligned} Y \cdot T(\Gamma, \mathbf{m}) &= (g.X) \cdot T(\Gamma, \mathbf{m}) \\ &= (X_1 \otimes \cdots \otimes X_d)(g_1 \otimes \cdots \otimes g_d) \cdot T(\Gamma, \mathbf{m}) \\ &= X \cdot T(\Gamma, \mathbf{m}) = T, \end{aligned}$$

and the dimension of the fiber is bounded by

$$\begin{aligned} \dim \Phi^{-1}(T) &= \dim\{Y : Y \cdot T(\Gamma, \mathbf{m}) = T\} \\ &= \dim\{g.X : g \in G(W_1, \dots, W_d), (g.X) \cdot T(\Gamma, \mathbf{m}) = T\} \\ &\geq \dim\{g.X : g \in \mathcal{G}_{\Gamma,\mathbf{m}}\} \\ &= \dim(\mathcal{G}_{\Gamma,\mathbf{m}} \cdot X). \end{aligned}$$

Therefore the dimension of the generic fiber is bounded from below by the dimension of the $\mathcal{G}_{\Gamma,\mathbf{m}}$-orbit of a generic element of $\mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)$. □

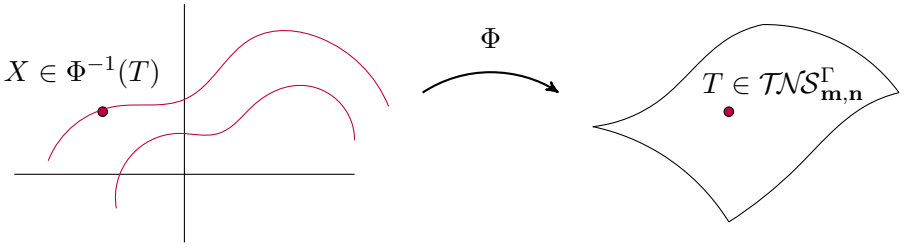


Figure 3.2: Pictorial representation of the tensor network variety map $\Phi : \mathrm{Hom}(W_1, \dots, W_d; V_1 \otimes \cdots \otimes V_d) \to \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$. Given $T \in \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$, the fiber $\Phi^{-1}(T)$ in red contains at least the orbit of $X \in \Phi^{-1}(T)$ under the action of the gauge subgroup $\mathcal{G}_{\Gamma,\mathbf{m}}$.

Applying the Theorem of the Dimension of the Fibers [Sha94, Thm. 1.25] to the $\mathcal{G}_{\Gamma,\mathbf{m}}$-orbit of a generic element $X \in \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)$, we deduce the following corollary, which completes the proof of Theorem 3.0.2.

Corollary 3.3.2. *Let* $(\Gamma, \mathbf{m}, \mathbf{n})$ *be a subcritical tensor network with no overabundant bond dimensions. Then*

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} \leq \left[\sum_{v\in\mathbf{v}(\Gamma)} N_v n_v - d + 1\right] - \sum_{e\in\mathbf{e}(\Gamma)}(m_e^2 - 1) + \dim \operatorname{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$$

where $N_v = \prod_{e\ni v} m_e$ *and* $X = X_1 \otimes \cdots \otimes X_d$ *with* $X_v \in \operatorname{Hom}(W_v, V_v)$ *generic.*

Proof. From (3.1) $\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \dim \operatorname{Hom}(W_1, \ldots, W_d, V_1, \ldots, V_d) - \dim \Phi^{-1}(T)$ where T is a generic element of $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$.

Now $\dim \operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d) = \sum_{v\in\mathbf{v}(\Gamma)} N_v n_v - d + 1$. By Theorem 3.3.1,

$$\begin{aligned}\dim \Phi^{-1}(T) \geq& \dim \mathcal{G}_{\Gamma,\mathbf{m}} \cdot X \\ =& \dim \mathcal{G}_{\Gamma,\mathbf{m}} - \dim \operatorname{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X) \\ =& \sum_{e\in\mathbf{e}(\Gamma)}(m_e^2 - 1) - \dim \operatorname{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X),\end{aligned}$$

where $X \in \operatorname{Hom}(W_1, \ldots, W_d, V_1, \ldots, V_d)$ is generic. □

3.4 Sharpening the upper bound

The term $\dim \operatorname{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ in Theorem 3.0.2 makes the statement not immediate to apply in full generality, as it describes the dimension of the tensor network variety in terms of the dimension of another object which is not immediate to compute. However, as explained in this section, the value $\dim \operatorname{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ can be bounded from above by the dimension of a potentially larger stabilizer which can be computed from the *local structure* of the graph, rather than from its global combinatorics. In fact, a consequence of Proposition 3.4.4 will be that the term $\dim \operatorname{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ is trivial in a wide range of cases.

Definition 3.4.1. Let G be an algebraic group acting on an algebraic variety V. We say that the action is *generically stable* if there exists an element $v \in V$ such that the stabilizer $\operatorname{Stab}_G(v)$ is a finite group.

In particular, the condition that the action of $\mathcal{G}_{\Gamma,\mathbf{m}}$ on $\operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$ is generically stable is equivalent to the fact that the value $\dim \operatorname{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ in Corollary 3.3.2 is zero.

A rich theory has been developed in the study of stable group actions (and more generally semistable actions) starting from [KN79] and related works. We refer to [MFK94] for the theory.

Proposition 3.4.2. *Let* $(C_d, \mathbf{m}, \mathbf{n})$ *be the tensor network on the cycle graph with constant bond dimension* $\mathbf{m} = (m, \ldots, m)$*. Assume* $n_j \geq 2$ *for at least one index. Then the action of* $\mathcal{G}_{C_d,\mathbf{m}}$ *on* $\operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$ *is generically stable, i.e.* $\dim \operatorname{Stab}_{\mathcal{G}_{C_d,\mathbf{m}}}(X) = 0$*, for* $X \in \operatorname{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$ *generic.*

Proof. Let $X_1 \otimes \cdots \otimes X_d \in \mathrm{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$ be a generic element. Write $W_j = U_j \otimes U_{j+1}^*$ with $U_j = U_{j+1} = \mathbb{C}^m$. Then X_j is a generic element of $U_j^* \otimes U_{j+1} \otimes V_j$, with $\dim V_j = n_j \geq 1$. For every j, write $X_j = \sum_{p=1}^{n_j} X_j^{(p)} \otimes v_p$ where $v_1, \ldots, v_{n_j}$ is a basis of V_j and $X_j^{(p)} \in U_j^* \otimes U_{j+1}$.

By genericity $X_j^{(1)}$ is a fixed isomorphism $X_j^{(1)} : U_j \to U_{j+1}$; after choosing bases in U_j, we write $X_j^{(1)} = \mathrm{Id}_m$ in coordinates for $j = 1, \ldots, d-1$ and $X_d^{(1)} : U_d \to U_1$ is a generic diagonal matrix.

The stabilizer $\mathrm{Stab}_{\mathcal{G}_{C_d,\mathbf{m}}}(X)$ is contained in the stabilizer of $X_1^{(1)} \otimes \cdots \otimes X_d^{(1)}$: this is the centralizer of $X_d^{(1)}$; in coordinates this is the maximal torus $\Theta_m \subseteq PGL_m$ of diagonal matrices in PGL_m^{Δ}, where $PGL_m^{\Delta} \subseteq \mathcal{G}_{C_d,\mathbf{m}} = \bigtimes_{j=1}^{d} PGL(U_j)$ lies on the diagonal of the direct factors. Therefore $\mathrm{Stab}_{\mathcal{G}_{C_d,\mathbf{m}}}(X) \subseteq \Theta_m$.

Now, there exists at least one index j such that $n_j \geq 2$. Correspondingly, there is a map $X_j^{(2)} : U_j \to U_{j+1}$. Therefore $\mathrm{Stab}_{\mathcal{G}_{C_d,\mathbf{m}}}(X) \subseteq \mathrm{Stab}_{\Theta_m}(X_j^{(2)})$. By genericity, $X_j^{(2)}$ has full rank and is not diagonal in the fixed basis, hence $\mathrm{Stab}_{\Theta_m}(X_j^{(2)})$ is trivial.

This shows that a generic $X \in \mathrm{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$ satisfies

$$\dim \mathrm{Stab}_{\mathcal{G}_{C_d,\mathbf{m}}}(X) = 0,$$

hence the action of $\mathcal{G}_{C_d,\mathbf{m}}$ on $\mathrm{Hom}(W_1, \ldots, W_d; V_1, \ldots, V_d)$ is generically stable. □

When the stabilizer is finite, $\dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X) = 0$, then the dimension of the generic fiber is bounded by the dimension of the gauge subgroup associated to the tensor network, $\dim \mathcal{G}_{\Gamma,\mathbf{m}} = \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1)$, see Section 2.2. The role of this group in the theory of tensor network is known and it is expected that it entirely controls the value of $\dim \mathcal{TNS}_{\mathbf{m},\mathbf{n}}^{\Gamma}$, c.f. [YL18, LQY12, HMOV14].

In the case of matrix product states associated to graphs *without loops*, the exact value of the dimension is given in [HMOV14] and coincides with our result

$$\min \left\{ \sum_{v \in \mathbf{v}(\Gamma)} (n_v \cdot \textstyle\prod_{e \ni v} m_e) - d + 1 - \displaystyle\sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1), \prod_{v \in \mathbf{v}(\Gamma)} n_v \right\}. \tag{3.3}$$

However, in Section 3.6, we will observe that there are at least some cases where the inequality is strict. They are in particular cases of matrix product states with loops.

For what concerns the stabilizer, we cannot extend the argument of Proposition 3.4.2 to the general case. Instead, we further localize the action, reembedding the gauge subgroup $\mathcal{G}_{\Gamma,\mathbf{m}}$ in the group $H = \mathrm{Im}\ \Psi$, where Ψ is the map described in Section 2.2. This will allow us to use results on the stability of the action on tensor spaces which in turn guarantee the stability of the action of $\mathcal{G}_{\Gamma,\mathbf{m}}$ in a wide range of cases. For this reason, we generally expect that $\dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X) = 0$, at least when the bond dimensions are balanced.

Since $\mathcal{G}_{\Gamma,\mathbf{m}} \subseteq H$, clearly $\mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X) \subseteq \mathrm{Stab}_H(X)$. Therefore, if the action of H on $\mathrm{Hom}(W_1,\dots,W_d,V_1,\dots,V_d)$ is generically stable, then the action of $\mathcal{G}_{\Gamma,\mathbf{m}}$ is generically stable on $\mathrm{Hom}(W_1,\dots,W_d,V_1,\dots,V_d)$.

We establish the following result, whose proof is immediate from the product structure of $\mathrm{Hom}(W_1,\dots,W_d,V_1,\dots,V_d)$ and of X.

Lemma 3.4.3. *If $X = X_1 \otimes \cdots \otimes X_d$ then*

$$\mathrm{Stab}_H(X) = \bigtimes_{v\in\mathbf{v}(\Gamma)} \mathrm{Stab}_{H_v}(X_v).$$

In particular, the action of H on $\mathrm{Hom}(W_1,\dots,W_d,V_1,\dots,V_d)$ is generically stable if and only if for every v the action of H_v on $\mathrm{Hom}(W_v,V_v)$ is generically stable.

Now, regard $X_v \in \mathrm{Hom}(W_v,V_v)$ as a tensor in $V_v \otimes W_v^* = V_v \otimes \left(\bigotimes_{e\ni v} U_e'\right)$ where $U_e' = U_e$ or $U_e' = U_e^*$ depending on whether U_e of U_e^* appears in W_v. The group H_v acts trivially on V_v; by Lemma 1.1.18, we deduce that $\mathrm{Stab}_{H_v}(X_v)$ is the point-wise stabilizer of $\mathrm{Im}\, X_v(V_v^*) \subseteq \bigotimes_{e\ni v} U_e'$. In particular, if X_v is generic, $\mathrm{Stab}_{H_v}(X_v)$ is the simultaneous stabilizer of n_v elements of W_v^*.

Therefore, we are reduced to studying the stability of the action of a product of special linear groups $SL(U_1)\times\cdots\times SL(U_k)$ on the space $U_1\otimes\cdots\otimes U_k\otimes V$. The study of the stability of this action is characterized in the recent [DM21, DMW20] and in the special case where $\dim V = 1$ it is characterized in [BRVR18].

Proposition 3.4.4. *Let $k \geq 3$ and consider vector spaces $U_1,\dots,U_k,V$ with $\dim U_\alpha = m$, $\dim V = n$. The action of $SL(U_1)\times\cdots\times SL(U_k)$ on $U_1\otimes\cdots\otimes U_k\otimes V$ is generically stable unless $(k,m,n) = (3,2,1)$.*

Proof. The case $(k,m,n) = (3,2,1)$ corresponds to the action of $SL_2\times SL_2\times SL_2$ on $\mathbb{C}^2\otimes\mathbb{C}^2\otimes\mathbb{C}^2$; this is not stable since

$$9 = \dim(SL_2\times SL_2\times SL_2) > \dim \mathbb{C}^2\otimes\mathbb{C}^2\otimes\mathbb{C}^2 = 8.$$

Except for this case, the result follows from [DMW20, Theorem 1.5 (Case 4)], since the inequality $m \leq \frac{1}{2}m^{k-1}n$ is always verified. □

Relaxing the hypothesis of Propositions 3.4.2 and 3.4.4, we state the following corollaries of Corollary 3.3.2, for MPS and PEPS respectively.

Corollary 3.4.5. *Let (C_d,m,n) be the tensor network on the cycle graph on d vertices with constant bond dimension m and constant local dimension $n \geq 2$. Then*

$$\dim \mathcal{TNS}_{m,n}^{C_d} \leq \min\{d(n-1)m^2+1, n^d\}.$$

Corollary 3.4.6. *Let Γ be a graph on d vertices such that all vertices of Γ have degree $k \geq 3$. Let (Γ, m, n) be the tensor network on Γ with constant bond dimension m and constant local dimension n. Then, unless $(k, m, n) = (3, 2, 1)$, we have*

$$\dim \mathcal{TNS}^{\Gamma}_{m,n} \leq \min \left\{ \sum_{v \in \mathbf{v}(\Gamma)} n m^{\deg(v)} - d + 1 - \sum_{e \in \mathbf{e}(\Gamma)} (m^2 - 1), n^d \right\}.$$

3.5 Critical case

In this section, we give the exact value of the dimension of the tensor network variety in the critical and supercritical ranges of the parameters. Indeed, in the critical range, the variety is the orbit-closure of the graph tensor under the action of $G(W_1, \ldots, W_d)$. The dimension is therefore controlled by the dimension of the stabilizer of the graph tensor that, by Corollary 2.3.4, coincides with the gauge subgroup. In particular, the dimension of the tensor network variety in the critical range equals the upper bound of Corollary 3.3.2 with $\dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X) = 0$. As a consequence, via Theorem 3.2.2, we obtain equality in the supercritical range.

Proposition 3.5.1. *Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a supercritical tensor network. Write $N_v = \prod_{e \ni v} m_e$. Then*

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \sum_{v \in \mathbf{v}(\Gamma)} n_v N_v - d + 1 - \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1).$$

Proof. First consider the *critical* case, that is $N_v = n_v$. In this case, a generic $X_v \in \mathrm{Hom}(W_v, V_v)$ is invertible. Therefore

$$\begin{aligned} \dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} &= \dim G(W_1, \ldots, W_d) \cdot T(\Gamma, \mathbf{m}) \\ &= \dim G(W_1, \ldots, W_d) - \dim \mathcal{G}_{\Gamma,\mathbf{m}} = \sum_{v \in \mathbf{v}(\Gamma)} N_v^2 - d + 1 - \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1). \end{aligned}$$

In the supercritical case, we apply Theorem 3.2.2. Write $\mathbf{N} = (N_v : v \in \mathbf{v}(\Gamma))$, so that the tensor network $(\Gamma, \mathbf{m}, \mathbf{N})$ is critical. Then

$$\begin{aligned} \dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} &= \dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{N}} + \sum_{v \in \mathbf{v}(\Gamma)} N_v (n_v - N_v) \\ &= \sum_{v \in \mathbf{v}(\Gamma)} N_v^2 - d + 1 - \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1) + \sum_{v \in \mathbf{v}(\Gamma)} N_v (n_v - N_v) \\ &= \sum_{v \in \mathbf{v}(\Gamma)} n_v N_v - d + 1 - \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1). \end{aligned}$$

□

Corollary 3.5.2. *Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a supercritical tensor network. Then*

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = \min \left\{ \sum_{v \in \mathbf{v}(\Gamma)} (n_v \cdot \textstyle\prod_{e \ni v} m_e) - d + 1 - \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1), \prod_{v \in \mathbf{v}(\Gamma)} n_v \right\}.$$

3.6 Analysis of small cases

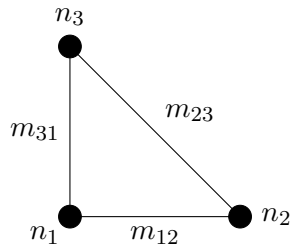
In this section, we analyze some cases of tensor network varieties for small graphs and small bond dimensions. If Γ only contains two vertices with local dimensions $\mathbf{n} = (n_1, n_2)$, then the tensor network variety is easily described as the variety of matrices whose rank is upper bounded by the bond dimension of the unique edge, $\mathbf{m} = m$. The variety is isomorphic to the affine cone over the m-secant variety of the two factor Segre, c.f. 1.1.1, and its projected dimension is bounded by $\dim \sigma_m(\mathbb{P}^{n_1-1} \times \mathbb{P}^{n_2-1}) \leq \min\{m(n_1+n_2-m)-1, n_1n_2-1\}$. This is the well-known case of $n_1 \times n_2$ complex matrices of rank bounded by m and it is completely understood, e.g. [Lan12].

Therefore, we start our analysis with the case of three vertices.

3.6.1 Triangular graph

The graph tensor associated to the triangular graph is the matrix multiplication tensor. This is the object of a rich literature, devoted to determining the value of the exponent of matrix multiplication. We refer to [Blä13, Lan17] for an overview on the subject.

Let C_3 be the triangular graph. Write $\{1,2,3\}$ for the three vertices and m_{12}, m_{23}, m_{31} for the three bond dimensions and (n_1, n_2, n_3) for the three local dimensions, ordered as follows:



If $\mathbf{m} = (m_{12}, m_{23}, m_{31}) = (a, b, 1)$ (in other words, the edge $\{3,1\}$ is erased) then every tensor in $W_1 \otimes W_2 \otimes W_3$ is a restriction of the graph tensor. In particular, if $\mathbf{n} = (n_1, n_2, n_3)$ with $n_1 \leq a$, $n_2 \leq ab$, $n_3 \leq b$, then

$$\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} = V_1 \otimes V_2 \otimes V_3.$$

Therefore, the first interesting case is the one with bond dimensions $\mathbf{m} = (2,2,2)$. We record the cases in the subcritical range in Table 3.1. For each of these cases, the lower bound for the dimension is obtained by computing explicitly the rank of the differential of the parametrization map Φ at a random point. We perform this calculation in Macaulay2 [GS20]. The scripts performing the calculation are available at `https://fulges.github.io/code/BDG-DimensionTNS.html`.

Since the point to compute the differential is chosen at random, we are confident that the number recorded as a lower bound is equal to the actual dimension of the tensor

	$\mathbf{n}$	lower bound	upper bound
	$(2,2,2)$	8	8
	$(2,2,3)$	12	12
	$(2,2,4)$	16	16
	$(2,3,3)$	18	18
*	$(2,3,4)$	22	24
*	$(2,4,4)$	26	29
	$(3,3,3)$	25	25
	$(3,3,4)$	29	29
	$(3,4,4)$	31	31
	$(4,4,4)$	37	37

Table 3.1: Upper and lower bound for $\dim \mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}}$. The lower bound is obtained via a direct calculation. The upper bound is the value obtained in Corollary 3.4.5. In the cases marked with * the two bounds do not coincide.

network variety $\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}}$. However, from a formal point of view, the sole calculation of the rank of the differential at a random point does not provide a complete proof.

The only cases where the lower bound does not match the upper bound given in Corollary 3.4.5 are the ones with $\mathbf{n} = (2,3,4)$ and $\mathbf{n} = (2,4,4)$. In these cases, we prove that the dimension of the tensor network variety equals the lower bound of Table 3.1. We provide the following result, that we prove in general and will be used in Theorem 3.6.2 in the cases $(a,b,r) = (3,4,2)$ and $(a,b,r) = (4,4,2)$.

Lemma 3.6.1. *Let V_1, V_2, V_3 be vector spaces with $\dim V_1 = 2$, $\dim V_2 = a$, $\dim V_3 = b$. Let $\sigma_r \subseteq \mathbb{P}(V_2 \otimes V_3)$ be the variety of elements of rank at most r. Define*

$$\mathcal{Z}_{a,b,r} = \overline{\left\{T \in V_1 \otimes V_2 \otimes V_3 : T(V_1^*) \cap \sigma_r \text{ contains at least two points}\right\}} \subseteq \mathbb{P}(V_1 \otimes V_2 \otimes V_3)$$

Then $\mathcal{Z}_{a,b,r}$ is an irreducible variety and

$$\dim \mathcal{Z}_{a,b,r} = 2r(a+b-r)+1$$

Proof. Define the variety of secant lines

$$\mathcal{S}_{a,b,r} = \overline{\left\{L \in \mathbf{G}(2, V_2 \otimes V_3) : \mathbb{P}L \cap \sigma_r \text{ contains at least two points}\right\}} \subseteq \mathbf{G}(2, V_2 \otimes V_3),$$

where $\mathbf{G}(2, V_2 \otimes V_3)$ denotes the Grassmannian of 2-planes in $V_2 \otimes V_3$.

Then $\mathcal{S}_{a,b,r}$ is an irreducible variety of dimension $2 \dim \sigma_r = 2[r(a+b-r)-1]$ [EH16].

The variety $\mathcal{Z}_{a,b,r}$ is an $SL(V_1)$-bundle on $\mathcal{S}_{a,b,r}$. This guarantees that $\mathcal{Z}_{a,b,r}$ is irreducible and provides

$$\dim \mathcal{Z}_{a,b,r} = \dim \mathcal{S}_{a,b,r} + 3 = 2[r(a+b-r)-1] + 3 = 2r(a+b-r)+1$$

as desired. □

Theorem 3.6.2. *Let* $\mathbf{m} = (2, 2, 2)$.

- *if* $\mathbf{n} = (2, 3, 4)$ *then* $\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} = \widehat{\mathcal{Z}}_{3,4,2}$*; in particular* $\dim \mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} = 22$*;*
- *if* $\mathbf{n} = (2, 4, 4)$ *then* $\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} = \widehat{\mathcal{Z}}_{4,4,2}$*; in particular* $\dim \mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} = 26$.

Proof. The lower bound on the dimension follows from Table 3.1.

By Lemma 3.6.1, we have

$$\dim \mathcal{Z}_{3,4,2} = 4 \cdot (3 + 4 - 2) + 1 = 21,$$
$$\dim \mathcal{Z}_{4,4,2} = 4 \cdot (4 + 4 - 2) + 1 = 25.$$

In the rest of the proof, we show that $\mathcal{TNS}^{C_3}_{\mathbf{m},(2,3,4)} \subseteq \widehat{\mathcal{Z}}_{3,4,2}$ and $\mathcal{TNS}^{C_3}_{\mathbf{m},(2,4,4)} \subseteq \widehat{\mathcal{Z}}_{4,4,2}$; here, if $Y \subseteq \mathbb{P}W$ is a projective variety, $\widehat{Y}$ denotes its affine cone in the space W, c.f. Definition 1.1.4.

Fix generic X_1, X_2, X_3 with $X_j \in \operatorname{Hom}(W_j, V_j)$ and let $T = X_1 \otimes X_2 \otimes X_3(T(C_3, \mathbf{m}))$. Let $L = T(V_1^*) \subseteq V_2 \otimes V_3$. It suffices to show that $\mathbb{P}L \cap \sigma_2$ contains at least two points in the two cases of interest.

We can normalize the linear maps X_1, X_2, X_3 using the gauge subgroup in $GL(W_1, W_2, W_3)$ and the action of $GL(V_1) \times GL(V_2) \times GL(V_3)$ on $V_1 \otimes V_2 \otimes V_3$.

Identify X_1 with a 2×2 matrix $B_1(v_1^{(1)}, v_2^{(1)})$ whose entries are linear combinations of the elements of a basis $\{v_1^{(1)}, v_2^{(2)}\}$ of V_1 and similarly for X_2 and X_3. In this way

$$X_1 \otimes X_2 \otimes X_3(T(C_3, \mathbf{m})) = \operatorname{Tr}\Big(B_1(v_1^{(1)}, v_2^{(1)}) \cdot B_2(v_1^{(2)}, \dots, v_3^{(2)}) \cdot B_3(v_1^{(3)}, \dots, v_4^{(3)})\Big).$$

Write $B_1(v_1^{(1)}, v_2^{(1)}) = B_1^1 v_1^{(1)} + B_1^2 v_2^{(1)}$ and similarly for the other matrices.

By genericity, the map X_3 is invertible: using the action of $GL(V_3)$, we may assume

$$B_3^1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \quad B_3^2 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \quad B_3^3 = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \quad B_3^4 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right).$$

Moreover, the linear space $\langle B_1^1, B_1^2 \rangle$ contains at least one matrix of rank 1; using the action of $GL(V_1)$ and of the gauge group, we may assume $B_1^1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

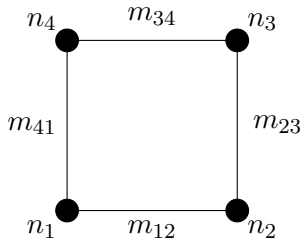
With these normalizations, it is possible to verify that the line $\mathbb{P}(T(V_1^*))$ contains two rank two matrices. We provide a Macaulay2 script determining the intersection $\mathbb{P}(T(V_1^*)) \cap \sigma_2$ at `https://fulges.github.io/code/BDG-DimensionTNS.html`.

If $\mathbf{n} = (2, 3, 4)$, this shows $\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} \subseteq \widehat{\mathcal{Z}}_{3,4,2}$; if $\mathbf{n} = (2, 4, 4)$, this shows $\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} \subseteq \widehat{\mathcal{Z}}_{4,4,2}$.

Finally, since $\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} \subseteq \widehat{\mathcal{Z}}_{3,4,2}$ and they are both irreducible varieties of dimension 22, equality holds. Similarly, equality holds in the inclusion $\mathcal{TNS}^{C_3}_{\mathbf{m},\mathbf{n}} \subseteq \widehat{\mathcal{Z}}_{4,4,2}$. □

3.6.2 Square graph

Consider the square graph C_4 with bond dimensions $\mathbf{m} = (m_{12}, m_{23}, m_{34}, m_{41})$ and local dimensions $\mathbf{n} = (n_1, \dots, n_4)$.



We focus on the case where all bond dimensions are equal to 2. As in the previous section, we record in Table 3.2 the lower bound obtained computing the differential of the parametrization at a random point and the upper bound obtained via Corollary 3.4.5. As before, because of the random choice of point, we are confident that the value recorded as lower bound coincides with the value of $\dim \mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$. We provide a formal proof for the case $\mathbf{n} = (2, 2, 2, 2)$ in Theorem 3.6.3.

Theorem 3.6.3. *Let* $\mathbf{m} = (2, 2, 2, 2)$ *and* $\mathbf{n} = (2, 2, 2, 2)$. *Then*

$$\dim \mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}} = 15;$$

more precisely $\mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$ *is a hypersurface of degree* 6.

Proof. The lower bound $\dim \mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}} \geq 15$ is obtained in Table 3.2.

Since $\dim V_1 \otimes V_2 \otimes V_3 \otimes V_4 = 16$, we obtain that either $\mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$ is the entire space or it is a hypersurface.

We determine an irreducible equation of degree 6 vanishing on $\mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$.

This equation is a degree 6 invariant for the action of $GL(V_1)\times\cdots\times GL(V_4)$ on $V_1\otimes\cdots\otimes V_4$. Its construction is described explicitly in [LT03, HLT12]. The evaluation of the invariant is performed by a Macaulay2 script [GS20] available at `https://fulges.github.io/code/BDG-DimensionTNS.html`.

We illustrate here how to construct it and how to exploit the action of $GL(V_1) \times \cdots \times GL(V_4)$ and of the gauge group to normalize the linear maps and reduce the degrees of freedom in order to allow the script to evaluate the invariant.

Given a tensor $T \in V_1 \otimes V_2 \otimes V_3 \otimes V_4$, consider the bilinear map $T^{1,3} : V_1^* \times V_3^* \to V_2 \otimes V_4$. This makes $V_2 \otimes V_4$ into a space of 2×2 matrices depending bilinearly on $V_1 \times V_3$. Let $F(T) = \det(T^{1,3})$ be the determinant (of the 2×2 matrix $V_2 \otimes V_4$) evaluated on the image of $T^{1,3}$. So $F(T)$ is a polynomial of bidegree $(2, 2)$ in $V_1 \times V_3$, therefore it can be regarded as a bilinear form on $S^2V_1 \times S^2V_2$, where S^2W denotes the second symmetric

	$\mathbf{n}$	lower bound	upper bound
*	$(2,2,2,2)$	15	16
*	$(2,2,2,3)$	20	21
*	$(2,2,2,4)$	24	25
	$(2,2,3,3)$	25	25
	$(2,2,3,4)$	29	29
	$(2,2,4,4)$	33	33
*	$(2,3,2,3)$	24	25
*	$(2,3,2,4)$	28	29
	$(2,3,3,3)$	29	29
	$(2,3,3,4)$	33	33
	$(2,3,4,3)$	33	33
	$(2,3,4,4)$	37	37
*	$(2,4,2,4)$	32	33
	$(2,4,3,4)$	37	37
	$(2,4,4,4)$	41	41
	$(3,3,3,3)$	33	33
	$(3,3,3,4)$	37	37
	$(3,3,4,4)$	41	41
	$(3,4,3,4)$	41	41
	$(3,4,4,4)$	45	45
	$(4,4,4,4)$	49	49

Table 3.2: Upper and lower bound for $\dim \mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$. The lower bound is obtained via a direct calculation. The upper bound is the value obtained in Corollary 3.4.5. In the cases marked with * the two bounds do not coincide.

power of a vector space W. Since $\dim S^2\mathbb{C}^2 = 3$, this bilinear form has an associated 3×3 matrix. The invariant I_6 that we are interested in is the determinant of such matrix, which is a polynomial of degree 6 in the coefficients of the original tensor T.

In order to prove that I_6 vanishes identically on $\mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$, we apply a normalization which reduces the total degrees of freedom, then we perform the calculation symbolically in Macaulay2.

Write $T \in \mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$ as

$$T = \mathrm{Tr}\Big(B_1(v_1^{(1)}, v_2^{(1)}) \cdots B_4(v_1^{(4)}, v_2^{(4)})\Big)$$

where $B_p(v_1^{(p)}, v_2^{(p)}) = B_p^1 v_1^{(p)} + B_p^2 v_2^{(p)}$ are 2×2 matrices depending linearly on a fixed basis of V_p.

Since $\mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$ is invariant under the action of $GL(V_1) \times \cdots \times GL(V_4)$ and the graph tensor is invariant under the action of the gauge subgroup, we may use these groups to normalize the matrices B_p^j. In particular, by the action of $GL(V_1)$ and $GL(V_3)$, we may assume B_1^1 and B_3^1 are rank one matrices; further, using the action of the gauge subgroup, we may assume $B_1^1 = B_3^1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

With this normalization, the evaluation of the invariant is performed and we can verify that $I_6(T) = 0$ whenever $T \in \mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$.

Since I_6 is irreducible, we conclude $\mathcal{TNS}^{C_4}_{\mathbf{m},\mathbf{n}}$ is a hypersurface of degree 6. □

If $d = 5, 6, 7$, the calculation of the differential at a random point shows that in the case of constant bond dimension 2 the dimension of tensor network varieties coincides with the upper bound of Corollary 3.4.5. Therefore, we propose the following conjecture:

Conjecture 3.6.4. *Let $d \geq 3$, $\mathbf{m} = (2, \ldots, 2)$ and $\mathbf{n} = (n_1, \ldots, n_d)$ with $n_j \geq 2$. Then*

$$\dim \mathcal{TNS}^{C_d}_{\mathbf{m},\mathbf{n}} = \min\left\{4\left(\textstyle\sum_{j=1}^d n_j - d\right) + 1, \prod_{j=1}^d n_j\right\}$$

except in the following cases:

- *if $d = 3$: $\mathbf{n} = (2, n_2, n_3)$, with $n_2 \geq 3$, $n_3 \geq 4$ and their cyclic permutations;*
- *if $d = 4$: $\mathbf{n} = (2, n_2, 2, n_4)$ with $n_2, n_4 \geq 2$ and their cyclic permutations.*

The results of this section, together with Theorem 3.2.2, confirm Conjecture 3.6.4 for $d = 3$. As mentioned above, a direct calculation confirms the conjecture for $d = 5, 6, 7$. In the case $d = 4$, the conjecture is confirmed in the case $\mathbf{n} = (2, 2, 2, 2)$, in all cases where the upper and lower bounds coincide in Table 3.2 and in the supercritical cases constructed from those.

Conclusions. In this chapter, we studied the dimension of the tensor network variety. We provided a completely general upper bound in Theorem 3.0.2 and we illustrated how to refine it in the cases relevant for applications in Corollary 3.4.5 and Corollary 3.4.6. In Corollary 3.5.2, we gave the exact value of the dimension of the tensor network variety in the supercritical range of parameters. Finally, we provided the calculation of the dimension of some cases, arising from small values of the parameters, in which the upper bound of Theorem 3.0.2 is not sharp.

Chapter 4

Linear span of uniform matrix product states

In this chapter, we study the linear span of uniform matrix product states. The results contained in the chapter are based on the paper [DLMS22a], co-authored with Harshit J. Motwani and Tim Seynnaeve, that is submitted at the moment of writing. From a quantum mechanics perspective, uniform matrix product model translation invariant physical systems of sites placed on a ring. The geometry of uniform matrix product states has been extensively studied [PGVWC07, HMOV14, CM14, CMS19], but our understanding of them is still far from complete, and several fundamental mathematical problems remain open.

Our geometric point of view is the following: we fix a tensor space $(\mathbb{C}^n)^{\otimes d}$, and consider the set of all tensors in this space that admit a translation-invariant matrix product state representation, with a given bond dimension m. After taking the closure of this set, we obtain an algebraic variety, which we denote by $\mathrm{uMPS}(m,n,d)$. The goal of this chapter is to determine which *linear* equations, if any, vanish on our variety, more precisely:

Problem 4.0.1. *Describe the linear span $\langle \mathrm{uMPS}(m,n,d)\rangle$ of the variety of uniform matrix product states, answering the following questions:*

- *What is the dimension of $\langle \mathrm{uMPS}(m,n,d)\rangle$? (Question 4.1.3)*
- *For which parameters $m,n,d \in \mathbb{N}$ does $\langle \mathrm{uMPS}(m,n,d)\rangle$ fill the ambient space? (Question 4.1.6)*
- *How does $\langle \mathrm{uMPS}(m,n,d)\rangle$ decompose as a GL_n-representation? (Question 4.1.9)*

The variety $\mathrm{uMPS}(m,n,d)$ does not only arise from tensor networks, it is also a very natural geometric construction in its own right. Indeed, as we will soon see, $\mathrm{uMPS}(m,n,d)$ is the closed image of the polynomial map that takes as input an n-tuple of $m \times m$-matrices, and outputs the traces of all d-fold products of the given matrices. In this way,

it is a natural generalization of the classical Veronese variety. Our main Problem 4.0.1 is therefore equivalent to the following:

Problem 4.0.2. *Let $A_0, \ldots, A_{n-1}$ be $m \times m$ matrices with generic entries. Which linear relations hold between the polynomials*

$$\mathrm{Tr}(A_{i_1} \cdots A_{i_d}),$$

where $(i_1, \ldots, i_d) \in [n]^d$?

The ring generated by all polynomials $\mathrm{Tr}(A_{i_1} \cdots A_{i_d})$, where the generic matrices are fixed but we allow d to vary, is known as the *trace algebra*. In his classical work [Pro76], Procesi described how to obtain all relations between the generators of this ring in principle. Note however that the question we are asking is slightly different: we are only interested in relations between traces of matrix products of a *fixed length* d.

The chapter is divided into three sections. In Section 4.1, we define the variety of uniform matrix product states and describe its natural symmetries. In Section 4.2 we undertake a computational study of the space $\langle \mathrm{uMPS}(m, n, d) \rangle$ in the smallest nontrivial case $m = n = 2$. Since the variety $\mathrm{uMPS}(m, n, d)$ is invariant under the action of GL_n, we approach the problem from a representation theory perspective. We recall the basic properties of the representation theory of GL_n and we describe an algorithm that can compute $\langle \mathrm{uMPS}(2, 2, d) \rangle$, viewed as a GL_2-representation. Based on this result, we obtain a conjectured formula for the character (and in particular: the dimension) of $\langle \mathrm{uMPS}(2, 2, d) \rangle$. In Subsection 4.2.3 we describe an algorithm that computes the highest weight vectors of the SL_2-representation $I_k(\mathrm{uMPS}(2, 2, d)$, i.e. the degree k-part of the ideal of the variety. This in particular allows finding explicitly higher degree equations of $\mathrm{uMPS}(2, 2, d)$. Section 4.3 contains our main theoretical results: we introduce a powerful method to find linear equations that vanish on $\langle \mathrm{uMPS}(m, n, d) \rangle$, based on the Cayley-Hamilton theorem. As a corollary, we show that for $d \geq (m + 1)(m + 2)$, the linear span $\langle \mathrm{uMPS}(m, n, d) \rangle$ does not fill its natural ambient space $\mathrm{Cyc}^d(\mathbb{C}^n)$, the space of cyclically symmetric tensors; significantly improving the state of the art. In Subsection 4.3.1 we study the special case $m = n = 2$, based on the computations done in Section 4.3. Using the trace parametrization we show an upper bound on the dimension of $\langle \mathrm{uMPS}(2, 2, d) \rangle$ which is close to optimal, and we take some first steps towards proving our conjectured character formula using again our Cayley-Hamilton technique.

4.1 Uniform matrix product states

We once and for all fix three parameters $m, n, d \in \mathbb{N}$. We will consider tensors in the space $(\mathbb{C}^n)^{\otimes d}$. The standard basis of $\mathbb{C}^n$ will be written as $\{e_0, \ldots, e_{n-1}\}$. We abbreviate the set $\{0, \ldots, n - 1\}$ to $[n]$.

Definition 4.1.1. The uniform matrix product state parametrization is given by the

map

$$\begin{aligned} \phi : (\mathbb{C}^{m\times m})^n &\to (\mathbb{C}^n)^{\otimes d} \\ (A_0, \ldots, A_{n-1}) &\mapsto \sum_{0\le i_1,\ldots,i_d\le n-1} \mathrm{Tr}(A_{i_1}\cdots A_{i_d})\; e_{i_1}\otimes\cdots\otimes e_{i_d}. \end{aligned} \tag{4.1}$$

We denote the image of this map by $\mathrm{uMPS}^\circ(m,n,d)$. The closure of $\mathrm{uMPS}^\circ(m,n,d)$, taken equivalently either in the Euclidean or the Zariski topology over the complex numbers, is the algebraic variety of uniform matrix product states, denoted by $\mathrm{uMPS}(m,n,d)$.

Remark 4.1.2 (Graphical representation)**.** If we think of $(\mathbb{C}^{m\times m})^n$ as the space of $m\times m\times n$ tensors, then the uMPS parametrization takes a tensor in this space and contracts it d times with itself in a circle; see Figure 4.1. Compared to the pictorial representation we give in the previous chapters, c.f. 2.2, from now on we add to every vertex $v\in\mathbf{v}(\Gamma)$ of the graph Γ another edge that reminds us of the local dimension associated to the vertex. This addition is of common use in the physics literature.

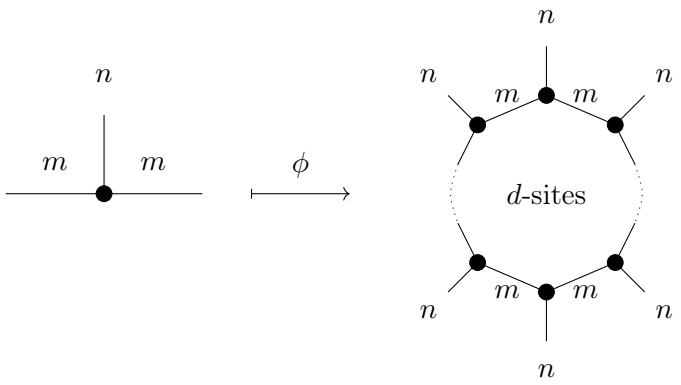


Figure 4.1: Graphical representation of map (4.1) defining $\mathrm{uMPS}(m,n,d)$. There are total d tensors involved with order $m\times m\times n$.

The main question we try to answer in this chapter is the following:

Question 4.1.3. *Determine the linear span of* $\mathrm{uMPS}(m,n,d)$*; i.e. the smallest vector subspace of* $(\mathbb{C}^n)^{\otimes d}$ *containing* $\mathrm{uMPS}(m,n,d)$*. In particular: what is the dimension of this space?*

4.1.1 Cyclic and symmetric invariance

The space $(\mathbb{C}^n)^{\otimes d}$ comes equipped with an action of the symmetric group $\mathfrak{S}_d$: for $\sigma\in\mathfrak{S}_d$ and $\omega = v_1\otimes\cdots\otimes v_d\in(\mathbb{C}^n)^{\otimes d}$ we have

$$\sigma\cdot(v_1\otimes\cdots\otimes v_d) = v_{\sigma^{-1}(1)}\otimes\cdots\otimes v_{\sigma^{-1}(d)}.$$

The symmetric group $\mathfrak{S}_d$ naturally contains the cyclic group C_d and the dihedral group D_{2d} as subgroups. To be precise: let $r,s\in\mathfrak{S}_d$ be the cyclic permutation and reflection

defined respectively by

$$r(i) = i + 1 (\mathrm{mod}\, d) \quad \text{and} \quad s(i) = d + 1 - i,$$

then $C_d \subseteq \mathfrak{S}_d$ is the cyclic subgroup generated by r, and D_{2d} is the subgroup generated by r and s.

Definition 4.1.4. The *cyclically symmetric tensors* and *dihedrally symmetric tensors* are then the elements of $(\mathbb{C}^n)^{\otimes d}$ that are invariant under the action of these subgroups:

$$\mathrm{Cyc}^d(\mathbb{C}^n) := \{\omega \in (\mathbb{C}^n)^{\otimes d} \mid \sigma \cdot \omega = \omega \quad \forall \sigma \in C_d\},$$
$$\mathrm{Dih}^d(\mathbb{C}^n) := \{\omega \in (\mathbb{C}^n)^{\otimes d} \mid \sigma \cdot \omega = \omega \quad \forall \sigma \in D_{2d}\}.$$

Note that both are linear subspaces of $(\mathbb{C}^n)^{\otimes d}$, and that $\mathrm{Dih}^d(\mathbb{C}^n) \subseteq \mathrm{Cyc}^d(\mathbb{C}^n)$, where the inclusion is strict unless $d \leq 2$ or $n = 2$ and $d \leq 5$.

Observation 4.1.5. The set $\mathrm{uMPS}(m, n, d)$ is a subset of the space of cyclically invariant tensors $\mathrm{Cyc}^d(\mathbb{C}^n) \subset (\mathbb{C}^n)^{\otimes d}$ because of the trace invariance under cyclic permutations of the matrices: given $M_1, \ldots, M_d \in \mathbb{C}^{m \times m}$ then

$$\mathrm{Tr}(M_1 \cdots M_d) = \mathrm{Tr}(M_{\sigma(1)} \cdots M_{\sigma(d)}),$$

for $\sigma \in C_d$.

In other words, we can think of $\mathrm{uMPS}(m, n, d)$ as a subvariety of the ambient space $\mathrm{Cyc}^d(\mathbb{C}^n)$. As noted in [CMS19, Corollary 3.18], if we fix n and d and let m grow, the space $\mathrm{uMPS}(m, n, d)$ will eventually fill the ambient space $\mathrm{Cyc}^d(\mathbb{C}^n)$.

Question 4.1.6. *For fixed n and d, what is the smallest m such that*

$$\langle \mathrm{uMPS}(m, n, d) \rangle = \mathrm{Cyc}^d(\mathbb{C}^m).$$

It is known that for $m = d$, equality holds [CMS19, Proposition 3.11]. On the other hand, it follows from a dimension count ([CMS19, Theorem 3.14], see also [NV18]) that if $d \gg m$, the inclusion $\langle \mathrm{uMPS}(m, n, d) \rangle \subset \mathrm{Cyc}^d(\mathbb{C}^n)$ is strict. In Section 4.3 we will prove that already for $d = O(m^2)$, we have a *strict* inclusion.

Observation 4.1.7. In the case $m = n = 2$, we have the stronger inclusion

$$\mathrm{uMPS}(2, 2, d) \subseteq \mathrm{Dih}^d(\mathbb{C}^n).$$

This is a consequence of the identity $\mathrm{Tr}(A_{i_1} \cdots A_{i_d}) = \mathrm{Tr}(A_{i_d} \cdots A_{i_1})$, which holds for any pair of 2×2 matrices A_0, A_1 and sequence $i_1, \ldots, i_d$ with $i_j \in \{0, 1\}$. See [Gre14, Theorem 1.1]

Example 4.1.8. Consider $\mathrm{uMPS}(2, 2, 6)$. For every $A_0, A_1 \in \mathbb{C}^{2 \times 2}$, it holds the trace relation $\mathrm{Tr}(A_0^2 A_1^2 A_0 A_1) = \mathrm{Tr}(A_1 A_0 A_1^2 A_0^2)$, that does not come from the invariance of trace under cyclic permutations of the matrices.

4.1.2 GL_n-invariance

The general linear group GL_n naturally acts on the space $(\mathbb{C}^n)^{\otimes d}$: given $g \in GL_n$ and $\omega = v_1 \otimes \cdots \otimes v_d \in (\mathbb{C}^n)^{\otimes d}$, we have

$$g \cdot (v_1 \otimes \cdots \otimes v_N) = (g \cdot v_1) \otimes \cdots \otimes (g \cdot v_N).$$

Clearly, $\mathrm{Cyc}^d(\mathbb{C}^n)$ and $\mathrm{Dih}^d(\mathbb{C}^n)$ are invariant under this action. The following computation shows that $\mathrm{uMPS}(m, n, d)$ is invariant under this action as well:

$$\begin{aligned}
g \cdot \phi(A_1, \ldots, A_n) &= \sum_{j_1,\ldots,j_d=1}^{n} \mathrm{Tr}(A_{j_1} \cdots A_{j_d}) \, (\sum_{i_1=1}^{n} g_{i_1,j_1} e_{i_1}) \otimes \cdots \otimes (\sum_{i_d=1}^{n} g_{i_d,j_d} e_{i_d}) \\
&= \sum_{j_1,\ldots,j_d=1}^{n} \sum_{i_1,\ldots,i_d=1}^{n} g_{i_1,j_1} \cdots g_{i_d,j_d} \, \mathrm{Tr}(A_{j_1} \cdots A_{j_d}) \; e_{i_1} \otimes \cdots \otimes e_{i_d} \\
&= \sum_{j_1,\ldots,j_d=1}^{n} \sum_{i_1,\ldots,i_d=1}^{n} \mathrm{Tr}(g_{i_1,j_1} A_{j_1} \cdots g_{i_d,j_d} A_{j_d}) \; e_{i_1} \otimes \cdots \otimes e_{i_d} \\
&= \sum_{i_1,\ldots,i_d=1}^{n} \mathrm{Tr}\Big[\Big(\sum_{j_1=1}^{n} g_{i_1,j_1} A_{j_1}\Big) \cdots \Big(\sum_{j_d=1}^{n} g_{i_d,j_d} A_{j_d}\Big)\Big] e_{i_1} \otimes \cdots \otimes e_{i_d} \\
&= \phi\Big(\sum_{j=1}^{n} g_{1,j} A_j, \ldots, \sum_{j=1}^{n} g_{n,j} A_j\Big). \qquad (4.2)
\end{aligned}$$

This means that the space $\langle \mathrm{uMPS}(m, n, d) \rangle$ we are interested in is naturally a representation of GL_n. We briefly recall what we will use about the representation theory of GL_n, and we fix the notation.

We once and for all fix a torus $T \subset GL_n$, consisting of all diagonal matrices; and identify T with $(\mathbb{C}^*)^n$. For $\lambda = (\lambda_0, \ldots, \lambda_{n-1}) \in \mathbb{Z}^n$ and $t = \mathrm{diag}(t_0, \ldots, t_{n-1}) \in T$, we write $t^\lambda = t_0^{\lambda_0} \cdots t_{n-1}^{\lambda_{n-1}} \in \mathbb{C}^*$. Let V be any representation of GL_n. For any $\lambda \in \mathbb{Z}^n$, the *weight space* V_λ is defined as

$$V_\lambda = \{v \in V \mid t \cdot v = t^\lambda v \quad \forall t \in T\}.$$

It is a well-known fact from representation theory that

$$V = \bigoplus_{\lambda \in \mathbb{Z}^n} V_\lambda$$

as vector spaces; and that knowing the dimensions of the weight spaces uniquely determines the representation V up to isomorphism. The polynomial

$$\chi_V = \sum_{\lambda \in \mathbb{Z}^n} (\dim V_\lambda) t^\lambda.$$

is known as the *character* of V. If we view our representation as a morphism $\rho : GL_n \to GL(V)$, the character χ_V is equal to $\mathrm{Tr}(\rho(\mathrm{diag}(t_0, \ldots, t_{n-1})))$.

So we can refine our main Question 4.1.3 to:

Question 4.1.9. *Let $V = \langle \mathrm{uMPS}(m,n,d) \rangle$, viewed as a GL_n-representation. For every weight $\lambda \in \mathbb{Z}^n$, determine the dimension of the weight space V_λ.*

4.1.3 Words, necklaces and bracelets

As a warm-up, let us consider the classical representation $V = (\mathbb{C}^n)^{\otimes d}$. Its character is equal to $(t_1 + \cdots + t_n)^d$, and by expanding we find that $\dim V_\lambda$ is equal to the multinomial coefficient $\binom{d}{\lambda_1,\ldots,\lambda_n}$ if $\sum \lambda_i = d$, and zero otherwise. We can also see this in terms of coordinates, and this will be useful later:

Definition 4.1.10. A *word* of length d on the alphabet $[n]$ is just an ordered tuple $I = (i_1, \ldots, i_d)$, with $i_j \in [n]$. The *weight* of a word I is a tuple $w(I) = (w_0, \ldots, w_{n-1}) \in \mathbb{Z}^n$, where w_i is the number of entries in I that are equal to i. We denote by $W(n,d)$ the set of words of length d on the alphabet $[n]$.

For every word $I = (i_1, \ldots, i_d)$, we can define a vector $e_I := e_{i_1} \otimes \cdots \otimes e_{i_d}$. The space $(\mathbb{C}^n)^{\otimes d}$ has a basis given by the e_I, where I runs over all words of length d on the alphabet $[n]$. In addition, every e_I is a weight vector of weight $w(I)$. So the dimension of the weight space V_λ is the number of words of weight λ, which is indeed the multinomial coefficient $\binom{d}{\lambda_1,\ldots,\lambda_n}$.

We move on to the subrepresentations $\mathrm{Cyc}^d(\mathbb{C}^n)$ and $\mathrm{Dih}^d(\mathbb{C}^n)$, the natural ambient spaces for uniform matrix product states.

Definition 4.1.11. A *necklace* (of length d on the alphabet $[n]$) is an equivalence class of words, where two words are equivalent if they agree up to the action C_d. A *bracelet* (of length d on the alphabet $[n]$) is an equivalence class of words, where two words are equivalent if they agree up to the action D_{2d}. For a fixed necklace or bracelet, all words in the equivalence class clearly have the same weight; this is the *weight* of the necklace or bracelet. We denote by $N(n,d)$, resp. $B(n,d)$, the set of necklaces, resp. bracelets, of length d on $[n]$ and by $N_\lambda(n,d) \subset N(n,d)$, resp. $B_\lambda(n,d) \subset B(n,d)$, the subset of elements of weight $\lambda \in \mathbb{Z}^n$.

To every necklace $N \in N(n,d)$, we associate a basis vector $e_N := \frac{1}{d} \sum_{\sigma \in C_d} \sigma \cdot e_I$, where I is any representative of N. Then $\mathrm{Cyc}^d(\mathbb{C}^n)$ has a basis given by $\{e_N : N \in N(n,d)\}$. Moreover, e_N is a weight vector of weight $w(N)$, hence we find that the dimension of the weight space of weight λ is given by the number $|N_\lambda(n,d)|$ of necklaces of weight λ.

Remark 4.1.12. The number $|N(n,d)|$ of necklaces of length d on $[n]$ can be counted using Polya's enumeration theorem, see for instance [Sta13, Theorem 7.10]:

$$\dim \mathrm{Cyc}^d(\mathbb{C}^n) = |N(n,d)| = \frac{1}{d} \sum_{l|d} \phi(l) n^{\frac{d}{l}},$$

where φ is Euler's totient function.

There is also a formula for $|N_\lambda(n,d)| = \dim \mathrm{Cyc}^d(\mathbb{C}^n)_\lambda$: it is equal to the coefficient of $x_0^{\lambda_0}\cdots x_{n-1}^{\lambda_{n-1}}$ in the polynomial

$$\frac{1}{d}\sum_{\ell|d}(x_0^{d/\ell}+\cdots+x_{n-1}^{d/\ell})^\ell\varphi\left(\frac{d}{\ell}\right).$$

To every bracelet $b \in B(n,d)$, we associate a basis vector $e_b := \frac{1}{2d}\sum_{\sigma\in D_{2d}}\sigma\cdot e_I$, where I is a representative of b. Then $\mathrm{Dih}^d(\mathbb{C}^n)$ has a basis given by $\{e_b : b \in B(n,d)\}$, and the dimension of the weight space of weight λ is given by the number $|B_\lambda(n,d)|$ of bracelets of weight λ.

Remark 4.1.13. The number of bracelets of length d on $[n]$ is given by

$$\dim \mathrm{Dih}^d(\mathbb{C}^n) = |B(n,d)| = \begin{cases} \frac{1}{2}|N(n,d)| + \frac{1}{4}(n+1)n^{d/2} & \text{for } d \text{ even,} \\ \frac{1}{2}|N(n,d)| + \frac{1}{4}n^{(d+1)/2} & \text{for } d \text{ odd.} \end{cases}$$

We only state the formula for $|B_\lambda(n,d)|$ in the case of binary bracelets (i.e. $n=2$), as that is the only case that is relevant to us:

$$|B_{(w,d-w)}(2,d)| = \begin{cases} \left(\frac{1}{2d}\sum_{l|\gcd(d,w)}\varphi(l)\binom{\frac{d}{l}}{\frac{w}{l}}\right) + \frac{1}{2}\binom{\frac{d}{2}-1}{\frac{w-1}{2}} & \text{for } w \text{ odd,} \\ \left(\frac{1}{2d}\sum_{l|\gcd(d,w)}\varphi(l)\binom{\frac{d}{l}}{\frac{w}{l}}\right) + \frac{1}{2}\binom{\frac{d}{2}}{\frac{w}{2}} & \text{for } w \text{ even.} \end{cases}$$

4.2 Computations

In this section, we describe how to computationally answer Question 4.1.9 for fixed parameters. We focus on the smallest interesting case $m = n = 2$. In this case we are dealing with representations of GL_2, so the weights are in $\mathbb{Z}^2$. Moreover, the only occurring weights in $(\mathbb{C}^2)^{\otimes d}$ are $t_0^w t_1^{d-w}$ for $w = 0,\ldots,d$. For subrepresentations of $(\mathbb{C}^2)^{\otimes d}$, we will abbreviate the weight spaces $V_{(w,d-w)}$ to V_w. Our goal is to determine the dimension of the weight spaces $\langle \mathrm{uMPS}(2,2,d)\rangle_w$.

All of our dimension counts in this section and the next use the following easy observation.

Observation 4.2.1. For $p_1,\ldots,p_N \in \mathbb{C}[y_1,\ldots,y_s]$ polynomials, and X the image of the polynomial map

$$\begin{aligned} \mathbb{C}^s &\to \mathbb{C}^N \\ (y_1,\ldots,y_s) &\mapsto (p_1(y_1,\ldots,y_s),\ldots,p_N(y_1,\ldots,y_s)), \end{aligned}$$

a linear equation $\sum \alpha_i x_i$ vanishes on X if and only if the identity $\sum \alpha_i p_i = 0$ holds in the polynomial ring $\mathbb{C}[y_1,\ldots,y_s]$. In particular, the dimension of the linear span of X is equal to the dimension of the subspace of $\mathbb{C}[y_1,\ldots,y_s]$ spanned by the p_i's.

4.2.1 The trace parametrization

If we directly use Definition 4.1.1, we see that uMPS$(2,2,d)$ is the closed image of a polynomial map $\mathbb{C}^8 \to \mathrm{Dih}^d(\mathbb{C}^2)$. However, in this specific case there is an alternative parametrization by $\mathbb{C}^5$ instead of $\mathbb{C}^8$: the *trace parametrization*, which appears to be computationally more efficient in practice. It is based on the connection between uniform matrix product states and invariant theory of matrices.

Definition 4.2.2. Let

$$R = \mathbb{C}[(a_{ij}^k)_{1\leq i,j\leq m;1\leq k\leq n}]$$

be the polynomial ring in $m^2 n$ variables, and for $k = 0,\ldots,n-1$, let $A_k := (a_{ij}^k)_{1\leq i,j\leq m}$ be generic $m \times m$ matrices. The *trace algebra* $\mathcal{C}_{m,n}$ is the subalgebra of R generated by the polynomials $\mathrm{Tr}(A_{i_1}\cdots A_{i_s})$, where $(i_1,\ldots,i_s)$ runs over all words (or equivalently: necklaces) in $[n]$.

Remark 4.2.3. The trace algebra is precisely the subring of R consisting of all elements that are invariant under simultaneous conjugation:

$$f \in \mathcal{C}_{m,n} \iff f(P^{-1}A_0P,\ldots,P^{-1}A_{n-1}P) = f(A_0,\ldots,A_{n-1}) \quad \forall P \in GL_m.$$

This is known as the *first fundamental theorem* in the invariant theory of matrices [Sib68, Pro76].

Proposition 4.2.4 ([Sib68, Corollary 2])**.** *The trace algebra $\mathcal{C}_{2,2}$ is generated by the following five polynomials:*

$$\mathrm{Tr}(A_0), \mathrm{Tr}(A_1), \mathrm{Tr}(A_0^2), \mathrm{Tr}(A_0A_1), \mathrm{Tr}(A_1^2),$$

and moreover, there are no polynomial relations between these generators.

Corollary 4.2.5. *For every bracelet $b = (b_1,\ldots,b_k)$, there is a unique polynomial*

$$P_b(T_0,T_1,T_{00},T_{01},T_{11}) \in \mathbb{C}[T_0,T_1,T_{00},T_{01},T_{11}]$$

such that for every pair (A_0,A_1) of 2×2 matrices, the following equality holds:

$$\mathrm{Tr}(A_{b_1}\cdots A_{b_k}) = P_b(\mathrm{Tr}(A_0),\mathrm{Tr}(A_1),\mathrm{Tr}(A_0^2),\mathrm{Tr}(A_0A_1),\mathrm{Tr}(A_1^2)).$$

Remark 4.2.6. If we give the ring $\mathbb{C}[T_0,T_1,T_{00},T_{01},T_{11}]$ a grading by putting $\deg(T_0) = \deg(T_1) = 1$ and $\deg(T_{00}) = \deg(T_{01}) = \deg(T_{11}) = 2$, then the polynomial P_b is homogeneous of degree $\mathrm{length}(b)$.

The above means that uMPS$(2,2,d)$ is the image of the polynomial map

$$\begin{aligned} \widetilde{\phi} : \mathbb{C}^5 &\to \mathrm{Dih}^d(\mathbb{C}^2) \\ (T_0,T_1,T_{00},T_{01},T_{11}) &\mapsto \sum_b P_b(T_0,T_1,T_{00},T_{01},T_{11})e_b, \end{aligned} \tag{4.3}$$

where b runs over all bracelets of length d. This is the trace parametrization.

In order to compute the polynomials P_b, for bracelets of length 3, one verifies that

$$P_{000} = -\frac{1}{2}T_0^3 + \frac{3}{2}T_0T_{00}, \qquad P_{100} = -\frac{1}{2}T_0^2T_1 + \frac{1}{2}T_1T_{00} + T_0T_{01},$$
$$P_{110} = -\frac{1}{2}T_0T_1^2 + \frac{1}{2}T_0T_{11} + T_1T_{01}, \qquad P_{111} = -\frac{1}{2}T_1^3 + \frac{3}{2}T_1T_{11}.$$

For bracelets of length ≥ 4, we can inductively use the following identity [Sib68], which holds for every quadruple (A, B, C, D) of 2×2-matrices:

$$\begin{aligned} 2\operatorname{Tr}(ABCD) = &\operatorname{Tr}(A)\left(\operatorname{Tr}(BCD) - \operatorname{Tr}(B)\operatorname{Tr}(CD)\right) + \operatorname{Tr}(B)\left(\operatorname{Tr}(CDA) - \operatorname{Tr}(C)\operatorname{Tr}(DA)\right) \\ &+ \operatorname{Tr}(C)\left(\operatorname{Tr}(DAB) - \operatorname{Tr}(D)\operatorname{Tr}(AB)\right) + \operatorname{Tr}(D)\left(\operatorname{Tr}(ABC) - \operatorname{Tr}(A)\operatorname{Tr}(BC)\right) \\ &- \operatorname{Tr}(AC)\operatorname{Tr}(BD) + \operatorname{Tr}(AB)\operatorname{Tr}(CD) + \operatorname{Tr}(AD)\operatorname{Tr}(BC) \\ &+ \operatorname{Tr}(A)\operatorname{Tr}(B)\operatorname{Tr}(C)\operatorname{Tr}(D). \end{aligned}$$

4.2.2 Computing the character

The weight space $\langle \text{uMPS}(2,2,d)\rangle_w$ is the image of the map

$$\begin{aligned} \mathbb{C}^5 &\to \operatorname{Dih}^d(\mathbb{C}^2)_w \\ (T_0, T_1, T_{00}, T_{01}, T_{11}) &\mapsto \sum_b P_b(T_0, T_1, T_{00}, T_{01}, T_{11})e_b, \end{aligned}$$

where b ranges over all bracelets of weight w. By Observation 4.2.1, we need to compute the dimension of the linear subspace of $\mathbb{C}[T_0, T_1, T_{00}, T_{01}, T_{11}]$ spanned by the P_b's. This can be done computing the rank of the matrix of coefficients of P_b's.

Putting everything together, we obtain Algorithm 1, which for a given d computes the character (in particular the dimension) of $\langle \text{uMPS}(2,2,d)\rangle$.

The results are summarized in Table 4.1, where we write $D_w := \dim\langle \text{uMPS}(2,2,d)\rangle_w$. For $d < 8$, the space $\langle \text{uMPS}(2,2,d)\rangle$ is equal to the ambient space $\operatorname{Dih}^2(\mathbb{C}^n)$.

We implemented this algorithm in SageMath [The20]. The code is available at `https://github.com/harshitmotwani2015/uMPS/`.

Algorithm 1: Linear span of $\langle \mathrm{uMPS}(2,2,d)\rangle$

Input: d

Output: Character of the representation $\langle \mathrm{uMPS}(2,2,d)\rangle$

$T_0 \leftarrow \mathrm{Tr}(A_0)$; $T_1 \leftarrow \mathrm{Tr}(A_1)$; $T_{00} \leftarrow \mathrm{Tr}(A_0^2)$; $T_{01} \leftarrow \mathrm{Tr}(A_0A_1)$; $T_{11} \leftarrow \mathrm{Tr}(A_1^2)$;

for $\ell = 3$ **to** d **do**

  bracelets = bracelets of length ℓ;

  for b *in bracelets* **do**

    P[b] $\leftarrow \mathrm{Tr}(A_{b_1}\cdots A_{b_\ell})$, expressed in the T_i;

  end

end

for $w = 0$ **to** d **do**

  List the monomials $\mathbf{y}^{\alpha_1},\ldots,\mathbf{y}^{\alpha_t}$ appearing in the $P[b]$, where b ranges over all bracelets of weight w;

  Write $P[b] = \sum_j \beta_{b,j}\mathbf{y}^{\alpha_j}$;

  Compute the rank of the matrix $(\beta_{b,j})_{b,j}$;

end

d	D_0	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	$\dim\langle \mathrm{uMPS}(2,2,d)\rangle$	$\dim \mathrm{Dih}^2(\mathbb{C}^n)$
8	1	1	4	5	7							29	30
9	1	1	4	6	8							40	46
10	1	1	5	7	11	11						61	78
11	1	1	5	8	12	14						82	126
12	1	1	6	9	15	17	20					118	224
13	1	1	6	10	16	20	23					154	380
14	1	1	7	11	19	23	29	29				211	687
15	1	1	7	12	20	26	32	35				268	1224
16	1	1	8	13	23	29	38	41	45			353	2250
17	1	1	8	14	24	32	41	47	51			438	4112
18	1	1	9	15	27	35	47	53	61	61		559	7685
19	1	1	9	16	28	38	50	59	67	71		680	14310
20	1	1	10	17	31	41	56	65	77	81	86	846	27012

Table 4.1: Character of $\langle \mathrm{uMPS}(2,2,d)\rangle$. Since $D_w = D_{d-w}$[1], we only list D_w for $w \leq \lceil \frac{d}{2} \rceil$.

[1]More generally: for every GL_2-representation V we have that $V_{(b_0,b_1)} = V_{(b_1,b_0)}$.

Based on these computations, we make the following conjecture:

Conjecture 4.2.7.

$$\dim\langle \mathrm{uMPS}(2,2,d)\rangle_w = \begin{cases} 1+\frac{d(v-1)v}{2}-\frac{2(v-1)v(2v-1)}{3}+v\lfloor\frac{d}{2}\rfloor-2v^2+v & \text{for } w=2v, \\ 1+\frac{dv(v+1)}{2}-\frac{2v(v+1)(2v+1)}{3} & \text{for } w=2v+1. \end{cases}$$

This would in particular imply

Conjecture 4.2.8 (Corollary of Conjecture 4.2.7)**.**

$$\dim\langle \mathrm{uMPS}(2,2,d)\rangle = \begin{cases} \frac{1}{192}(d^4-4d^2+192d+192) & \text{for } d \text{ even}, \\ \frac{1}{192}(d^4-10d^2+192d+201) & \text{for } d \text{ odd}. \end{cases}$$

4.2.3 Higher order equations of $\mathrm{uMPS}(2,2,d)$

Equations of tensor varieties, such as Secant varieties of the Segre variety or Veronese variety, are hard to compute. For uniform matrix product states, Critch and Morton gave a complete description of the ideal of the varieties $\mathrm{uMPS}(2,2,4)$ and $\mathrm{uMPS}(2,2,5)$ and, in [CM14], several linear equations of $\mathrm{uMPS}(2,2,d)$ are given for d until 12. The generators of the ideal of $\mathrm{uMPS}(2,2,d)$ for $d=4,5,6$ are given in [CMS19]. Our method does not find equations for the variety via an implicitization problem as in the previously cited work. Such problems are difficult in general. We instead exploit representation theory, computing directly the highest weight vectors of the SL_2-representation $I_k(\mathrm{uMPS}(2,2,d))$, the degree k part of the ideal of the variety. Our method turns out to be based on the solution of a linear system of equations in unknown coefficients, i.e. a linear algebra problem.

Defining equations. Let V be a complex vector space of dimension S and V^* its dual space. Let $\{v_1,\dots,v_S\}$ and $\{x_1,\dots,x_S\}$ be the respective canonical bases. The ring of polynomial functions of V over $\mathbb{C}$ is denoted by $\mathbb{C}[V]\simeq\mathbb{C}[x_1,\dots,x_S]$. It is isomorphic to the symmetric algebra of V^*; $\mathbb{C}[V]\simeq\mathrm{Sym}(V^*)=\bigoplus_{k\geq 0}\mathrm{Sym}^k(V^*)$. Sometimes we will denote $\mathrm{Sym}^k(V^*)=\mathbb{C}[V]_k$. Let $X\subseteq V$ be an algebraic variety. Its defining ideal decomposes by degrees

$$I(X)=\{f\in\mathbb{C}[V]: f(p)=0,\forall p\in X\}=\bigoplus_{k>0} I_k(X),$$

where $I_k(X)$ is the vector space of homogeneous polynomials of degree k vanishing on X.

The variables of the polynomial ring $\mathbb{C}[\mathrm{Cyc}^d(\mathbb{C}^n)]$ can be written as x_N, with $N\in N(n,d)$. In the particular case $m=n=2$, the variables of $\mathbb{C}[\mathrm{Dih}^d(\mathbb{C}^2)]$ are x_B, with $B\in B(2,d)$. The variety $\mathrm{uMPS}(2,2,d)$ is invariant under the action of the general linear group given

in Equation (4.2). The action naturally induces an action on the ring of polynomial functions $\mathbb{C}[\mathrm{Dih}^d(\mathbb{C}^2)] \simeq \mathrm{Sym}(\mathrm{Dih}^d(\mathbb{C}^2)^*)$ and on the ideal of the variety

$$I(\mathrm{uMPS}(m,n,d)) = \bigoplus_{k>0} I_k(\mathrm{uMPS}(2,2,d)).$$

We therefore consider the space $I_k(\mathrm{uMPS}(2,2,d)) \subset \mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*)$ as a representation of SL_2 and we implement an algorithm that explicitly computes its *highest weight vectors* in the ranges $k = 1$ and $d \leq 16$, $k = 2$ and $d \leq 9$, $k = 3$ and $d \leq 8$. The highest weight vectors are in particular equations of the variety. On the other hand, the collection of highest weight vectors allows to determine the SL_2-decomposition, degree by degree, of the ideal into irreducible subrepresentations. The decomposition of the *degree one* part $I_1(\mathrm{uMPS}(2,2,d))$, seen as a vector space isomorphic to the dual of the linear span of the variety, is coherent with the computations done in Section 4.2.

Weight vectors. Let $s = \mathrm{diag}(s_0, s_1) \in T \subseteq GL_2$ be an element of the torus. Let $I = (i_1, \ldots, i_d) \in W(2,d)$ be a word of length d in the alphabet $[2]$. The action of the torus on the tensor product $V = (\mathbb{C}^2)^{\otimes d}$ is given by

$$s(e_I) = s(e_{i_1}) \otimes \cdots \otimes s(e_{i_d}) = s_0^{d-w} s_1^w e_I,$$

where w is the number of entries in I that are equal to 1. The element $e_I = e_{i_1} \otimes \cdots \otimes e_{i_d}$ is therefore a weight vector of weight $w(I) = (d-w, w)$, $w \in \{0, \ldots, d\}$.

Define $t_0 := s_0^{-1}$, $t_1 := s_1^{-1}$ and $t = \mathrm{diag}(t_0, t_1) \in T \subseteq GL_2$. The action of the torus on $x_I \in V^*$ is therefore given by

$$s(x_I) = (s_0^{-1})^{d-w}(s_1^{-1})^w x_I = t_0^{d-w} t_1^w x_{i_1,\ldots,i_d}. \tag{4.4}$$

We want to determine the action on the k-symmetric power of V^*. An element of $\mathbb{C}[V]_k$ is a monomial $\underline{x} = x_{I_1} \ldots x_{I_k}$, with $I_j = (i_1^j, \ldots, i_d^j) \in W(2,d)$, for $j = 1, \ldots, k$. Let $w_j := w(I_j) \in \mathbb{Z}^2$ be the weight of I_j. By Equation (4.4) the action on the monomial is

$$s(\underline{x}) = \prod_{j=1}^k s(x_{I^j}) = \prod_{j=1}^k s(x_{i_1^j,\ldots,i_d^j}) = \prod_{j=1}^k t_0^{d-w_j} t_1^{w_j} \underline{x} = \left(t_0^{d-\sum_{j=1}^k w_j} t_1^{\sum_{j=1}^k w_j} \right) \underline{x} \tag{4.5}$$

Define $\omega := \sum_{j=1}^k w_j$. Clearly $\omega \in \{0, \ldots, kd\}$ and it is the number of 1's appearing in $\underline{x} \in \mathbb{C}[V]_k$. We have that every monomial $\underline{x} \in \mathbb{C}[V]_k$ is a weight vector of weight $t_0^{dk-\omega} t_1^\omega$, for some $\omega \in \{0, \ldots, dk\}$. We abbreviate again "weight $\lambda = (dk - \omega, \omega) \in \mathbb{Z}^2$" to "weight ω". The action naturally restricts to $\mathbb{C}[\mathrm{Dih}^d(\mathbb{C}^2)]_k \subseteq \mathbb{C}[V]_k$.

Highest weight vectors. We summarize the facts we need here about the representation theory of $\mathfrak{sl}_2$, the Lie algebra of SL_2, c.f. Chapter 1, Subsection 1.1.2.

The Lie algebra of SL_2, denoted by $\mathfrak{sl}_2$, is the space of 2×2 trace-zero matrices with complex entries. It is $\mathbb{C}$-span by three operators

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

which satisfy the commutation relations $[H, X] = 2X$,$[H, Y] = -2Y$ and $[X, Y] = H$. For every non-negative integer n there is an *irreducible* SL_2-representation V_n of dimension $(n + 1)$; isomorphic to $\mathrm{Sym}^n(\mathbb{C}^2)$. If $v \in V(n) := \{v \in \mathbb{C}^2 : H(w) = n \cdot v\}$ then $X(v) = 0$ and $\{v, Y(v), \ldots, Y^n(v)\}$ spans V_n. A vector $v \in V_n$, eigenvector of H, such that $X(v) = 0$ is called *highest weight vector* and its corresponding eigenvalue is called highest weight.

The operator X, called the *raising operator*, acts on the elements of the basis as $X(e_0) = 0, X(e_1) = e_0$ and the induced action on the tensor product is given by the Leibniz's rule

$$X(e_I) = X(e_{i_1}) \otimes e_{i_2} \otimes \cdots \otimes e_{i_d} + \cdots + e_{i_1} \otimes \cdots \otimes e_{i_{d-1}} \otimes X(e_{i_d}) = \sum_{j_1, \ldots, j_d} e_I,$$

where $(j_1, \ldots, j_d)$ are all the possible words such that 1 is replaced by 0.

The action on the coordinate $x_I \in V^*$ is

$$x_I \mapsto - \sum_{j_1, \ldots, j_d} x_{j_1, \ldots, j_d}, \tag{4.6}$$

where $(j_1, \ldots, j_d)$ are all the possible words such that 1 is replaced by 0.

The opposite holds for the so called *lowering operator* Y, which acts on the elements of the basis as $Y(e_0) = e_1, Y(e_1) = 0$, where $(j_1, \ldots, j_d)$ are all the possible words such that 0 is replaced by 1.

Given the monomial $\underline{x} = x_{I_1} \ldots x_{I_k} \in \mathbb{C}[V]_k$, with $I_j = (i_1^j, \ldots, i_d^j) \in W(2, d)$, for $j = 1, \ldots, k$, by Equation (4.6) we have

$$X(\underline{x}) = \sum_{1 \leq j \leq k} x_{I_1} \cdots X(x_{I_j}) \cdots x_{I_k}. \tag{4.7}$$

The action naturally restricts to $\mathbb{C}[\mathrm{Dih}^d(\mathbb{C}^2)]_k$.

We have that the highest weight vectors for the SL_2-representation $\mathbb{C}[\mathrm{Dih}^d(\mathbb{C}^2)]_k \simeq \mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*)$ are the polynomials in the variables of $\mathrm{Dih}^d(\mathbb{C}^2)^*$ which are linear combinations of monomials of degree k of the same weight $\omega \in \{0, \ldots, \lceil \frac{dk}{2} \rceil\}$, described in Equation (4.5), and which are killed by (the induced action of) the raising operator X as in Equation (4.7). The number of highest weight vectors of a fixed weight

$\omega \in \{0, \dots, \lceil \frac{dk}{2} \rceil\}$ gives the multiplicity, denoted by $c_w \in \mathbb{N}$, of the corresponding irreducible module, denoted by $V_\omega \subset \mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*)$. The elements of V_ω are recovered by applying the powers of the lowering operator on the respective highest weight vector. Applying X, resp. Y, on an element of weight ω we obtain an element of weight $\omega - 1$, resp. $\omega + 1$.

We can write the decomposition into irreducible SL_2-representation:

$$\mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*) = \bigoplus_{\omega=0}^{\lceil \frac{dk}{2} \rceil} V_\omega^{c_\omega}. \tag{4.8}$$

Lemma 4.2.9. *Let V_ω be the irreducible SL_2-module of the SL_2-representation $\mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*)$ associated to the weight $\omega \in \{0, \dots, \lceil \frac{dk}{2} \rceil\}$. Then*

$$\dim V_\omega = dk + 1 - 2\omega, \quad \textit{for} \quad \omega = 0, \dots, \left\lceil \frac{dk}{2} \right\rceil,$$

that is the module $V_\omega (= V_{(dk-\omega,\omega)})$ is isomorphic to the vector space $M_{dk-2\omega+1}$ of dimension $dk - 2\omega + 1$.

Proof. An irreducible SL_2-subrepresentation of $\mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*)$ is isomorphic to the module $V_\omega \simeq \mathrm{Sym}^{dk-2\omega}(\mathbb{C}^2)$. The highest weight vector associated to the weight $\omega = 0$ generates the irreducible SL_2-module $V_{dk} \simeq \mathrm{Sym}^{dk}(\mathbb{C}^2)$ of dimension $dk + 1$. Fixed $\omega \in \{0, \dots, \lceil \frac{dk}{2} \rceil\}$ we have $\dim V_\omega = \dim V_{dk} - 2\omega = dk + 1 - 2\omega$. □

Decomposition (4.8) can be written as

$$\mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*) = \bigoplus_{\omega=0}^{\lceil \frac{dk}{2} \rceil} (M_{dk-2\omega+1})^{\oplus c_\omega}.$$

Determining the highest weight vectors of the SL_2-representations

$$I_k(\mathrm{uMPS}(2, 2, d)) \subseteq \mathrm{Sym}^k(\mathrm{Dih}^d(\mathbb{C}^2)^*)$$

reveals the decomposition into irreducible SL_2-modules, i.e. the multiplicities $a_\omega \in \mathbb{N}$ such that

$$I_k(\mathrm{uMPS}(2, 2, d)) = \bigoplus_{\omega=0}^{\lceil \frac{dk}{2} \rceil} (M_{dk-2\omega+1})^{\oplus a_\omega}, \quad a_\omega \leq c_\omega.$$

Algorithm for highest weight vectors. The algorithm we describe does not find equations of the variety via an implicitization problem. The method is based instead on the solution of a linear system of equations in unknown coefficients. More precisely, a highest weight vector of the SL_2-representation $I_k(\text{uMPS}(2,2,d))$ is a polynomial in the variables of $\text{Dih}^d(\mathbb{C}^2)^*$ which is a *linear combination* of monomials of degree k and of the same weight $\omega \in \{0,\dots,\lceil\frac{dk}{2}\rceil\}$ (see Equation (4.5)) such that the following two conditions are satisfied:

1. the polynomial is killed by (the induced action of) the raising operator X (see Equation (4.7));
2. it is zero, evaluated in the generic point of the variety $\text{uMPS}(2,2,d)$.

Imposing these two conditions to the linear combination of monomials generates a linear system of equations in the unknown coefficients. We solve the linear system in order to obtain the highest weight vectors of the fixed weight ω.

Let k be a fixed degree and $\omega \in \{0,\dots,\lceil\frac{dk}{2}\rceil\}$ be a weight. Consider the linear combination of all the monomials $\underline{x}_\omega$ of $\mathbb{C}[\text{Dih}^d(\mathbb{C}^2)]_k$ of degree k and weight ω

$$f = \sum_h \gamma_h^\omega \underline{x}_\omega^h \in \mathbb{C}[\text{Dih}^d(\mathbb{C}^2)]_k, \tag{4.9}$$

where $\gamma_h^\omega \in \mathbb{C}$ are unknown coefficients. We impose the conditions that f is a highest weight vector, i.e. $X(f)=0$, and that it is an element of the ideal of $\text{uMPS}(2,2,d)$, i.e. $f(P)=0$, for every $P \in \text{uMPS}(2,2,d)$. This is equivalent to solving a linear system of equations in the indeterminates γ_h^ω.

Solve the linear system. We report a method proposed in [BM05] by N. Bray and J. Morton. The general problem is the following: given a polynomial map g with ring map $g^* : \mathbb{C}[x_1,\dots,x_{t'}] \to \mathbb{C}[y_1,\dots,y_{s'}]$, we want to find the ideal I_g of relations among the x_i. Let $x^a = \prod_{i=1}^{t'} x_i^{a_i} \in \mathbb{C}[x_1,\dots,x_{t'}]$ and denote $g^*(x^a) = \mathbf{g}^a \in \mathbb{C}[y_1,\dots,y_{s'}]$. Then I_g is the ideal of polynomials $\sum_a \alpha_a x^a \in \mathbb{C}[x_1,\dots,x_{t'}]$ such that $\sum_a \alpha_a \mathbf{g}^a = 0$.

Fix the degree k and let $\mathcal{P} = \{x^a : x^a \in \mathbb{C}[x_1,\dots,x_{t'}]_k\} = \{\mathbf{x}_1,\dots,\mathbf{x}_t\} \subset \mathbb{C}[x_1,\dots,x_{t'}]_k$ be the set of all monomials of degree k in the variables $\{x_i\}_{i=1,\dots,t'}$. Denote the image of the set by $g^*(\mathcal{P}) = \{\mathbf{g}_1,\dots,\mathbf{g}_t\} \subset \mathbb{C}[y_1,\dots,y_{s'}]$. Notice that the element $\mathbf{g}_i$ for $i=1,\dots,t$ can be a linear combination of monomials in $\{y_i\}_{i=1,\dots,s'}$. If we denote by $\{\mathbf{y}_1,\dots,\mathbf{y}_s\}$ the set of all *monomials* in $\{y_i\}_{i=1,\dots,s'}$ that appear in $g^*(\mathcal{P})$ then we can write

$$\mathbf{g}_i = \sum_{j=1}^{s} \beta_{ij}\mathbf{y}_j,$$

for every $i=1,\dots,t$ and for some coefficients $\beta_{ij} \in \mathbb{C}$.

Consider the matrix of coefficients $B = (\beta_{ij}) \in \mathbb{C}^{t\times s}$ and its transpose $B^t = (\beta'_{ij}) \in \mathbb{C}^{s\times t}$ with $\beta'_{ij} = \beta_{ji}$. We compute a set of generators of the kernel of B^t with Macaulay2

[GS20]. They consist in q vectors: $\alpha^l = (\alpha^l_1, \dots, \alpha^l_t)$, for $l = 1, \dots, q$ and they are such that, for every for $l = 1, \dots, q$, we have

$$\sum_{i=1}^{t} \alpha_i^l \mathbf{g}_i = \sum_{i=1}^{t} \alpha_i^l \left(\sum_{j=1}^{s} \beta_{ij} \mathbf{y}_j \right) = \sum_{j=1}^{s} \left(\sum_{i=1}^{t} \beta_{ij} \alpha_i^l \right) \mathbf{y}_j = \sum_{j=1}^{s} \left(\sum_{i=1}^{t} \beta'_{ji} \alpha_i^l \right) \mathbf{y}_j = 0.$$

In particular, we obtain the relations among $\{\mathbf{x}_i\}_{i=1,\dots,t}$, i.e. among the monomials of degree k in the variables $\{x_i\}_{i=1,\dots,t'}$, given by

$$\sum_{i=1}^{t} \alpha_i^l \mathbf{x}_i \in I_g \cap \mathbb{C}[x_1, \dots, x_{t'}]_k, \quad \text{for every } l = 1, \dots, q.$$

We apply this technique, degree by degree and weight by weight, to find the coefficients γ_h^ω's of Equation (4.9). In our case the map g^* is first the linear map induced by the rising operator X on $\mathbb{C}[\mathrm{Dih}^d(\mathbb{C}^2)]_k$ and then the evaluation map in the coordinates of the trace parametrization.

Decomposition tables. We do not report all the equations of the highest weight vectors obtained with our code. They are provided using the code available at `https://github.com/claudia-dela/uMPS_highest-weight-vectors/`. Instead, we display the decomposition into irreducible SL_2-representations of the degree 1, 2 and 3 part of the ideal in Tables 4.2, 4.3 and 4.4, respectively. The decomposition is given by

$$I_k(\mathrm{uMPS}(2,2,d)) = \bigoplus_{\omega=0}^{\lceil \frac{dk}{2} \rceil} (M_{dk-2\omega+1})^{\oplus a_\omega}.$$

where $M_{dk-2\omega+1}$ is a $dk - 2\omega + 1$-dimensional irreducible module of SL_2 and a_ω its multiplicity. In particular M_i is an i-dimensional irreducible SL_2-module associated to a_i highest weight vectors of weight $\omega = \frac{dk+1-i}{2}$.

As an example, consider the degree 2 part of the ideal of $\mathrm{uMPS}(2,2,8)$ given in Table 4.3. The last term of the third line of Table 4.3 is $5M_9$. The corresponding weight of the module M_9 is $\omega = \frac{dk+1-9}{2} = \frac{8\cdot 2+1-9}{2} = 4$. One of the 5 equations of degree 2 of $\mathrm{uMPS}(2,2,8)$ of weight 4, that we denote by $p^4_{8,3}$, has the expression:

$$\begin{aligned} p^4_{8,3} =& x_{00000000}x_{11101000} - x_{00000000}x_{11011000} - x_{00000000}x_{11100100} + \\ & - x_{00000000}x_{11010100} + x_{00000000}x_{10110100} + x_{00000000}x_{11001100}. \end{aligned}$$

d	$I_1(\mathrm{uMPS}(2,2,d))$
8	M_1
9	$M_2 + M_4$
10	$0M_1 + 4M_3 + M_5$
11	$4M_2 + 6M_4 + 2M_6$
12	$9M_1 + 7M_3 + 11M_5 + 3M_7$
13	$16M_2 + 18M_4 + 15M_6 + 4M_8$
14	$7M_1 + 41M_3 + 28M_5 + 23M_7 + 5M_9$
15	$44M_2 + 68M_4 + 49M_6 + 29M_8 + 7M_{10}$
16	$61M_1 + 92M_3 + 124M_5 + 69M_7 + 41M_9 + 8M_{11}$

Table 4.2: Decomposition of the degree one part of the ideal of $\mathrm{uMPS}(2,2,d)$ into irreducible SL_2-representations, for $d = 8, \ldots, 16$.

d	$I_2(\mathrm{uMPS}(2,2,d))$
6	$M_1 + M_5$
7	$0M_1 + 3M_3 + 2M_5 + M_7$
8	$10M_1 + 7M_3 + 15M_5 + 5M_7 + 5M_9$
9	$0M_1 + 31M_3 + 31M_5 + 27M_7 + 14M_9 + 6M_{11} + M_{13}$

Table 4.3: Decomposition of the degree 2 part of the ideal of $\mathrm{uMPS}(2,2,d)$ into irreducible SL_2-representations, for $d = 6, 7, 8, 9$.

d	$I_3(\mathrm{uMPS}(2,2,d))$
6	$6M_3 + 3M_5 + 6M_7 + M_9 + M_{11}$
7	$10M_2 + 18M_4 + 20M_6 + 14M_8 + 11M_{10} + 3M_{12} + M_{14}$
8	$54M_1 + 64M_3 + 122M_5 + 94M_7 + 97M_9 + 50M_{11} + 37M_{13} + 9M_{15} + 5M_{17}$

Table 4.4: Decomposition of the degree 3 part of the ideal of $\mathrm{uMPS}(2,2,d)$ into irreducible SL_2-representations, for $d = 6, 7, 8$.

In particular, comparing the decomposition of the ambient space $\mathrm{Dih}^d(\mathbb{C}^2)$ (Table 4.5) with the decomposition of the degree one part of the ideal (Table 4.2) we obtain the decomposition of the linear span of $\mathrm{uMPS}(2,2,d)$, for $d = 8, \ldots, 16$, whose dimension agrees with the dimension given in Table 4.1 of the previous subsection.

d	$\mathrm{Dih}^d(\mathbb{C}^2)$
8	$3M_1 + M_3 + 3M_5 + M_9$
9	$3M_2 + 3M_4 + 3M_6 + M_{10}$
10	$0M_1 + 8M_3 + 3M_5 + 4M_7 + M_{11}$
11	$6M_2 + 10M_4 + 5M_6 + 4M_8 + M_{12}$
12	$12M_1 + 9M_3 + 17M_5 + 6M_7 + 5M_9 + M_{13}$
13	$19M_2 + 22M_4 + 21M_6 + 8M_8 + 5M_{10} + M_{14}$
14	$7M_1 + 47M_3 + 32M_5 + 31M_7 + 9M_9 + 6M_{11} + M_{15}$
15	$47M_2 + 74M_4 + 55M_6 + 37M_8 + 12M_{10} + 6M_{12} + M_{16}$
16	$65M_1 + 95M_3 + 133M_5 + 75M_7 + 51M_9 + 13M_{11} + 7M_{13} + M_{17}$

Table 4.5: Decomposition of the space of dihedrally symmetric tensors into irreducible SL_2-representations, for $d = 8, \ldots, 16$.

d	Linear span $\langle \mathrm{uMPS}(2, 2, d) \rangle$
8	$2M_1 + M_3 + 3M_5 + M_9$
9	$2M_2 + 2M_4 + 3M_6 + M_{10}$
10	$0M_1 + 4M_3 + 2M_5 + 4M_7 + M_{11}$
11	$2M_2 + 4M_4 + 3M_6 + 4M_8 + M_{12}$
12	$3M_1 + 2M_3 + 6M_5 + 3M_7 + 5M_9 + M_{13}$
13	$3M_2 + 4M_4 + 6M_6 + 4M_8 + 5M_{10} + M_{14}$
14	$0M_1 + 6M_3 + 4M_5 + 8M_7 + 4M_9 + 6M_{11} + M_{15}$
15	$3M_2 + 6M_4 + 6M_6 + 8M_8 + 5M_{10} + 6M_{12} + M_{16}$
16	$4M_1 + 3M_3 + 9M_5 + 6M_7 + 10M_9 + 5M_{11} + 7M_{13} + M_{17}$

Table 4.6: Decomposition of the linear span of $\mathrm{uMPS}(2, 2, d)$ into irreducible SL_2-representations, for $d = 8, \ldots, 16$.

4.3 Linear relations via Cayley-Hamilton

We show that the linear span of the space of cyclically invariant matrix product states $\mathrm{uMPS}(m, n, d)$ is a proper subspace of the space of cyclically invariant tensors $\mathrm{Cyc}^d(\mathbb{C}^n)$ for $n \geq m + 2$ and $d \geq \frac{(m+1)(m+2)}{2}$. In particular, Theorem 4.3.6 gives nontrivial trace relations that do not follow either from cyclic permutations or reflections. We recall the classical Cayley-Hamilton theorem. We apply this result in order to find linear equations of $\mathrm{uMPS}(m, n, d + t) \subseteq \mathrm{Cyc}^{d+t}(\mathbb{C}^n)$, $t \geq m$ based on linear equations of $\mathrm{uMPS}(m, n, N) \subseteq \mathrm{Cyc}^N(\mathbb{C}^n)$, $N = d, d + 1, \ldots, d + m - 1$, in Lemma 4.3.3.

Theorem 4.3.1 (Cayley-Hamilton)**.** *Let $A \in \mathbb{C}^{m \times m}$ be an $m \times m$ complex matrix and $p_A(\lambda) = \det(\lambda \mathrm{Id}_m - A)$ be its characteristic polynomial. Then $p_A(A) = 0$.*

In fact, the only thing we will use is the following statement, which (since $\deg p_A = m$) immediately follows from Theorem 4.3.1.

Corollary 4.3.2. *Let $A \in \mathbb{C}^{m\times m}$. Then A^q, for $q \geq m$ can be written as a linear combination of its previous powers.*

Proof. Consider $A^q = A^m A^{q-m}$ with $q \geq m$. Let p_A be the characteristic polynomial. If $q = m$ then by the Cayley-Hamilton Theorem

$$p_A(A) = A^m + \gamma_{m-1}A^{m-1} + \cdots + \gamma_1 A + \gamma_0 \mathrm{Id}_m = 0.$$

$A^q = A^m = -(\gamma_{m-1}A^{m-1} + \cdots + \gamma_1 A + \gamma_0 \mathrm{Id}_m)$, i.e. A^q, $q = m$ can be written as a linear combination of its previous powers A^i, for $i = 0, \ldots, m-1$. Assume that $q > m$, then

$$\begin{aligned} A^m A^{q-m} &= A^m A A^{q-m-1} = -(\gamma_{m-1}A^{m-1} + \cdots + \gamma_1 A + \gamma_0 \mathrm{Id}_m) A A^{q-m-1} \\ &= -(\gamma_{m-1}A^m + \cdots + \gamma_1 A^2 + \gamma_0 A) A^{q-m-1}, \end{aligned}$$

and after $q - m - 1$ steps we have

$$A^q = -(\gamma_{m-1}A^{q-1} + \cdots + \gamma_1 A^{q-m+1} + \gamma_0 A^{q-m}).$$

This concludes the proof. □

Lemma 4.3.3. *Let $c = (c_1, \ldots, c_s) \in \mathbb{C}^s$ be a vector of coefficients and $\{i_\ell^j\}_{1\leq \ell \leq d, 1 \leq j \leq s}$ be indices, with $i_\ell^j \in [n]$. Assume that for every n-tuple $(A_0, \ldots, A_{n-1})$ of $m \times m$ matrices and every $k < m$ the following identity holds:*

$$\sum_{j=1}^{s} c_j \operatorname{Tr}(A_{i_1^j} \cdots A_{i_d^j} A_0^k) = 0. \tag{4.10}$$

Then the same identity holds for arbitrary $k \in \mathbb{N}$.

Proof. We use induction on $k \geq m$. By Corollary 4.3.2, A_0^k can be written as a linear combination of its previous powers A_0^q, for $q = 0, \ldots, m-1$. Therefore, we have

$$\begin{aligned} \sum_{j=1}^{s} c_j \operatorname{Tr}(A_{i_1^j} \cdots A_{i_d^j} A_0^k) &= \sum_{j=1}^{s} c_j \operatorname{Tr}\Big[A_{i_1^j} \cdots A_{i_d^j} \Big(\sum_{l=0}^{m-1} \gamma_l A_0^q \Big)\Big] \\ &= \sum_{j=1}^{s} c_j \sum_{l=0}^{m-1} \gamma_l \big(\operatorname{Tr}(A_{i_1^j} \cdots A_{i_d^j} A_0^q)\big) \\ &= \sum_{l=0}^{m-1} \gamma_l \Big(\sum_{j=1}^{s} c_j \big(\operatorname{Tr}(A_{i_1^j} \cdots A_{i_d^j} A_0^q)\big)\Big) = 0. \end{aligned}$$

□

Corollary 4.3.4. *Let $c = (c_1, \dots, c_s) \in \mathbb{C}^s$ be a vector of coefficients and let $l_j = (i_1^j \dots i_d^j) \in W(n,d)$, for $j = 1, \dots, s$, Let $0^m := \underbrace{(0 \dots 0)}_{m\text{-times}}$ be the word of m zeros and let*

$$l_j 0^m = (i_1^j \dots i_d^j \underbrace{0 \dots 0}_{m\text{-times}})$$

be the word obtained concatenating l_j and 0^m. Then, if the following relations hold

$$\begin{aligned} h_0 &:= c_1 x_{l_1} + \dots + c_s x_{l_s} \in I_1(\mathrm{uMPS}(m,n,d)), \\ h_1 &:= c_1 x_{l_1 0} + \dots + c_s x_{l_s 0} \in I_1(\mathrm{uMPS}(m,n,d+1)) \\ &\vdots \\ h_{m-1} &:= c_1 x_{l_1 0^{m-1}} + \dots + c_s x_{l_s 0^{m-1}} \in I_1(\mathrm{uMPS}(m,n,d+m-1)), \end{aligned}$$

then, for every $q \geq m$, it holds

$$h_t := c_1 x_{l_1 0^q} + \dots + c_s x_{l_s 0^q} \in I_1(\mathrm{uMPS}(m,n,d+t)).$$

Proof. Simply notice that, for every $q \in \mathbb{N}$, $h_q \in I_1(\mathrm{uMPS}(m,n,d+q)$ if and only if

$$\sum_{j=1}^{s} c_j \operatorname{Tr}(A_{i_1^j} \cdots A_{i_d^j} A_0^q) = 0.$$

The result descends directly from Lemma 4.3.3. □

The usefulness of Lemma 4.3.3 stems from the fact that one can find expressions of the form (4.10) which are trivial for small k, in the sense that they follow from the cyclic invariance of the trace, but nontrivial for large k. We illustrate this in the example below.

Example 4.3.5. We show that for any 2×2 matrices A_0, A_1, A_2, A_3 and any $k \geq 0$, the following identity holds

$$\begin{aligned} &\operatorname{Tr}(A_1 A_2 A_0 A_3 A_0^k) + \operatorname{Tr}(A_2 A_3 A_0 A_1 A_0^k) + \operatorname{Tr}(A_3 A_1 A_0 A_2 A_0^k) \\ =&\operatorname{Tr}(A_1 A_0 A_2 A_3 A_0^k) + \operatorname{Tr}(A_2 A_0 A_3 A_1 A_0^k) + \operatorname{Tr}(A_3 A_0 A_1 A_2 A_0^k). \end{aligned}$$

By the above argument, it suffices to show the identity for $k = 0$ and $k = 1$. But these both follow from cyclic invariance of the trace:

$$\begin{aligned} &\underline{\operatorname{Tr}(A_1 A_2 A_0 A_3)} + \underline{\operatorname{Tr}(A_2 A_3 A_0 A_1)} + \underline{\operatorname{Tr}(A_3 A_1 A_0 A_2)} \\ =&\underline{\operatorname{Tr}(A_1 A_0 A_2 A_3)} + \underline{\operatorname{Tr}(A_2 A_0 A_3 A_1)} + \underline{\operatorname{Tr}(A_3 A_0 A_1 A_2)} \end{aligned}$$

$$\begin{aligned} &\underline{\operatorname{Tr}(A_1 A_2 A_0 A_3 A_0)} + \underline{\operatorname{Tr}(A_2 A_3 A_0 A_1 A_0)} + \underline{\operatorname{Tr}(A_3 A_1 A_0 A_2 A_0)} \\ =&\underline{\operatorname{Tr}(A_1 A_0 A_2 A_3 A_0)} + \underline{\operatorname{Tr}(A_2 A_0 A_3 A_1 A_0)} + \underline{\operatorname{Tr}(A_3 A_0 A_1 A_2 A_0)}. \end{aligned}$$

This immediately gives us a nontrivial linear relation on uMPS$(2,4,d)$, for $d \geq 6$. But we can also find linear relations on uMPS$(2,2,d)$. For instance, if we put $k=2$, $A_2 = A_1^2$ and $A_3 = A_0A_1$, we find

$$\begin{aligned}&\operatorname{Tr}(A_1^3A_0^2A_1A_0^2) + \operatorname{Tr}(A_1^2A_0A_1A_0A_1A_0^2) + \operatorname{Tr}(A_1^2A_0^3A_1^2A_0)\\ =&\operatorname{Tr}(A_1^2A_0A_1A_0^2A_1A_0) + \operatorname{Tr}(A_1^2A_0^2A_1^2A_0^2) + \operatorname{Tr}(A_1^3A_0^3A_1A_0),\end{aligned}$$

which is the unique linear relation on uMPS$(2,2,8)$ that doesn't follow from dihedral symmetry.

Theorem 4.3.6. *Let $A_0,\dots,A_m,B$ be $m\times m$ matrices. Then for every $\ell\in\mathbb{N}$ it holds that*

$$\sum_{\sigma\in\mathfrak{S}_m,\tau\in C_{m+1}} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\operatorname{Tr}(A_{\tau(0)}B^{\sigma(0)}A_{\tau(1)}B^{\sigma(1)}\cdots A_{\tau(m-1)}B^{\sigma(m-1)}A_{\tau(m)}B^{\ell}) = 0. \tag{4.11}$$

Here $\mathfrak{S}_m$ is the symmetric group acting on $\{0,1,\dots,m-1\}$, and C_{m+1} is the cyclic group acting on $\{0,1,\dots,m\}$.

Proof. We will first show the statement for $\ell\in\{0,1,\dots,m-1\}$. So let us fix such an ℓ. We will write

$$T(\sigma,\tau) := \operatorname{Tr}(A_{\tau(0)}B^{\sigma(0)}A_{\tau(1)}B^{\sigma(1)}\cdots A_{\tau(m-1)}B^{\sigma(m-1)}A_{\tau(m)}B^{\ell}).$$

Let us write c_a for the permutation that cyclically permutes the first a elements. Precisely

$$c_a(i) = \begin{cases} i+1 & \text{for } i < a-1\\ 0 & \text{for } i = a-1\\ i & \text{for } i > a-1.\end{cases}$$

Step 1. For $\sigma\in\mathfrak{S}_m$ and $\tau\in C_{m+1}$, we define

$$\begin{aligned}\widetilde{\sigma} &:= \sigma\circ c_{\sigma^{-1}(\ell)+1}^{-1}\circ c_m^{\sigma^{-1}(\ell)+1}\\ \widetilde{\tau} &:= \tau\circ c_{m+1}^{\sigma^{-1}(\ell)+1}\end{aligned}$$

we have $T(\sigma,\tau) = T(\widetilde{\sigma},\widetilde{\tau})$.

Indeed, if we write $k=\sigma^{-1}(\ell)$, then

$$\begin{aligned}T(\sigma,\tau) &= \operatorname{Tr}(A_{\tau(0)}B^{\sigma(0)}\cdots A_{\tau(k)}B^{\sigma(k)}A_{\tau(k+1)}B^{\sigma(k+1)}\cdots A_{\tau(m-1)}B^{\sigma(m-1)}A_{\tau(m)}B^{\sigma(k)})\\ &= \operatorname{Tr}(A_{\tau(k+1)}B^{\sigma(k+1)}\cdots A_{\tau(m)}B^{\sigma(k)}A_{\tau(0)}B^{\sigma(0)}\cdots A_{\tau(k-1)}B^{\sigma(k-1)}A_{\tau(k)}B^{\sigma(k)})\\ &= T(\widetilde{\sigma},\widetilde{\tau}).\end{aligned}$$

Where for the last step, note that

- $k+1 = c_{m+1}^{k+1}(0)$, …, $m = c_{m+1}^{k+1}(m-k-1)$, $0 = c_{m+1}^{k+1}(m-k)$, …, $k = c_{m+1}^{k+1}(m)$.

- $k{+}1 = c_{k+1}^{-1}(c_m^{k+1}(0)), \ldots, m{-}1 = c_{k+1}^{-1}(c_m^{k+1}(m{-}k{-}2)), k = c_{k+1}^{-1}(c_m^{k+1}(m{-}k{-}1)),$ $0 = c_{k+1}^{-1}(c_m^{k+1}(m-k)), \ldots, k-1 = c_{k+1}^{-1}(c_m^{k+1}(m-1))$.

Step 2. Note that the assignment

$$\begin{aligned} \mathfrak{S}_m \times C_{m+1} &\to \mathfrak{S}_m \times C_{m+1} \\ (\sigma, \tau) &\mapsto (\widetilde{\sigma}, \widetilde{\tau}) \end{aligned}$$

is an involution. Indeed, we have

$$\widetilde{\sigma}^{-1}(\ell) = c_m^{-\sigma^{-1}(\ell)-1}(c_{\sigma^{-1}(\ell)+1}(\sigma^{-1}(\ell))) = c_m^{-\sigma^{-1}(\ell)-1}(0) = m - \sigma^{-1}(\ell) - 1.$$

So, again writing $k = \sigma^{-1}(\ell)$

$$\begin{aligned} \widetilde{\widetilde{\sigma}} &= \widetilde{\sigma} \circ c_{\widetilde{\sigma}^{-1}(\ell)+1}^{-1} \circ c_m^{\widetilde{\sigma}^{-1}(\ell)+1} \\ &= \sigma \circ c_{k+1}^{-1} \circ c_m^{k+1} \circ c_{m-k}^{-1} \circ c_m^{m-k} \\ &= \sigma. \end{aligned}$$

To see the last equality:

- For $i < k$: $c_{k+1}^{-1}(c_m^{k+1}(c_{m-k}^{-1}(c_m^{m-k}(i)))) = c_{k+1}^{-1}(c_m^{k+1}(c_{m-k}^{-1}(m-k+i))) = c_{k+1}^{-1}(c_m^{k+1}(m-k+i)) = c_{k+1}^{-1}(i+1) = i.$
- For $i = k$: $c_{k+1}^{-1}(c_m^{k+1}(c_{m-k}^{-1}(c_m^{m-k}(k)))) = c_{k+1}^{-1}(c_m^{k+1}(c_{m-k}^{-1}(0))) = c_{k+1}^{-1}(c_m^{k+1}(m-k-1)) = c_{k+1}^{-1}(0) = k.$
- For $i > k$: $c_{k+1}^{-1}(c_m^{k+1}(c_{m-k}^{-1}(c_m^{m-k}(i)))) = c_{k+1}^{-1}(c_m^{k+1}(c_{m-k}^{-1}(i-k))) = c_{k+1}^{-1}(c_m^{k+1}(i-k-1)) = c_{k+1}^{-1}(i) = i.$

And furthermore $\widetilde{\widetilde{\tau}} = \tau \circ c_{m+1}^{k+1} \circ c_{m+1}^{m-k} = \tau$.

Step 3. Note that

$$\begin{aligned} \operatorname{sgn}(\widetilde{\sigma}) \operatorname{sgn}(\widetilde{\tau}) &= (-1)^{k+(k+1)(m-1)+(k+1)m} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \\ &= -\operatorname{sgn}(\sigma) \operatorname{sgn}(\tau). \end{aligned}$$

From Step 1 we have that $T(\sigma, \tau) = T(\widetilde{\sigma}, \widetilde{\tau})$. There will be cancellations of terms in (4.11) as $\operatorname{sgn}(\widetilde{\sigma}) \operatorname{sgn}(\widetilde{\tau}) = -\operatorname{sgn}(\sigma) \operatorname{sgn}(\tau)$ from Step 3. Finally using Step 2 we ensure that all terms will cancel out therefore establishing the given identity for $0 \leq l \leq m-1$. Now using Lemma 4.3.3, we conclude that the given identity (4.11) holds for all $l \in \mathbb{N}$. □

Corollary 4.3.7. *If $n \geq 3$ and $d \geq \frac{(m+1)(m+2)}{2}$, then* uMPS$(m, n, d)$ *is contained in a proper linear subspace of the space of cyclically invariant tensors.*

Proof. Let $\ell \geq m$ and let $\mathfrak{S}_{m+1}$ denote the symmetric group acting on $\{0, 1, \ldots, m-1, \ell\}$. Then we can rewrite (4.11) as follows:

$$\sum_{\sigma \in \mathfrak{S}_{m+1}} \operatorname{sgn}(\sigma) \operatorname{Tr}(A_0 B^{\sigma(0)} A_1 B^{\sigma(1)} \cdots A_{m-1} B^{\sigma(m-1)} A_m B^{\sigma(\ell)}) = 0. \tag{4.12}$$

Let X_0, X_1, X_2 be $m \times m$ matrices, and in (4.12) substitute $A_0 = X_0$, $B = X_1$, and $A_i = X_2$ for $i = 1, \ldots, m$. Note that even after that substitution, the ternary bracelets corresponding to the $(m+1)!$ terms in (4.12) are all distinct. Hence no two terms will cancel, and we get a nontrivial linear relation on $\mathrm{uMPS}(m, 3, d)$, where $d = 1 + 2 + \cdots + (m-1) + \ell + (m+1) \geq 1 + 2 + \cdots + (m+1) = \binom{m+2}{2}$. □

Remark 4.3.8. With a bit more care, one can also get nontrivial relations on $\mathrm{uMPS}(m, 2, d)$ in this way. For instance if we take $\ell = m$ and in (4.12) we substitute $A_0 = X_0 X_1^{m+1} X_0$, $B = X_1$, and $A_i = X_0$ for $i = 1, \ldots, m$, one verifies that again no terms cancel, and hence we found a nontrivial linear relation on $\mathrm{uMPS}(m, 2, d)$, where $d = \binom{m+3}{2}$.

4.3.1 Linear equations for $\mathrm{uMPS}(2, 2, d)$

From the trace parametrization, we can give an upper bound on $\dim\langle \mathrm{uMPS}(2, 2, d)\rangle$.

Theorem 4.3.9. *For every $d \in \mathbb{N}$, we have the inequality*

$$\dim\langle \mathrm{uMPS}(2, 2, d)\rangle \leq \begin{cases} \frac{1}{192}(d+6)(d+4)^2(d+2) & \text{for } d \text{ even}, \\ \frac{1}{192}(d+7)(d+5)(d+3)(d+1) & \text{for } d \text{ odd}. \end{cases}$$

Proof. It follows from (4.3) and Remark 4.2.6 that $\dim\langle \mathrm{uMPS}(2, 2, d)\rangle$ can be at most the number of degree d monomials in $\mathbb{C}[T_0, T_1, T_{00}, T_{01}, T_{11}]$. Counting these monomials gives the above formula. □

Note that asymptotically for $d \to \infty$, the above bound agrees with our conjectured formula in Conjecture 4.2.8.

As in the previous section, we abbreviate "weight $\lambda = (w, d-w) \in \mathbb{Z}^2$" to "weight w". In the rest of this section, we provide a proof of Conjecture 4.2.7 in the cases $w = 0, 1, 2, 3$.

Consider the parametrization of $\mathrm{uMPS}(2, 2, d)$ in coordinates

$$\begin{aligned} \phi : (\mathbb{C}^{2\times 2})^2 &\to \mathrm{Dih}^d(\mathbb{C}^2) \\ (A_0, A_1) &\mapsto (\operatorname{Tr}(A_0^d), \operatorname{Tr}(A_0^{d-1} A_1), \ldots, \operatorname{Tr}(A_1^d)). \end{aligned}$$

It is in particular a polynomial map in the unknown entries of the matrices $A_0, A_1 \in \mathbb{C}^{2\times 2}$, we denote by

$$A_0 = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \quad A_1 = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}.$$

We will write

$$T_{i_1\dots i_d} := \mathrm{Tr}(A_{i_1}\dots A_{i_d}) \in \mathbb{C}[a_1,\dots,b_4]_d \quad \text{and} \quad W_w := \langle T_b : b \in B_w(2,d)\rangle$$

By Observation 4.2.1, we have

$$\dim\langle \mathrm{uMPS}(2,2,d)\rangle_w = \dim W_w.$$

The cases $w = 0$ and $w = 1$ are easy:

Proposition 4.3.10. *The space W_0 is a 1-dimensional vector space generated by the polynomial $T_{0,\dots,0} = \mathrm{Tr}(A_0^d)$. The space W_1 is a 1-dimensional vector space generated by the polynomial $T_{10\dots0} = \mathrm{Tr}(A_1A_0^{d-1})$.*

Proof. If $w = 0$ then $b = (0\dots0) \in B_0(2,d)$ is the only binary bracelet of weight zero, and if $w = 1$ then $b = (10\dots0) \in B_1(2,d)$ is the only binary bracelet of weight 1. □

We now turn to the case $w = 2$. Then Conjecture 4.2.7 states that $\dim W_w = \lfloor \frac{d}{2} \rfloor$. But $\lfloor \frac{d}{2} \rfloor$ is exactly the number $B_2(2,d)$ of generators T_b of W_w; hence we need to show that they are linearly independent.

Proposition 4.3.11. *The polynomials $\{T_b : b \in B_2(2,d)\}$ are linearly independent.*

Proof. Note that

$$W_2 = \langle T_{10^i10^{d-2-i}} \mid i = 0,\dots \left\lfloor \frac{d}{2} \right\rfloor - 1\rangle.$$

If we make the following substitutions:

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}$$

our generators T_b become

$$T_{10^i10^{d-2-i}} = \mathrm{Tr}(A_1A_0^iA_1A_0^{d-2-i}) = x^i + x^{d-2-i} \quad i = 0,\dots \left\lfloor \frac{d}{2} \right\rfloor - 1. \tag{4.13}$$

Since for the given choice of A_0, A_1 the polynomials (4.13) are $\lfloor \frac{d}{2} \rfloor$ linearly independent polynomials, the same holds for a generic choice of matrices. □

Finally, we prove the case $w = 3$. In this case our conjectured formula states that $\dim W_w = d - 3$. Consider the following subset of $B_3(2,d)$:

$$\widetilde{B}_3 := \{b \in B_3(2,d) : b \text{ contains } 11 \text{ or } 101\} \subset B_3(2,d).$$

Lemma 4.3.12. *The cardinality of $\widetilde{B}_3$ equals $d - 3$.*

Proof. The cardinality of $\widetilde{B}_3$ is the sum of the number of binary bracelets of weight 3 containing 11 and the number of binary bracelets of weight 3 containing 101 but not 11, that are $\lceil \frac{d-2}{2} \rceil$ and $\lceil \frac{d-5}{2} \rceil$ respectively. Therefore the cardinality of $\widetilde{B}_3$ is

$$\left\lceil \frac{d-2}{2} \right\rceil + \left\lceil \frac{d-5}{2} \right\rceil = d-3. \qquad \square$$

In order to prove the case $w = 3$ we need to show that $\{T_b : b \in \widetilde{B}_3\}$ is a basis of W_3. We first show linear independence:

Lemma 4.3.13. *The polynomials $\{T_b : b \in \widetilde{B}_3\}$ are linearly independent.*

Proof. We will show that the polynomials are linearly independent even after the following substitution:

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}.$$

Then W_3 is spanned by the following polynomials:

$$f_b := T_{110^b10^{d-b-3}} = x^b + x^{d-b-3} + 2x^{d-3} \quad b \in \{0, \ldots, \left\lfloor \frac{d-3}{2} \right\rfloor\},$$
$$g_b := T_{1010^b10^{d-b-4}} = x^{b+1} + x^{d-b-3} + x^{d-4} + x^{d-3} \quad b \in \{1, \ldots, \left\lfloor \frac{d-4}{2} \right\rfloor\}.$$

We now simply have to put the coefficients of these polynomials in a matrix and show it has full rank. For d even the matrix of coefficients is given by

$$S = \begin{pmatrix}
1 & & & & \cdots & & & & & 3 \\
 & 1 & & & & & & & 1 & 2 \\
 & & 1 & & & & & 1 & & 2 \\
\vdots & & & \ddots & & & \iddots & & & \vdots \\
0 & & & & 1 & 1 & & & & 2 \\
0 & 0 & 1 & 0 & & \cdots & & & 2 & 1 \\
 & & 0 & 1 & & & & 1 & 1 & 1 \\
\vdots & & & & \ddots & & \iddots & & & \vdots \\
0 & & & \cdots & & 2 & \cdots & 0 & 1 & 1
\end{pmatrix}$$

and for d odd, given by

$$S = \begin{pmatrix}
1 & & & & & & \cdots & & & & 0 & 3 \\
 & 1 & & & & & & & & 0 & 1 & 2 \\
 & & 1 & & & & & & & 1 & 0 & 2 \\
\vdots & & & \ddots & & & & & \iddots & & \vdots & \vdots \\
 & & & & 1 & & & 1 & & & 0 & 2 \\
 & & & & & 2 & & & & & 0 & 2 \\
0 & 0 & 1 & 0 & & & \cdots & & & 0 & 2 & 1 \\
 & & 0 & 1 & 0 & & & & & 1 & 1 & 1 \\
 & & & & \ddots & & & & \iddots & & \vdots & \vdots \\
0 & 0 & & & & 1 & 1 & & & 0 & 1 & 1
\end{pmatrix}.$$

By elementary row operations, we can reduce the left upper part to a diagonal matrix of order $\lfloor \frac{d-1}{2} \rfloor$. The left lower part is filled with zeros. The (rectangular) right lower block of dimension $\lfloor \frac{d-4}{2} \rfloor \times \lfloor \frac{d-2}{2} \rfloor$ can be put in the following upper triangular forms, for d even and odd respectively

$$\begin{pmatrix}
0 & & \cdots & -1 & 2 & -1 \\
\vdots & & -1 & 1 & 1 & -1 \\
 & \iddots & \iddots & 0 & 1 & -1 \\
-1 & 1 & & \vdots & \vdots & \vdots \\
2 & 0 & \cdots & 0 & 1 & 1
\end{pmatrix} \to \begin{pmatrix}
2 & 0 & \cdots & & 1 & 1 \\
0 & 2 & & & 3 & -1 \\
 & & \ddots & & 5 & -3 \\
\vdots & & & & \vdots & \vdots \\
0 & \cdots & & 0 & * & * \\
0 & \cdots & & 0 & * & *
\end{pmatrix} \quad \text{for } d \text{ even},$$

$$\begin{pmatrix}
0 & & \cdots & -1 & 2 & -1 \\
\vdots & & -1 & 1 & 1 & -1 \\
 & \iddots & \iddots & 0 & 1 & -1 \\
 & & & \vdots & \vdots & \vdots \\
-1 & 1 & & & & -1 \\
1 & 0 & \cdots & 0 & 1 & 0
\end{pmatrix} \to \begin{pmatrix}
1 & 0 & \cdots & & 0 & 1 & 0 \\
0 & 1 & 0 & \cdots & 0 & 2 & -1 \\
\vdots & 0 & \ddots & & & 3 & -2 \\
 & & & & & \vdots & \vdots \\
 & & & 0 & 1 & * & * \\
 & & & & 0 & * & *
\end{pmatrix} \quad \text{for } d = 2 \text{ odd}.$$

Both the obtained blocks have rank $\lfloor \frac{d-4}{2} \rfloor$. We have that the rank of S is $\lfloor \frac{d-1}{2} \rfloor + \lfloor \frac{d-4}{2} \rfloor = d - 3$, and this concludes the proof. □

We finish our proof by showing that $\{T_b : b \in \widetilde{B}_3\}$ spans W_3:

Lemma 4.3.14. *Every polynomial $T_{10^a10^b10^c} = \mathrm{Tr}(A_1A_0^aA_1A_0^bA_1A_0^c)$, with $1 < a \leq b \leq c$, $a + b + c = d - 3$ is an element of the linear span $\langle T_b : b \in \widetilde{B}_3 \rangle$.*

Proof. Notice that the elements of $B_3(2, n) \setminus \widetilde{B}_3$ can be written without loss of generality in the form

$$10^a10^b10^c, \quad \text{with } 1 < a \leq b \leq c.$$

We use induction on a. If $a = 0$ and $a = 1$ then $(10^a10^b10^c) \in \widetilde{B}_3$ and we are done. If we substitute $A_1 \to A_1A_0^{a-1}$, $A_2 \to A_1A_0^b$ and $A_3 \to A_1$ in the equation given by Theorem 4.3.6, we get

$$\begin{aligned}\operatorname{Tr}(A_1A_0^{a-1}A_1A_0^{b+1}A_1A_0^c) + \operatorname{Tr}(A_1A_0^bA_1A_0A_1A_0^{a+c-1}) + \operatorname{Tr}(A_1^2A_0^aA_1A_0^{b+c}) = \\ \operatorname{Tr}(A_1A_0^aA_1A_0^bA_1A_0^c) + \operatorname{Tr}(A_1A_0^{b+1}A_1^2A_0^{a+c-1}) + \operatorname{Tr}(A_1A_0A_1A_0^{a-1}A_1A_0^{b+c}).\end{aligned}$$

Reordering the summands we obtain

$$T_{10^a10^b10^c} = (T_{10^b1010^{a+c-1}} + T_{110^a10^{b+c}} - T_{10^{b+1}110^{a+c-1}} - T_{1010^{a-1}10^{b+c}}) + T_{10^{a-1}10^{b+1}10^c}.$$

All terms in the parenthesis have as subscript an element of $\widetilde{B}_3$, and the last term is in $\langle\{T_b : b \in \widetilde{B}_3\}\rangle$ by the induction hypothesis. This concludes the proof. □

Conclusions. In this chapter, we studied the linear span $\langle \text{uMPS}(m, n, d)\rangle$. Theorem 4.3.6 introduces a new method to find linear equations that vanish on $\langle \text{uMPS}(m, n, d)\rangle$, based on the Cayley-Hamilton theorem. As a corollary, we provided that for $d \geq (m + 1)(m + 2)$, the linear span $\langle \text{uMPS}(m, n, d)\rangle$ does not fill its natural ambient space $\text{Cyc}^d(\mathbb{C}^n)$, significantly improving the state of the art. For the special case $m = n = 2$, we described an algorithm that computes $\langle \text{uMPS}(2, 2, d)\rangle$, viewed as a GL_2-representation and we obtained a conjectured formula for its dimension that we proved in the first cases, using our Cayley-Hamilton technique. Moreover, using the trace parametrization we gave an upper bound on the dimension of $\langle \text{uMPS}(2, 2, d)\rangle$ which is close to optimal.

Chapter 5

Nonlinear conjugate gradient method on MPS

In quantum physics, every physical system is associated to a Hilbert space. Moreover, a system composed of several subsystems is associated to the tensor product of the Hilbert spaces associated to the subsystems. One of the main problems in studying or simulating these many-body quantum systems is the exponential growth of the dimension of the tensor product in the number of its factors.

Consider a quantum many-body system of d particles. Let $\mathcal{H}_k = \mathbb{C}^{n_k}$, for $k = 1, \dots, d$ be the state space of the single k-th component and let $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_d$ be the Hilbert space associated to the composite system. For every $k = 1, \dots, d$, fix a basis $\{e_{j_k}^{(k)}\}_{j_k=1}^{n_k}$ of $\mathcal{H}_k$. An element $\psi \in \mathcal{H}$ can be written as

$$\psi = \sum_{j_1,\dots,j_d=1}^{n_1,\dots,n_d} \psi_{j_1,\dots,j_d} \; e_{j_1}^{(1)} \otimes \cdots \otimes e_{j_d}^{(d)}.$$

The tensor ψ is usually called wave function and it is completely characterized by its coefficients $\psi_{j_1,\dots,j_d}$. We immediately notice that the dimension of the Hilbert space

$$\dim(\mathcal{H}) = \prod_{i=1}^{d} \dim(\mathcal{H}_i) = \prod_{i=1}^{d} n_i,$$

increases exponentially in the number of factors of the tensor product, meaning that the number of parameters needed to describe the state $\psi \in \mathcal{H}$ is exponentially large in the system size. From a computational point of view, this means that the representation of ψ is inefficient. This leads to the problem of finding an efficient representation of ψ that also provides an accurate physical description of the system, in particular of the expected entanglement properties of the state [Orú14].

Bond dimensions and entanglement. Tensor network varieties play a major role in this context, where they are used as a variational ansatz class to describe strongly correlated quantum systems, whose entanglement structure is given by the underlying graph. A state of the tensor network is the result of a contraction of tensors prescribed by the edges of the graph. By the physical interpretation of this construction, the edges of the graph encode the structure of the entanglement of the state ψ, and the bond dimension associated to every edge is considered a quantitative measure of the amount of quantum correlation in the wave function. For example, low-energy eigenstates of gapped Hamiltonians with local interactions exhibit little entanglement relative to typical states and therefore they can be efficiently approximated with matrix product states and projected entangled pair states with finite bond dimensions [ECP10, Has07, VWPGC06]. Particularly relevant and studied are therefore *lattice graphs*. We provide examples of matrix product states (MPS) and projected entangled pair states (PEPS) in Figure 5.1. Matrix product states (projected entangled pair states) with periodic boundary conditions are said to be translation invariant if the same tensor is associated to every vertex of the underlying graph. If the graph is the path graph, the associated matrix product state (projected entangled pair states) with open boundary conditions can be translation invariant only in the so called thermodynamic limit, where the number of vertices approaches infinity.

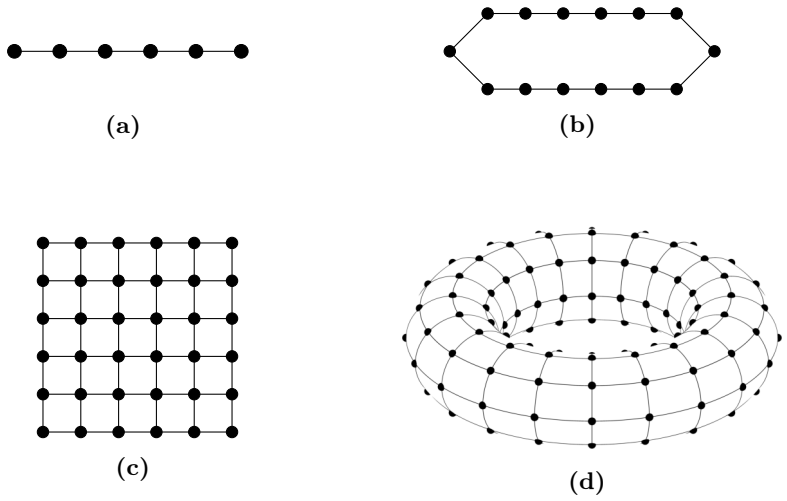


Figure 5.1: Examples of lattice graphs: MPS with open (a) and periodic (b) boundary conditions; PEPS with open (c) and periodic (d) boundary conditions.

We refer to [Orú14, STG+19, PGVWC07, CLVW20] for a full description of the subject from the viewpoint of quantum physics and for details on the construction of MPS, PEPS and other related entanglement structures. The entanglement theory behind the construction of the MPS ansatz led to the introduction of variational MPS algorithms applied to PEPS [VWPGC06, VC04] and tensor network varieties associated to higher dimensional graphs.

Dimension. Moreover, the dimension of tensor network varieties is polynomial in the system size allowing for an efficient parametrization of states of a composite system.

Let $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} \subset \mathcal{H}$ be the tensor network variety associated to $(\Gamma, \mathbf{m}, \mathbf{n})$. By Theorem 3.0.2 in Chapter 3 we have the following upper bound on the dimension of the variety

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} \leq \sum_{v \in \mathbf{v}(\Gamma)} (n_v \cdot N_v) - d + 1 - \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1) + \dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X),$$

where $N_v = \prod_{e \ni v} m_e$, $X = X_1 \otimes \cdots \otimes X_d$ with $X_v \in \mathrm{Hom}(W_v, V_v)$ generic and $\mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$ is the stabilizer of X under the action of the gauge subgroup.

By naive computation, if we denote by $N_{\max} = \max_{v \in \mathbf{v}(\Gamma)} N_v$, then

$$\dim \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} = O(\mathrm{poly}(d)\, \mathrm{poly}(N_{\max})).$$

Therefore, if we approximate the state ψ to be an element of $\mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}} \subset \mathcal{H}$, the number of parameters needed to describe ψ grows polynomially in the system size, allowing an efficient parametrization of the wave function.

State of the art. The original motivation in quantum physics is the description of quantum spin chains [AKLT88, FNW92, ÖR95]. In particular, I. Affleck, T. Kennedy, E. H. Lieb and H. Tasaki [AKLT88] proved that ground states of the AKLT model admitted an analytic solution corresponding to a matrix product state representation. While in general analytical solutions of ground states are too hard to be found, the AKLT model suggested the use of matrix product states as variational classes of tensors for the *approximation* of ground states.

Let $\mathcal{H} = \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$. Let $H : \mathcal{H} \to \mathcal{H}$ be the Hamiltonian describing the dynamics of a quantum physical system. We denote the lowest eigenvalue of H by $\lambda_0 \in \mathbb{R}$, and the associated eigenspace by E_{λ_0}. Vectors of E_{λ_0} are the so called ground state.

The expectation value of the Hamiltonian in $v \in \mathcal{H}$ is the real valued operator $\rho : \mathcal{H} \to \mathbb{R}$, defined as

$$\rho(v) := \langle H \rangle_v = \frac{v^\dagger H v}{v^\dagger v}.$$

The Courant-Fischer-Weyl min-max theorem [RS78] ensures that the minimum eigenvalue of H corresponds to

$$\lambda_0 = \min \{\rho(v) : \ v \in \mathcal{H}\}.$$

Therefore, the ground states of H correspond to eigenvectors that realize the minimum. Denote a ground state by v_0.

Problem 5.0.1. *The variational ansatz consists in approximating the ground state v_0 with a tensor $\widetilde{v}_0 \in \mathcal{H}$ of a chosen tensor network representation, i.e. fixed $(\Gamma, \mathbf{m}, \mathbf{n})$ tensor network, in finding*

$$\widetilde{v}_0 = \arg\min \left\{\rho(v) : \ v \in \mathcal{TNS}^{\Gamma\circ}_{\mathbf{m},\mathbf{n}}\right\}.$$

Matrix product states are the first nontrivial class of tensor network varieties and they received great attention in the recent years. Several techniques have been developed in order to study their geometry and to build ad hoc algorithms for exploiting the efficiency of the MPS representation.

Problem 5.0.2. *Given $\phi : D_{\mathrm{MPS}} \to \mathrm{MPS} \subset \mathcal{H}$, the parametrization of the variety of matrix product states, problem 5.0.1 can be restated as: finding*

$$\arg\min_{A \in D_{\mathrm{MPS}}}(\rho \circ \phi(A)) = \arg\min_{A \in D_{\mathrm{MPS}}} \frac{\phi(A)^\dagger H \phi(A)}{\phi(A)^\dagger \phi(A)}.$$

Equivalently, finding $A \in D_{\mathrm{MPS}}$ which realizes the minimum of the functional $f = \rho \circ \phi : D_{\mathrm{MPS}} \to \mathbb{R}$, given by

$$f(A) = \frac{\phi(A)^\dagger H \phi(A)}{\phi(A)^\dagger \phi(A)}. \tag{5.1}$$

Variational ansatz. White [Whi93] introduced an algorithm to find an approximate ground states that nowadays is the well-known Density Matrix Renormalization Group (DMRG) variational algorithm [Vid04, Sch11]. The DMRG is an adaptive algorithm that optimizes the MPS in order to make it an approximation of the lowest eigenvalue of a given Hamiltonian. The algorithm consists in the sequential solution of minor diagonalization problems: all but one or two tensors of the MPS are fixed, an eigensolver is performed on the reduced eigenvalue problem and then the MPS form (that can have been lost) is restored through a Singular Value Decomposition. The DMRG has been generalized to arbitrary loop-free tensor network formats (Tree-Tensor Networks) [STG+19] and MERA (Multiscaled Entanglement Renormalization Ansatz) [Vid07, Vid08, EV11, EV14].

The best known and most powerful method for approximating the time-evolving wave function within the MPS manifold is the Time-Evolving Block Decimation (TEBD) developed by Vidal [Vid04]. Given a Hamiltonian $H : \mathcal{H} \to \mathcal{H}$, the time evolution of a quantum state $\psi(t) \in \mathcal{H}$ is dictated by the Shrödinger equation (1.4)

$$i\hbar \frac{d}{dt}\psi(t) = H(t)\psi(t).$$

whose solution is given by

$$\psi(t) = e^{-itH/\hbar}\psi(t_0), \quad t_0 = 0.$$

The new idea is to use a Lie-Trotter-Suzuki decomposition for the evolution operator $e^{-i\delta H/\hbar}$, for small time steps $\delta \ll 1$. Given a decomposition $H = H^{(1)} + H^{(2)}$ such that $H^{(1)}$ and $H^{(2)}$ separately contain local terms that all commute, we can write

$$e^{-i\delta H} = e^{-i\delta H^{(1)}} e^{-i\delta H^{(2)}} + O(\delta^2) \approx_{\delta \to 0} e^{-i\delta H^{(1)}} e^{-i\delta H^{(2)}}.$$

The decomposed evolution operator is sequentially applied to the wave function in the MPS representation. Again, after every application, a *truncation step* (SVD or Schmidt decomposition) is performed to restore the wave function to the MPS format.
Assuming that H is *independent* of time and considering the substitution $\tau = it$

$$\frac{d}{d\tau}\psi(\tau) = -\frac{1}{\hbar}H\psi(\tau),$$

the solution in imaginary time is given by

$$\psi(\tau) = e^{-H\tau/\hbar}\psi(\tau_0), \quad \tau_0 = 0.$$

and the evolution of the initial state $\psi(\tau_0) \in \mathcal{H}$ will tend to the ground state, for time going to infinity. The TEBD can now be applied in combination with imaginary time evolution, in order to find a ground state approximation within the variational class.
Notice that, before the SVD, the state leaves the variational manifold and a representative from the manifold must be found that best approximates the new time-evolved state. In general, the truncation step is not guaranteed to be optimal.
Moreover, as in all sequential methods, either the evolution and re-embedding or the optimization of the state update some but not all the tensors of the network. The drawback of this *local update* in translational invariance systems is the breaking of the symmetry, immediately after a single step.

The Time-Dependent Variational Principle (TDVP) was proposed [HCO+11, HOV13, HLO+16] to overcome these limits, in the framework of translation invariant matrix product states in the thermodynamic limit. The algorithm relies on the concept of tangent space of the uMPS manifold. Moreover, in [ZSVF+18], an alternative scheme has been proposed that applies *global updates* in order to preserve the translation invariance of the MPS. The method combines the DMRG with MPS tangent space concepts.
The TDVP transforms the linear Schrödinger equation $\mathcal{H}$ into a non-linear set of differential equations in the parameter space of the variational manifold. The right-hand side of the Schrödinger equation is projected onto the tangent space of a chosen MPS with fixed bond dimension so that the evolution never leaves the manifold. The TDVP describes the best direction in which the quantum state can evolve without leaving the variational manifold in order to approximate the time-dependent Schrödinger equation. The simulation of time evolution consists in the integration of a set of non-linear coupled differential equations and produces an *approximation to a gradient descent* in the full Hilbert space.

Conjugate gradient methods have been introduced in [PVV11, VHCV16] in order to approximate ground states of translation invariant systems with periodic boundary conditions. Moreover, based on the TDVP, a variational conjugate gradient method has been proposed in [MHO13] and it has been applied to critical quantum field theory. Very recently Riemannian gradient-based optimization has been proved to be a competitive method applied to the tensor network ansatz [HVDH21].

The outlined methods sequentially minimize the functional on a selected subspace of the parameter space of the variety. We review and apply a global minimization method instead, namely the nonlinear conjugate gradient method, with the aim to keep the structure of the variety unchanged, minimizing over all the parameters simultaneously. Moreover, the global approach allows us to exploit further the knowledge on the dimension of the variety.
The first sections of the chapter are meant to describe our framework: the application of the nonlinear conjugate gradient method constrained to the variational class of matrix product state *with open boundary conditions*, in order to approximate the ground state of the AKLT Hamiltonian.
The nonlinear conjugate gradient method is described in detail in Section 5.1. Matrix product states with open boundary conditions, that we denote by $\mathrm{MPS}(\mathbf{m}, \mathbf{n}, d)$ (usually referring to general matrix product states) are defined in Section 5.2. We also study the case of homogeneous matrix product states with open boundary conditions which are subvarieties of matrix product states constructed via site-independent tensors and a boundary condition, c.f. [NV18]. We denote the variety by $\mathrm{hMPS}(m, n, d)$ (usually referring to general homogeneous matrix product states). The general setting is developed for matrix product states with open boundary conditions and then specialized to the homogeneous case.
In Sections 5.3 and 5.4, we review the AKLT model and the matrix product operator representation (MPO) of the AKLT Hamiltonian, respectively. In Section 5.5, we compute the gradient of the functional f (5.1) needed to implement the variational nonlinear conjugate gradient method on the matrix product states class.
The second part of the chapter, consisting of Sections 5.6 and 5.7, introduces a theory that is meant to reduce the number of coordinates of the gradient.
In Section 5.6 we describe a first attempt we made that consists in the selection of a linear subspace of the domain of the matrix product state map of dimension equal to the codimension of the variety. The subspace of the parameter space contains a finite number of points of each fiber of the map. It takes the role of the domain of a reparametrization of the variety, of the minimal dimension, i.e. the dimension of the variety. Even if the attempt led to the implementation of an algorithm with a higher time of convergence compared to the standard algorithm, we report it in the thesis since we believe that it is interesting from a geometric point of view.
In Section 5.7, we study in more detail the fiber of the matrix product state map and its tangent space. In the case of matrix product states with open boundary conditions, our results in Chapter 3 and the results in [HMOV14], imply that the fiber of the map in a given point coincides exactly with the gauge orbit of the point. Therefore the domain of the map admits a natural pointwise decomposition in the tangent space to the gauge orbit and its complementary vector space.

5.1 Nonlinear conjugate gradient

In this section, we describe the linear conjugate gradient method and the nonlinear adaptation, based on [NW06] which we refer to for a detailed treatment of the topic. One of the routines that the method invokes is the *line search*, c.f. Algorithm 3, that is the most expensive part of the algorithm in terms of time.

Linear Conjugate Gradient method (CG). Consider the Hilbert space $\mathcal{H} = \mathbb{C}^q$. The CG method is an iterative method for solving a linear system of equations

$$Hx = b,$$

where $x \in \mathbb{C}^q$, $H \in \mathbb{C}^{q\times q}$ Hermitian and positive-definite, and $b \in \mathbb{C}^q$.

If $f(x) := \frac{1}{2}x^\dagger Hx - x^\dagger b + c$, with $c \in \mathbb{C}$, the problem is equivalent to the following minimization problem

$$\min_x f(x).$$

The gradient of f is

$$\nabla f(x) = Hx - b,$$

and it coincides with the residual of the linear system $Hx - b$, i.e. if $x = x_k$, then the residual is defined by $r_k = Hx_k - b$ and it coincides with $\nabla f(x_k)$.

Two vectors $x, y \in \mathbb{C}^q$ are said to be *conjugate* with respect to H if and only if $x^\dagger Hy = 0$. Any two such vectors are moreover linearly independent.

Consider a starting point $x_0 \in \mathbb{C}^q$ and a basis of $\mathbb{C}^q$ given by a conjugate set $\{p_0, \dots, p_{q-1}\}$, i.e. $p_i H p_j = \delta_{ij}$. Define

$$x_{k+1} = x_k + \alpha_k p_k,$$

where $\alpha_k = \arg\min_\alpha f(x_k + \alpha p_k)$. The coefficient $\alpha_k \in \mathbb{R}$ is the one-dimensional minimizer of the quadratic function f along $x_k + \alpha p_k$; it is analytically computed by solving $\frac{\partial}{\partial_\alpha} f(x_k + \alpha p_k) = 0$ and it is explicitly given by

$$\alpha_k = -\frac{r_k^\dagger p_k}{p_k^\dagger H p_k}. \tag{5.2}$$

Moreover, each direction p_k can be chosen to be a linear combination of $-r_k$, the deepest descent direction for f, and only the previous one p_{k-1}

$$p_k = -r_k + \beta_k p_{k-1},$$

if β_k is such that $p_k^\dagger H p_{k-1} = 0$, that is given by

$$\beta_k = \frac{p_{k-1}^\dagger H p_k}{p_{k-1}^\dagger H p_{k-1}}.$$

Remark 5.1.1. For a general quadratic function successive minimization along coordinate axes may not find the solution in q iterations. The importance of conjugacy lies in the fact that f is minimized in exactly q steps by successively minimizing it along the directions in the conjugate set.

Nonlinear Conjugate Gradient method (NLCG). A modified version of the linear conjugate gradient method was proposed by Fletcher and Reeves [FR64], in order to extend the method to general nonlinear functions, $f : \mathcal{H} \to \mathbb{R}$, approximately quadratic near the stationary points. The changes to the CG algorithm are the following.

1. In the CG method, the one-dimensional minimizer $\alpha_k \in \mathbb{R}$ such that

   $$\alpha_k = \arg\min_{\alpha} f(x_k + \alpha p_k),$$

   has the analytic solution given in Equation (5.2). In the NLCG method instead an approximate line search is performed, which consists in an iterative method that finds an approximate minimum of f in direction p_k.

2. The residual r is replaced by the gradient of the nonlinear function f

   $$r_k = \nabla f(x_k),$$

   and β is defined as

   $$\beta_{k+1} = \beta_{k+1}^{\mathrm{FR}} := \frac{\nabla f(x_{k+1})^\dagger \nabla f(x_{k+1})}{\nabla f(x_k)^\dagger \nabla f(x_k)}.$$

3. After q iterations, $\beta_{q+1}^{\mathrm{FR}} = 0$ (and $p_{q+1} = -\nabla f(x_{q+1})$), i.e. there is a *restart* of the method after every q steps.

The NLCG method is given in Algorithm 2. We fix in input the tolerance $t \in \mathbb{R}$ that we want to be reached by the norm of gradient.

Line search. Consider the one-dimensional minimization problem

$$\alpha_k = \arg\min_{\alpha \in \mathbb{R}} f(x_k + \alpha p_k).$$

An approximate line search is an iterative method that determines a step length $\alpha_k \in \mathbb{R}$ such that $f(x_k + \alpha_k p_k)$ is a good approximation of $\min_{\alpha \in \mathbb{R}} f(x_k + \alpha p_k)$. The approximate line search we implement satisfies the *strong Wolfe conditions*

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k \nabla f(x_k)^\dagger p_k, \tag{5.3}$$

$$|\nabla f(x_k + \alpha_k p_k)^\dagger p_k| \leq -c_2 \nabla f(x_k)^\dagger p_k, \tag{5.4}$$

where $0 < c_1 < c_2 < \frac{1}{2}$. Typically, for the NLCG method, $c_1 = 0.1$ and $c_2 = 0.0001$ [NW06]. Condition (5.3) is called *sufficient decrease condition* and it imposes a reduction

Algorithm 2: NLCG

Input: function f; initial point $x_0 \in \mathbb{C}^q$; tolerance $t \in \mathbb{R}$.
Output: $x^* = x_{k+1}$ approximation of $\arg\min_{x \in \mathbb{C}^q} f(x)$.

Evaluate $f_0 = f(x_0)$;
$p_0 \leftarrow -\nabla f(x_0)$;
$k \leftarrow 0$;
while $|\nabla f(x_k)| > t$ **do**
  $\boxed{\alpha_k = \arg\min_\alpha f(x_k + \alpha p_k)}$;
  Set $x_{k+1} = x_k + \alpha_k p_k$;
  Evaluate $\nabla f(x_{k+1})$;
  if $\boxed{k = q}$ **then**
    $\beta_{k+1}^{\mathrm{FR}} \leftarrow 0$;
    $k \leftarrow 0$;
  else
    $\boxed{\beta_{k+1}^{\mathrm{FR}} \leftarrow \frac{\nabla f(x_{k+1})^\dagger \nabla f(x_{k+1})}{\nabla f(x_k)^\dagger \nabla f(x_k)}}$;
  end
  $p_{k+1} \leftarrow -\nabla f(x_{k+1}) + \beta_{k+1}^{\mathrm{FR}} p_k$;
  $k \leftarrow k + 1$;
end

of f proportional to both the step length α_k and the directional derivative $\nabla f(x_k)^\dagger p_k$. It therefore imposes a sufficient decrease of f. Condition (5.4) is called *curvature condition* and it imposes to $\nabla f(x_k + \alpha_k p_k)^\dagger p_k$, called *slope*, to be greater than the initial slope $\nabla f(x_k)^\dagger p_k$. It ensures to avoid too short steps, moreover it does not allows the gradient to be too positive and this excludes from the search points that are far to be stationary for f. Moreover, consider the hermitian product

$$\nabla f(x_k)^\dagger p_k = -|\nabla f(x_k)|^2 + \beta_k^{\mathrm{FR}} \nabla f(x_k)^\dagger p_{k-1}. \tag{5.5}$$

In an exact line search $\nabla f(x_k)^\dagger p_{k-1} = 0$. In an approximate line search instead it can occur $\nabla f(x_k)^\dagger p_{k-1} < 0$ if the right hand part of (5.5) is dominated by the second term. In this case p_k is not a descent direction. The curvature condition (5.4) ensures that p_k is always a descent direction.

The line search method, whose pseudo-code is given in Algorithm 3, begins with a trial step length α_1, and keeps increasing it until it finds either an acceptable step length, i.e. a step length that satisfies conditions (5.3) and (5.4), or an interval that contains the acceptable step length. In the latter case, it is invoked the so called **zoom** function, given in Algorithm 4, which successively decreases the size of the interval until an acceptable step length is reached.

We always refer to [NW06], in particular to Section 3.5 for a detailed discussion on the line search method and the zoom function which the line search invokes and to Section 5.2 for their use in the NLCG.

We define

$$\varphi(\alpha) := f(x_k + \alpha p_k).$$

The derivative with respect to α is therefore

$$\varphi'(\alpha) = \nabla f(x_k + \alpha p_k)^\dagger p_k.$$

Notice that both routines require several evaluations of the function and computations of gradients in order to find an acceptable step length.

Algorithm 3: Line search with strong Wolfe conditions

Input: Initial point $x \in \mathbb{C}^q$, direction $p \in \mathbb{C}^q$.

Output: $\alpha_* \sim \arg\min_{\alpha\in\mathbb{R}} \varphi(\alpha)$ with $\varphi(\alpha) := f(x + \alpha p)$, i.e. step length satisfying strong Wolfe conditions.

Set the lower step length $\alpha_0 \leftarrow 0$;
Choose upper step length $\alpha_{\max} > 0$;
Choose an initial step length $\alpha_1 \in (0, \alpha_{\max})$;
$i \leftarrow 1$;
while 1 **do**
  Evaluate $\varphi(\alpha_i)$;
  if $\varphi(\alpha_i) > \varphi(0) + c_1\alpha_i\varphi'(0)$ *or* $(\varphi(\alpha_i) \geq \varphi(\alpha_{i-1})$ *and* $i > 1)$ **then**
    $\alpha_* \leftarrow$ **zoom**(α_{i-1}, α_i) and **stop**;
  end
  Evaluate $\varphi'(\alpha_i)$;
  if $|\varphi'(\alpha_i)| \leq -c_2\varphi'(0)$ **then**
    set $\alpha_* \leftarrow \alpha_i$ and **stop;**
  end
  if $\varphi'(\alpha_i) \geq 0$ **then**
    set $\alpha_* \leftarrow$ **zoom**(α_i, α_{i-1}) and **stop;**
  end
  Choose $\alpha_{i+1} \in (\alpha_i, \alpha_{\max})$: $\alpha_0 = \alpha_1$, $\alpha_1 = \min(\alpha_{\max}, 3\alpha_1)$;
  $i \leftarrow i + 1$;
end

Algorithm 4: Zoom function

Input: interval of step length $(\alpha_{\text{lo}}, \alpha_{\text{hi}})$

Output: $\alpha_* \sim \arg\min_{\alpha\in\mathbb{R}} \varphi(\alpha)$ with $\varphi(\alpha) := f(x + \alpha p)$, i.e. step length satisfying strong Wolfe conditions.

Fix a bound on the number of iterations $\mathrm{t} \in \mathbb{N}$;

for $i \leftarrow 1$ **to** t **do**

  $\alpha_j = \frac{\alpha_{\text{lo}}+\alpha_{\text{hi}}}{2}$;

  Evaluate $\varphi(\alpha_j)$;

  if $\varphi(\alpha_j) > \varphi(0) + c_1\alpha_j\varphi'(0)$ **or** $\varphi(\alpha_j) \geq \varphi(\alpha_{lo})$ **then**

    $\alpha_{\text{lo}} \leftarrow \alpha_j$;

  else

    Evaluate $\varphi'(\alpha_j)$;

    if $|\varphi'(\alpha_j)| \leq -c_2\varphi'(0)$ **then**

      Set $\alpha_* \leftarrow \alpha_j$ and **stop**

    end

    if $\varphi'(\alpha_j)(\alpha_{hi} - \alpha_{lo}) \geq 0$ **then**

      $\alpha_{\text{hi}} \leftarrow \alpha_{\text{lo}}$

    end

    $\alpha_{\text{lo}} \leftarrow \alpha_j$;

  end

end

5.2 MPS with open boundary conditions

Matrix product states have been introduced in Chapter 2, Subsection 2.1.2 as an example of tensor network variety and their dimension has been studied in Chapter 3 for particular ranges of the parameters. We give here the well-known basis-dependent definition, already obtained in Subsection 2.1.2, then we define the subvariety of matrix product states with open boundary conditions and the homogeneous matrix product states with open boundary conditions, that we are going to use as variational classes.

Let $C_d = (\mathbf{v}(C_d), \mathbf{e}(C_d))$ be the cyclic graph on d vertices, $\mathbf{v}(C_d) = \{1, \ldots, d\}$ and $\mathbf{e}(C_d) = \{e_1, \ldots, e_d\}$ with $e_i = \{i, i+1\}$. Let $\mathbf{m} = (m_1, \ldots, m_d)$ the set of bond dimensions: each m_i is associated to e_i. Let $\mathbf{n} = (n_1, \ldots, n_d)$ be the sets of local dimensions.

Definition 5.2.1. The parametrization associated to the tensor network $(C_d, \mathbf{n}, \mathbf{m})$ is given by

$$\begin{aligned} \overline{\phi} : \bigtimes_{k=1}^{d} \mathbb{C}^{m_{k-1}\times n_k \times m_k} &\to V_1 \otimes \cdots \otimes V_d \\ (A_1, \ldots, A_d) &\mapsto \sum_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d} \mathrm{Tr}\left(A_1^{i_1} \cdots A_d^{i_d}\right) e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_d}^{(d)}, \end{aligned} \tag{5.6}$$

where $\{e_{i_j}^{(j)} : i_j = 1, \ldots, n_j\}$ is the canonical basis of V_j. The variety of matrix product states is either the Zariski or the Euclidean closure of the image of the map $\overline{\phi}$. According to Chapters 2 and 3, we denote the variety by $\mathcal{TNS}_{\mathbf{m},\mathbf{n}}^{C_d}$.

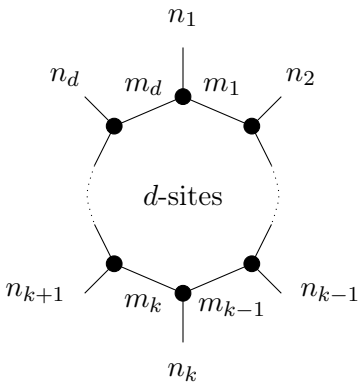


Figure 5.2: Graphical representation of a tensor in $\mathcal{TNS}_{\mathbf{m},\mathbf{n}}^{C_d}$. There are total d tensors involved with order $m_{i-1} \times m_i \times n_i$, for $i = 1, \ldots, d$ and $m_0 := m_d$.

Matrix product states with open boundary conditions are a particular class of matrix product states. Consider the tensor network $(C_d, \mathbf{n}, \mathbf{m})$ given in Definition 5.2.1. Assume that $e_d = \{d, 1\}$ has associated bond dimension $m_d = 1$. By Remark 2.1.3, the edge e_d can be removed, c.f. Figure 5.3. We obtain the following tensor network: $(P_d, \widetilde{\mathbf{m}} = (m_1, \ldots, m_{d-1}), \mathbf{n})$, where P_d denotes the path graph with d vertices.

Definition 5.2.2. The parametrization associated to the tensor network $(P_d, \widetilde{\mathbf{m}}, \mathbf{n})$, $(m_d = m_0 = 1)$, is given by

$$\begin{aligned} \phi : \bigtimes_{k=1}^{d} \mathbb{C}^{n_k \times m_{k-1} \times m_k} &\to V_1 \otimes \cdots \otimes V_d \\ (A_1, \ldots, A_d) &\mapsto \sum_{i_1, \ldots, i_d = 1}^{n_1, \ldots, n_d} A_1^{i_1} \cdots A_d^{i_d} \; e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_d}^{(d)}, \end{aligned} \tag{5.7}$$

where $\{e_{i_j}^{(j)} : i_j = 1, \ldots, n_j\}$ is the canonical basis of V_j. The variety of matrix product states *with open boundary conditions* is either the Euclidean or the Zariski closure of the image of the map ϕ. We denote matrix product states with open boundary conditions by $\mathrm{MPS}(\mathbf{m}, \mathbf{n}, d)$ (usually referring to general matrix product states).

We denote the domain of ϕ by $D_{\mathrm{MPS}} = \bigtimes_{k=1}^{d} \mathbb{C}^{m_{k-1} \times n_k \times m_k}$ (with $m_0 = m_d = 1$) and its dimension by $N_{\mathrm{MPS}} = \dim D_{\mathrm{MPS}}$.

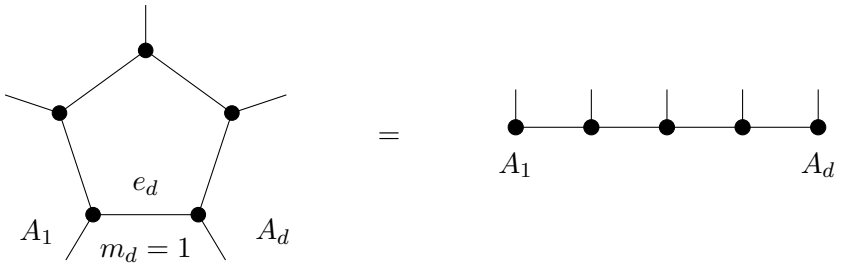


Figure 5.3: Underlying graph associated to matrix product states with *open boundary conditions*. As an example, we show the cyclic graph (left) $\Gamma = C_d$ with $d = 5$ vertices.Tensors A_1 and A_d, associated to vertices $1, d \in \mathbf{v}(\Gamma)$ respectively, are connected by the edge e_d of weight $m_d = 1$. The graph obtained removing the edge e_d is the path graph with $d = 5$ vertices (right). The tensor network varieties associated to the two graphs are equivalent: matrix product states with open boundary conditions.

Remark 5.2.3. In the case of matrix product states with open boundary conditions, a complete result regarding the dimension of the tensor network variety is given in [HMOV14, Thm 14]. In the language of Theorem 3.3.2, the result of [HMOV14] is (formally setting $m_d := 1 =: m_0$)

$$\dim \mathrm{MPS}(\mathbf{m}, \mathbf{n}, d) = \sum_{i=1}^{d} n_i m_{i-1} m_i - \sum_{i=1}^{d-1} m_i^2, \tag{5.8}$$

that coincides with the expected value of the dimension given in Equation (3.3)

$$\dim \mathrm{MPS}(\mathbf{m}, \mathbf{n}, d) = \left((n_1 m_1) + (n_d m_{d-1}) + \sum_{i=2}^{d-1} n_i m_{i-1} m_i \right) - d + 1 - \sum_{i=1}^{d-1} (m_i^2 - 1).$$

Therefore, for matrix product states with open boundary conditions, the generic fiber of the map ϕ (5.7) is *isomorphic* to the orbit of a generic element in the fiber, under the action of the gauge subgroup.

Example 5.2.4. Fix all bond dimensions equal to $m \geq 2$ and all local dimensions equal to $n \geq 2$, then

$$\begin{aligned} \dim \mathrm{MPS}(m, n, d) &= 2(mn - 1) + (d - 2)(m^2 n - 1) + 1 - (d - 1)(m^2 - 1) \\ &= 2mn + (d - 2)m^2 n - (d - 1)m^2. \end{aligned}$$

Remark 5.2.5 (Vector reshape)**.** The element $A = (A_1, \ldots, A_d) \in D_{\mathrm{MPS}}$ is a collection of d tensors of order 3. Fixed $k \in \{1, \ldots, d\}$, the k-th tensor $A_k \in \mathbb{C}^{n_k \times m_{k-1} \times m_k} \simeq \mathbb{C}^{n_k} \otimes \mathbb{C}^{m_{k-1}} \otimes \mathbb{C}^{m_k}$ has coordinates $A_k = (a_k)^{i_k}_{\gamma_k \delta_k}$. Therefore the coordinates of D_{MPS} are the collection of the entries of all the A_k, for $k = 1, \ldots, d$:

$$\{ a_k{}^{i_k}_{\gamma_k \delta_k} \ : \ k = 1, \ldots, d, i_k = 1, \ldots, n_k, \gamma_k = 1, \ldots, m_{k-1}, \delta_k = 1, \ldots, m_k, \}.$$

We fix once for all the linear map

$$\begin{aligned} L_k : \mathbb{C}^{n_k} \otimes \mathbb{C}^{m_{k-1}} \otimes \mathbb{C}^{m_k} &\to \mathbb{C}^{n_k m_{k-1} m_k} \\ e_{i_k} \otimes e_{\gamma_k} \otimes e_{\delta_k} &\mapsto e_{l_k(i_k,\gamma_k,\delta_k)}, \end{aligned}$$

with $l_k : [n_k] \times [m_{k-1}] \times [m_k] \to [n_k m_{k-1} m_k]$ bijection of indices. Our choice of the bijection l_k is such that $L_k(A_k)$ has coordinates

$$L_k(A_k) = \left(a_{k\,11}^{\;\;1}, \ldots, a_{k\,1m_k}^{\;\;1}, \ldots, a_{k\,m_{k-1}m_k}^{\;\;1}, \ldots, a_{k\,m_{k-1}m_k}^{\;\;n_k} \right).$$

We give a pictorial representation in Figure 5.4.

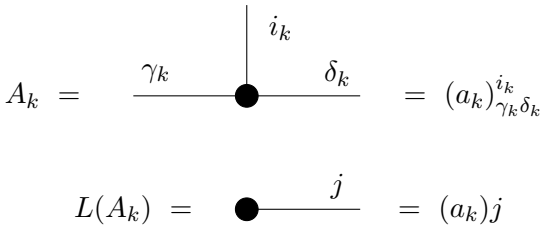


Figure 5.4: Top: graphical representation of a 3-order tensor. Bottom: tensor A_k is reshaped into a vector with coordinate $(a_{kj})_{j=1}^{n_k m_{k-1} m_k}$. The index $j = l_k(i_k, \gamma_k, \delta_k)$ is such that $L_k(A_k)$ has coordinates $(a_{k1}, \ldots, a_{km_{k-1}m_k n_k}) = (a_{k\,11}^{\;\;1}, \ldots, a_{k\,m_{k-1}m_k}^{\;\;n_k})$.

Given $A = (A_1, \ldots, A_d) \in D_{\mathrm{MPS}}$, the natural extension to D_{MPS} is given by

$$\begin{aligned} L : \bigtimes_{k=1}^{d} \mathbb{C}^{n_k} \otimes \mathbb{C}^{m_{k-1}} \otimes \mathbb{C}^{m_k} &\to \mathbb{C}^{N_{\mathrm{MPS}}} \\ \bigtimes_{k=1}^{d} e_{i_k} \otimes e_{\gamma_k} \otimes e_{\delta_k} &\mapsto e_{l(k,i_k,\gamma_k,\delta_k)}, \end{aligned}$$

with bijection l such that the N_{MPS} coordinates of $L(A) = L(A_1, \ldots, A_d)$ are

$$(a_j)_{j=1,\ldots,N_{\mathrm{MPS}}} = (\underbrace{a_{1\,11}^{\;\;1}, a_{1\,12}^{\;\;1}, \ldots, a_{1\,m_0m_1}^{\;\;n_1}}_{L_1(A_1)}, \underbrace{a_{2\,11}^{\;\;1}, \ldots, a_{2\,m_1m_2}^{\;\;n_2}}_{L_2(A_2)}, \ldots, \underbrace{a_{d\,11}^{\;\;1}, \ldots, a_{d\,m_{d-1}m_d}^{\;\;n_d}}_{L_d(A_d)}).$$

With abuse of notation, we denote $l_k(i_k, \gamma_k, \delta_k)$ and $l(k, i_k, m_{k-1}, m_k, n_k)$ simply by l_k and l respectively.

On the other hand, it is sometimes useful to think about the tensor $A_k \in \mathbb{C}^{n_k \times m_{k-1} \times m_k}$ as a collection of n_k matrices in $\mathbb{C}^{m_{k-1} \times m_k}$. Therefore the tensor A_k is identified with the collection of matrices $(A_k^1, \ldots, A_k^{n_k})$, with $A_k^i = (a_k^i)_{\gamma_k \delta_k}$.

5.2.1 Homogeneous MPS with open boundary conditions

Homogeneous matrix product states are matrix product states constructed via site-independent tensors and a boundary condition, c.f. [NV18]. Fix $m_i = m$, $n_i = n$ for every $i = 1, \dots, d$ and one tensor $A \in \mathbb{C}^{n \times m \times m}$. Fixed the canonical basis, we identify the tensor A with a tuple of n matrices $A_j \in \mathbb{C}^{m \times m}$ for $1 \leq j \leq n$. Denote $A = (A_1, \dots, A_n) \in (\mathbb{C}^{m \times m})^{\times n}$ and let $\omega \in \mathbb{C}^m$.

Definition 5.2.6. The Zariski or Euclidian closure of the image of the map

$$\begin{aligned} \overline{\psi} : \mathbb{C}^{m \times m} \times (\mathbb{C}^{m \times m})^{\times n} &\to (\mathbb{C}^n)^{\otimes d} \\ p = (\omega, A) &\mapsto \sum_{i_1, \dots, i_d = 1}^{n} \operatorname{Tr}\Big(\omega\ A_{i_1} \cdots A_{i_d}\Big) e_{i_1} \otimes \cdots \otimes e_{i_d}, \end{aligned}$$

is the variety of *homogeneous* matrix product states.

Notice that uniform matrix product states, c.f. Definition 4.1.1, that we have studied in Chapter 4 can be seen as subvarieties of homogeneous matrix product states, imposing the condition $\omega = \mathrm{Id}_m$. Here we are interested instead in the subvariety determined by the condition $\mathrm{rk}(\omega) = 1$.

Definition 5.2.7. Let $\omega = v_L^* \otimes v_R$ be the rank one decomposition of the matrix $\omega \in \mathbb{C}^{m \times m}$. Under this assumption parametrization given in Definition 5.2.6 becomes

$$\begin{aligned} \psi : \mathbb{C}^m \times (\mathbb{C}^{m \times m})^{\times n} \times \mathbb{C}^m &\to (\mathbb{C}^n)^{\otimes d} \\ p = (v_L, A, v_R) &\mapsto \sum_{i_1, \dots, i_d = 1}^{n} \Big(v_L^\dagger A_{i_1} \cdots A_{i_d} v_R\Big) e_{i_1} \otimes \cdots \otimes e_{i_d}, \end{aligned} \tag{5.9}$$

The associated variety is the homogeneous matrix product states *with open boundary conditions*. Again with abuse of notation, the variety is denoted by $\mathrm{hMPS}(m, n, d)$ (usually referring to general homogeneous matrix product states).

Remark 5.2.8. Homogeneous matrix product states with open boundary conditions are in particular matrix product states with open boundary conditions associated to the path graph P_{d+2} with $d + 2$ vertices, see Figure 5.5 for a pictorial representation. In this Chapter, every computation done for matrix product states with open boundary conditions naturally holds for homogeneous matrix product states with open boundary conditions.

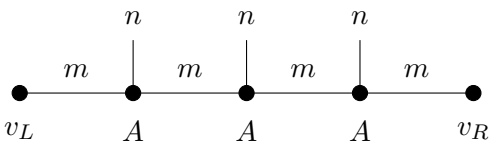


Figure 5.5: Example of a homogeneous matrix product state with open boundary conditions with bond dimensions equal to m and local dimensions equal to n. Each inner vertex is associated with the same tensor $A \in (\mathbb{C}^{m \times m})^{\times n}$. The boundaries are vectors.

We must be careful with homogeneous matrix product states with open boundary conditions. They are clearly subvarieties of matrix product states with open boundary conditions. On the other hand, they are subvarieties of homogeneous matrix product states and their dimension is site-independent. Uniform matrix product states are subvarieties of homogeneous matrix product states as well, but there are no containment relations with homogeneous matrix product states with open boundary conditions. In conclusion, we cannot use either our results on the dimension of general matrix product states with open boundary conditions or the results on general uniform matrix product states. However, from a simple count of parameters, analogous to that of our main result of Chapter 3, the dimension of homogeneous matrix product states with open boundary conditions, with $n > m \geq 2$, can be bounded from above by

$$\begin{aligned} \dim \mathrm{hMPS}(m,n,d) &\leq 2(m-1) + (m^2 n - 1) + 1 - (m^2 - 1) \\ &= 2m + m^2(n-1) - 1, \end{aligned} \tag{5.10}$$

which is the count of parameters of the domain of the parametrization minus the dimension of the gauge subgroup. Indeed, in this case, we can say that the stabilizer of the gauge subgroup is trivial since the stabilizer of n matrices in $\mathbb{C}^{m\times m}$ is trivial for $n > m \geq 2$.

5.3 AKLT Model

We describe in detail the AKLT model, following the paper by I. Affleck, T.Kennedy, E.H. Lieb and H. Tasaki [AKLT88]. We are interested in their results since ground states of the AKLT Hamiltonian admit an analytic solution that is exactly a matrix product state representation.

Consider a quantum system consisting of a chain of d spin 1 particles. The composite system of d spin 1 particles is associated to a state space $\mathcal{H} = \bigotimes_{j=1}^{d} \mathcal{H}_j$, with $\mathcal{H}_j = \mathbb{C}^3$ the state space associated to the j-th spin 1 particle, c.f. Section 1.2.

Figure 5.6: Chain of spin 1 particles, represented as vertices of a graph. The vertices are not linked by edges, meaning that the bond dimensions between vertices are equal to 1 and the tensor network variety associated to the graph is $\mathcal{H} = \bigotimes_{j=1}^{d} \mathcal{H}_j$, with $\mathcal{H}_j = \mathbb{C}^3$. The local dimension associated to each vertex is $\dim \mathcal{H}_j = 3$.

Consider the spin 1 operators whose matrix representation in the canonical basis is given in Equation (1.5) and that we recall here

$$S^1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, S^2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, S^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

and denote $\vec{S} = (S^1, S^2, S^3) \in (\mathbb{C}^{3\times 3})^{\times 3}$. The action of S^α, $\alpha \in \{1,2,3\}$, on the site j of the chain is given by

$$S_j^\alpha := \mathrm{Id}_3^{\boxtimes(j-1)} \boxtimes S^\alpha \boxtimes \mathrm{Id}_3^{(d-j)} \in \mathbb{C}^{3^d\times 3^d},$$

where $\boxtimes$ denotes the Kronecker product given in Definition 1.1.2. We denote the scalar product of vectors of matrices by

$$\vec{S}_j \cdot \vec{S}_{j+1} := \sum_{\alpha=1}^{3} S_j^\alpha S_{j+1}^\alpha = \sum_{\alpha=1}^{3} (\mathrm{Id}^{\boxtimes(j-1)} \boxtimes S^\alpha \boxtimes S^\alpha \boxtimes \mathrm{Id}^{(d-j-1)}) \in \mathbb{C}^{3^d\times 3^d}. \tag{5.11}$$

The AKLT Hamiltonian is the operator $H : \mathcal{H} \to \mathcal{H}$ whose matrix representation is given by

$$H = \sum_{i=1}^{d-1} \left(\vec{S}_j \cdot \vec{S}_{j+1} + \frac{1}{3}(\vec{S}_j \cdot \vec{S}_{j+1})^2 \right) \in \mathbb{C}^{3^d\times 3^d} \simeq \mathrm{End}(\mathcal{H}). \tag{5.12}$$

More compactly, we define the matrix

$$M = S_1^1 S_2^1 + S_1^2 S_2^2 + S_1^3 S_2^3 = S^1 \boxtimes S^1 + S^2 \boxtimes S^2 + S^3 \boxtimes S^3, \tag{5.13}$$

and the local operator acting on two sites by

$$h := M + \frac{1}{3} M^2 \in \mathbb{C}^{9\times 9}.$$

Then, the extension of $h \in \mathbb{C}^{9\times 9}$ to the whole space $\mathbb{C}^{3^d\times 3^d}$ is

$$h_j = \mathrm{Id}^{\boxtimes(j-1)} \boxtimes h \boxtimes \mathrm{Id}^{(d-j-1)} \in \mathbb{C}^{3^d\times 3^d}, \tag{5.14}$$

and the Hamiltonian (5.12) can be written as a sum of local terms

$$H = \sum_{i=1}^{d-1} h_j.$$

Construction of Hamiltonian and ground states. We summarize the steps which bring to the construction of the AKLT Hamiltonian (5.12) and its ground states. Recall that $V^{(m)} \simeq \mathbb{C}^{m+1}$ corresponds to spin $\frac{m}{2}$, see Remark 1.2.16.

1. Consider the spin 1 chain. Each couple of adjacent spin 1 sites, j and $j+1$, admits a decomposition into irreducible $\mathfrak{sl}_2$-representations

   $$\mathbb{C}^3 \otimes \mathbb{C}^3 \simeq V^{(0)} \oplus V^{(2)} \oplus V^{(4)} \simeq \mathbb{C}^1 \oplus \mathbb{C}^3 \oplus \mathbb{C}^5, \tag{5.15}$$

   where the first $\simeq$ is an isomorphism of $\mathfrak{sl}_2$ representations, given by Theorem 1.2.15, and the second $\simeq$ is an isomorphism of vector spaces. This means that $\mathbb{C}^3 \otimes \mathbb{C}^3$ decomposes into spin 0, 1 and 2.

2. In [AKLT88], they consider a local Hamiltonian $\widetilde{H}$, i.e. defined as the sum over $j = 1, \dots, d-1$ of local operators $\widetilde{h}_j : \mathcal{H} \to \mathcal{H}$, acting on nearby pairs of spins. The local operator is explicitly expressed in terms of the spin operators

$$\widetilde{h}_j := \frac{1}{2}\vec{S}_j \cdot \vec{S}_{j+1} + \frac{(\vec{S}_j \cdot \vec{S}_{j+1})^2}{6} + \frac{1}{3}\mathrm{Id}_{3^d}.$$

Thus in particular

$$\widetilde{H} := \sum_{j=1}^{d-1} \widetilde{h}_j = \sum_{j=1}^{d-1} \left(\frac{1}{2}\vec{S}_j \cdot \vec{S}_{j+1} + \frac{1}{6}(\vec{S}_j \cdot \vec{S}_{j+1})^2 + \frac{1}{3}\mathrm{Id}_{3^d} \right).$$

The local operator $\widetilde{h}_j$ is proved to be the projector on spin 2, i.e. $V^{(4)} \simeq \mathbb{C}^5$ in Decomposition (5.15). In particular $\ker(\widetilde{h}_j) = V^{(0)} \oplus V^{(2)} \simeq \mathbb{C}^1 \oplus \mathbb{C}^3$.

3. The Hamiltonian $\widetilde{H} = \sum_{j=1}^{d-1} \widetilde{h}_j$, as a sum of projectors, is positive definite therefore its lowest eigenvalue is $\lambda_0 = 0$ and its ground states are elements of its kernel. The kernel of $\widetilde{H}$ is constructed from the kernel of the local operator $\widetilde{h}_j$. A basis of $\ker(\widetilde{h}_j) \simeq V^{(0)} \oplus V^{(2)}$ is explicitly determined, for every $j = 1, \dots, d-1$ and a 4-dimensional family of degenerate ground states of $\widetilde{H}$ is built.

4. Finally, notice that h_j given in Equation (5.14) is

$$h_j = 2\widetilde{h}_j - \frac{2}{3}\mathrm{Id}_{3^d} = \vec{S}_j \cdot \vec{S}_{j+1} + \frac{1}{3}(\vec{S}_j \cdot \vec{S}_{j+1})^2$$

and Hamiltonian (5.12) is the shifted operator

$$H = \sum_{j=1}^{d-1} h_j = \sum_{j=1}^{d-1} \left(2\widetilde{h}_j - \frac{2}{3}\mathrm{Id}_{3^d} \right) = 2\sum_{j=1}^{d-1} \widetilde{h}_j - \frac{2}{3}\sum_{j=1}^{d-1} \mathrm{Id}_{3^d} = 2\widetilde{H} - \frac{2}{3}(d-1)\mathrm{Id}_{3^d}.$$

Therefore, the 4-dimensional family of degenerate ground states of $\widetilde{H}$ is a 4-dimensional family of degenerate ground states of H, with (shifted) lowest eigenvalue $\lambda_{0,d} = 0 - \frac{2}{3}(d-1)$.

Local operator and Hamiltonian. Assume $d = 2$. We study the local Hamiltonian between sites $(1, 2)$ of the chain

$$\begin{aligned} \widetilde{h} := \widetilde{h}_1 &= \frac{1}{2}\left(\vec{S}_1 \cdot \vec{S}_2 + \frac{1}{3}(\vec{S}_1 \cdot \vec{S}_2)^2 \right) + \frac{1}{3}\mathrm{Id}_9 \\ &= \frac{1}{2}\left(M + \frac{1}{3}M^2 \right) + \frac{1}{3}\mathrm{Id}_9 \in \mathbb{C}^{9\times 9}, \end{aligned} \tag{5.16}$$

where $M \in \mathbb{C}^{9\times 9}$ is given in Equation (5.13). We show that it is the projector on $V^{(4)} \simeq \mathbb{C}^5$ [AKLT88].

The characteristic polynomial of $M \in \mathbb{C}^{9\times 9}$ (5.13) is $p_M(x) = (x+2)(x+1)^3(x-1)^5$ and gives $D = \mathrm{diag}(-2,-1,-1,-1,1,1,1,1,1)$, the diagonal matrix of eigenvalues of M.

Let λ be an eigenvalue of M and E_λ the associated eigenspace generated by the collection of orthonormal vectors $\{v_\lambda^k, k = 1,\dots,d_\lambda\}$, of dimension d_λ. Denote by $\mathcal{B} = \{v_\lambda^j : \lambda \in \{-2,-1,1\}, j = 1,\dots,d_\lambda\}$ the basis of $\mathbb{C}^9$ given by the collection of eigenvectors.

The projector on E_λ is

$$P_{\{\lambda\}} = \sum_{i=1}^{d_\lambda} v_\lambda^i {v_\lambda^i}^\dagger.$$

We denote $P_0 := P_{\{-2\}}$, $P_1 := P_{\{-1\}}$ and $P_2 := P_{\{1\}}$, since these are actually the projectors on $V^{(0)}, V^{(2)}$ and $V^{(4)}$, i.e. on spin 0, spin 1 and spin 2 respectively. Then

$$M = -2P_0 - P_1 + P_2.$$

Substituting M in (5.16), the local Hamiltonian diagonalizes in the $\mathcal{B}$ basis as

$$\begin{aligned}\widetilde{h} &= \frac{1}{2}\left(-2P_0 - P_1 + P_2 + \frac{1}{3}(-2P_0 - P_1 + P_2)^2\right) + \frac{1}{3}\mathrm{Id}_9 \\ &= \frac{1}{2}\left(-\frac{2}{3}P_0 - \frac{2}{3}P_1 + \frac{4}{3}P_2\right) + \frac{1}{3}(P_0 + P_1 + P_2) = P_2.\end{aligned}$$

The operator $\widetilde{h}$ has eigenvalues $\lambda_1 = 1$ of degeneration 5 and $\lambda_0 = 0$ of degeneration 4. The latter is the lowest eigenvalue of $\widetilde{h}$ and its eigenspace is given by $\ker(P_2) = E_{-2} \oplus E_{-1} = V^{(0)} \oplus V^{(2)}$.

For $d \geq 2$, $\widetilde{h}_j$ is simply the Kronecker product of P_2 and $d-2$ copies of the identity matrix I_3:

$$\widetilde{h}_j = (\mathrm{Id}^{\boxtimes(j-1)} \boxtimes \widetilde{h} \boxtimes \mathrm{Id}^{(d-j-1)}) \in \mathbb{C}^{3^d\times 3^d}.$$

The Hamiltonian $\widetilde{H} = \sum_{j=1}^{d-1} \widetilde{h}_j$ is a sum of projectors and it is therefore *positive definite*: the lowest eigenvalue of $\widetilde{H}$ is $\lambda_0 = 0$. Since $d(\lambda_0) = 4$, the associated eigenspace $E_{\lambda_0} \subseteq (\mathbb{C}^3)^{\otimes d}$ has dimension 4. It corresponds to the kernel of $\widetilde{H}$ and it is the space of ground states.

Since Hamiltonian (5.12) is the shifted operator

$$H = 2\widetilde{H} - \frac{2}{3}(d-1)\mathrm{Id}_{3^d},$$

its lowest eigenvalue is the shifted value

$$\lambda_{0,d} = -\frac{2}{3}(d-1) \tag{5.17}$$

with associated eigenspace E_{λ_0}, that is therefore the space of ground states of H.

Kernel of the local operator. We explain the construction of an element of the kernel $\widetilde{h}_j$ (5.16) in order to build the elements of $\ker(\widetilde{H})$ (and $\ker(H)$) which are the ground states of $\widetilde{H}$ (and H).

We consider the spaces $\mathbb{C}^2$ with basis $\{e_0, e_1\}$ and the tensor products $\mathbb{C}^2 \otimes \mathbb{C}^2$ with basis given by $\{e_{00}, e_{01}, e_{10}, e_{11}\}$, where we recall that $e_{ij} := e_i \otimes e_j$. The space $\mathbb{C}^2 \otimes \mathbb{C}^2$ can be seen as the state space of a system of two spin $\frac{1}{2}$ particles. As highlighted in Example 1.2.17, we have its decomposition into irreducible $\mathfrak{sl}_2$-representation

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \simeq V^{(0)} \oplus V^{(2)} \simeq \mathbb{C}^1 \oplus \mathbb{C}^3. \tag{5.18}$$

A basis of the spin 1, $V^{(2)}$, is given by taking the symmetric part of the tensor product of two spin 1/2: $\{v_{00} = e_{00}, v_{10} = v_{01} = \frac{e_{01}+e_{10}}{\sqrt{2}}, v_{11} = e_{11}\}$. Moreover, we fix a basis of $V^{(0)} = \mathbb{C}^1$, given by the element

$$\omega := \frac{e_{01} - e_{10}}{\sqrt{2}} = \frac{1}{\sqrt{2}} \sum_{\gamma,\delta=0,1} (e_{\gamma\delta}\epsilon^{\gamma\delta}) \in V^{(0)} \subset \mathbb{C}^2 \otimes \mathbb{C}^2,$$

with $\epsilon \in \mathbb{C}^2 \otimes \mathbb{C}^2$ the Levi-Civita antisymmetric tensor, i.e.

$$\epsilon^{\gamma\delta} := \begin{cases} 1 & \text{if } \operatorname{sgn}(\gamma,\delta) \equiv 0 \mod 2 \\ -1 & \text{if } \operatorname{sgn}(\gamma,\delta) \equiv 1 \mod 2 \\ 0 & \text{if } \gamma = \delta \end{cases} \tag{5.19}$$

Remark 5.3.1. The element ω is called the spin *singlet*. It is a generator of $V^{(0)} \simeq \mathbb{C}^1$ in Decomposition (5.18), i.e. of the spin 0 component of $\mathbb{C}^2 \otimes \mathbb{C}^2$. Notice that it is a rank-2 tensor.

Now consider four spins 1/2 particles, i.e. $(\mathbb{C}^2 \otimes \mathbb{C}^2) \otimes (\mathbb{C}^2 \otimes \mathbb{C}^2) \simeq \mathbb{C}^4 \otimes \mathbb{C}^4$ which decomposes into irreducible $\mathfrak{sl}_2$-representation as follows

$$\begin{aligned} \mathbb{C}^4 \otimes \mathbb{C}^4 &\simeq \mathbb{C}^2 \otimes (\mathbb{C}^2 \otimes \mathbb{C}^2) \otimes \mathbb{C}^2 \\ &\simeq \mathbb{C}^2 \otimes (\mathbb{C}^1 \oplus \mathbb{C}^3) \otimes \mathbb{C}^2 \\ &\simeq (\mathbb{C}^2 \otimes \mathbb{C}^2) \oplus (\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2) \\ &\simeq (\mathbb{C}^1 \oplus \mathbb{C}^3) \oplus (\mathbb{C}^1 \oplus \mathbb{C}^3 \oplus \mathbb{C}^3 \oplus \mathbb{C}^5) \end{aligned} \tag{5.20}$$

In physics, fixing a *valence bond* in $\mathbb{C}^4 \otimes \mathbb{C}^4 \simeq \mathbb{C}^2 \otimes (\mathbb{C}^2 \otimes \mathbb{C}^2) \otimes \mathbb{C}^2$ means fixing the spin singlet tensor ω in the middle pair $(\mathbb{C}^2 \otimes \mathbb{C}^2)$, i.e. considering the following element of $\overline{\Omega} \in \mathbb{C}^4 \otimes \mathbb{C}^4$:

$$\overline{\Omega} := \sum_{\alpha,\beta=0,1} e_\alpha \otimes \omega \otimes e_\beta \in \mathbb{C}^2 \otimes (\mathbb{C}^2 \otimes \mathbb{C}^2) \otimes \mathbb{C}^2. \tag{5.21}$$

We pictorially represent the constraint given by fixing $\omega \in \mathbb{C}^2 \otimes \mathbb{C}^2$ in Figure 5.7.

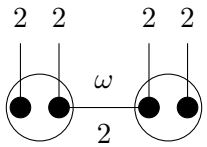


Figure 5.7: Consider two sites. On each site, there is a couple of spin $\frac{1}{2}$ (meaning $\mathbb{C}^2$). The *spin singlet* is fixed between nearby pair of spin 1/2 in different sites. The pair of spin 1/2 inside each site can have only either spin 0 or 1. The number 2 under the edge connecting the 1/2 spins indicates the rank of the spin singlet.

We have that $\overline{\Omega} \in \mathbb{C}^4 \otimes \mathbb{C}^4$ can be written as

$$\begin{aligned}
\overline{\Omega} &= \sum_{\alpha,\beta=0,1} e_\alpha \otimes \omega \otimes e_\beta \\
&= \frac{1}{\sqrt{2}} \sum_{\alpha,\beta=0,1} e_\alpha \otimes \Big[\sum_{\gamma,\delta=0,1} (e_{\gamma\delta}\epsilon^{\gamma\delta}) \Big] \otimes e_\beta \\
&= \frac{1}{\sqrt{2}} \sum_{\alpha,\beta=0,1} \sum_{\gamma,\delta=0,1} (e_{\alpha\gamma} \otimes e_{\delta\beta}\epsilon^{\gamma\delta}) \\
&= \frac{1}{\sqrt{2}} \sum_{\alpha,\beta=0,1} (e_{\alpha 0} \otimes e_{1\beta} - e_{\alpha 1} \otimes e_{0\beta}).
\end{aligned}$$

Defined the following matrix

$$E = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 \end{pmatrix},$$

then $\overline{\Omega}$ takes the form

$$\overline{\Omega} = \sum_{\alpha,\beta=0,1} \sum_{\gamma,\delta=0,1} E_{\gamma\delta} e_{\alpha\gamma} \otimes e_{\delta\beta}.$$

Remark 5.3.2. By construction, $\overline{\Omega}_{\alpha\beta} = \omega \in \mathbb{C}^1 \subseteq \mathbb{C}^2 \otimes \mathbb{C}^2$, therefore $\overline{\Omega} \in \mathbb{C}^1 \oplus \mathbb{C}^3 \subset \mathbb{C}^4 \otimes \mathbb{C}^4$. This means that $\overline{\Omega}$ is an element of the left vector space of the Decomposition (5.20), i.e. it can have only either spin 0 or spin 1. In particular, it is in the kernel of the projector $\widetilde{h} = P_2$ (5.16).

In order to recover the spin 1 pair, the element $\overline{\Omega} \in \mathbb{C}^4 \otimes \mathbb{C}^4$ is mapped to $\mathbb{C}^3 \otimes \mathbb{C}^3$. This is done via the map from $\mathbb{C}^4$ to its symmetrized part (isomorphic to $\mathbb{C}^3$); see Decomposition (5.18). Fix the basis $\{e^{ij}\}$ of $\mathbb{C}^{2*} \otimes \mathbb{C}^{2*}$ and the canonical basis $\{v_1, v_2, v_3\}$ of $\mathbb{C}^3$ given by the eigenvectors of S^3 of eigenvalues $1, 0, -1$ respectively; the map is the following

$$\mathcal{P}_j = v_1 \otimes e^{00} + v_2 \otimes (\frac{e^{01} + e^{10}}{\sqrt{2}}) + v_3 \otimes e^{11},$$

for $j = 1, \dots, d$. It can be written in the form

$$\mathcal{P}_j = \sum_{\substack{\alpha,\beta=0,1 \\ \gamma=1,2,3}} P^{\gamma}_{\alpha\beta} v_\gamma \otimes e^{\alpha\beta}, \tag{5.22}$$

with the following matrices

$$P^1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, P^2 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix}, P^3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

In what follows we omit the summation over repeated indices until last expression

$$\begin{aligned} \mathcal{P}_j \otimes \mathcal{P}_{j+1}(\overline{\Omega}) &= (P^{j_1}_{\alpha_1\beta_1} E_{\gamma\delta} P^{j_2}_{\alpha_2\beta_2}) e^{\alpha_1\beta_1}(e_{\alpha\gamma}) e^{\alpha_2\beta_2}(e_{\delta\beta})\ v_{j_1} \otimes v_{j_2} \\ &= (P^{j_1}_{\alpha_1\beta_1} E_{\gamma\delta} P^{j_2}_{\alpha_2\beta_2}) \delta^{\alpha_1\beta_1}_{\alpha\gamma} \delta^{\alpha_2\beta_2}_{\delta\beta}\ v_{j_1} \otimes v_{j_2} \\ &= \sum_{j_1,j_2=1,2,3} \sum_{\alpha,\beta=0,1} \sum_{\alpha_2,\beta_1=0,1} (P^{j_1}_{\alpha\beta_1} E_{\beta_1\alpha_2} P^{j_2}_{\alpha_2\beta})\ v_{j_1} \otimes v_{j_2} =: \Omega \in \mathbb{C}^3 \otimes \mathbb{C}^3. \end{aligned}$$

Remark 5.3.2 immediately implies that $\Omega \in \mathbb{C}^3 \otimes \mathbb{C}^3$ can have either spin 0 or 1 and it is therefore in the kernel of $\widetilde{h}$, the projector on spin 2.

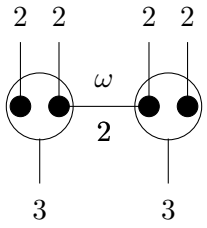


Figure 5.8: The pairs of 1/2 spins are projected to the symmetric part $\mathbb{C}^3$. The circle containing the pair of spin 1/2 can be thought of as the map $\mathcal{P}_j$ acting on the j-th site.

Remark 5.3.3. Denote

$$\Omega_{\alpha\beta} := \sum_{j_1,j_2=1,2,3} \sum_{\alpha_2,\beta_1=0,1} (P^{j_1}_{\alpha\beta_1} E_{\beta_1\alpha_2} P^{j_2}_{\alpha_2\beta})\ v_{j_1} \otimes v_{j_2}.$$

If we identify $v_1 = v_{00}, v_2 = v_{01} (= v_{10}), v_3 = v_{11}$ via $\mathcal{P}_j|_{\mathbb{C}^3}$, and we define $\psi_{00} = \frac{v_{00}}{\sqrt{2}}$, $\psi_{01} = \psi_{10} = v_{01}$ and $\psi_{11} = \frac{v_{11}}{\sqrt{2}}$, then we recover

$$\Omega_{\alpha\beta} = \frac{1}{\sqrt{2}} \sum_{\gamma,\delta=0,1} \psi_{\alpha\gamma} \otimes \psi_{\delta\beta} \epsilon^{\gamma\delta},$$

which is the expression given in the original article [AKLT88].

Ground state. A 4-dimensional family of ground states $\psi_{\text{aklt}} \in (\mathbb{C}^3)^{\otimes d}$ of $\widetilde{H}$ is constructed. It is obtained from a chain of couples of spin 1/2 in each site $j = 1, \dots, d$, fixing $\omega \in \mathbb{C}^2 \otimes \mathbb{C}^2$ between every pair of adjacent spin 1/2 (one from site j and one from site $j+1$); then mapping the pair of spin 1/2 in the site j onto the symmetric part, isomorphic to $\mathbb{C}^3$.

Generalizing Equation (5.21) we take an element $\overline{\psi} \in (\mathbb{C}^4)^{\otimes d}$ with fixed spin singlets between all pairs of spin 1/2, connecting adjacent sites. In coordinates, it is written as

$$\begin{aligned}\overline{\psi} &= \frac{1}{\sqrt{2}} \sum_{\alpha,\beta=0,1} \sum_{\substack{\alpha_j=0,1 \\ j=2,\dots d}} \sum_{\substack{\beta_l=0,1 \\ l=1,\dots,d-1}} e_{\alpha\beta_1} \otimes e_{\alpha_2\beta_2} \otimes \cdots \otimes e_{\alpha_d\beta} \epsilon^{\beta_1\alpha_2} \dots \epsilon^{\beta_{d-1}\alpha_d} \\ &= \sum_{\alpha,\beta=0,1} \sum_{\substack{\alpha_j=0,1 \\ j=2,\dots d}} \sum_{\substack{\beta_l=0,1 \\ l=1,\dots,d-1}} (E_{\beta_1\alpha_2} E_{\beta_2,\alpha_3} \dots E_{\beta_{d-1}\alpha_d})\ e_{\alpha\beta_1} \otimes e_{\alpha_2\beta_2} \otimes \cdots \otimes e_{\alpha_d\beta},\end{aligned}$$

where E is the matrix

$$E = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}.$$

For any two adjacent sites j and $j+1$, there is a spin 1/2 at site i and a spin 1/2 at site $i+1$ which are contracted with an ϵ tensor (c.f. Equation 5.19) to form a singlet. Thus when $\overline{\psi}$ is restricted to sites j and $j+1$, it has only spins 0 and 1. Hence $\widetilde{h}_j \overline{\psi} = 0$ and therefore $\overline{\psi}$ is a ground state of $\widetilde{H} = \sum_{j=1}^{d-1} \widetilde{h}_j$.

Figure 5.9: On each site, we consider a pair of spin 1/2. Every nearby pair of spin 1/2 in different sites are put in a *spin singlet*. The pair of spin 1/2 inside each site can have only spin either 0 or 1.

In order to recover the spin 1 chain, we consider the projector from $(\mathbb{C}^4)^{\otimes d}$ to $(\mathbb{C}^3)^{\otimes d}$ given by the tensor product of projectors (5.22), for $j = 1, \dots, d$

$$\mathcal{P} = \bigotimes_{j=1}^{d} \mathcal{P}_j = \sum_{\substack{\alpha_j,\beta_j=0,1 \\ \gamma_j=1,2,3, \\ j=1\dots,d}} (P^{\gamma_1}_{\alpha_1\beta_1} \cdots P^{\gamma_d}_{\alpha_d\beta_d}) e_{\gamma_1} \otimes e^{\alpha_1\beta_1} \otimes \cdots \otimes e_{\gamma_d} \otimes e^{\alpha_d\beta_d}$$

and applied to element $\overline{\psi} \in (\mathbb{C}^4)^{\otimes d}$ it gives

$$\psi = \sum_{\alpha,\beta=0,1} \sum_{\substack{\alpha_j,\beta_j=0,1 \\ \gamma_j=1,2,3 \\ j=1\dots,d}} \left(P^{\gamma_1}_{\alpha,\beta_1} E_{\beta_1\alpha_2} P^{\gamma_2}_{\alpha_2\beta_3} \cdots E_{\beta_{d-1}\alpha_d} P^{\gamma_d}_{\alpha_d\beta} \right) e_{\gamma_1\dots\gamma_d} \in (\mathbb{C}^3)^{\otimes d}. \tag{5.23}$$

Notice that 4 degrees of freedom remain free at the boundaries of the chain.

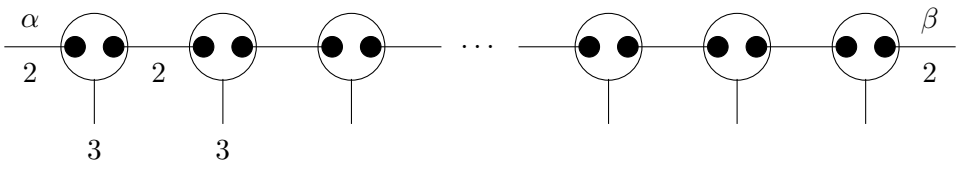


Figure 5.10: The spin 1 chain obtained by the spin 3/2 chain (pairs of 1/2 spins), after projecting $\mathbb{C}^2 \otimes \mathbb{C}^2$ into the symmetric part $\mathbb{C}^3$. The boundaries have 2 degrees of freedom left corresponding to the indices α and β respectively.

Expression (5.23) is simplified introducing $\widetilde{A}^i = P^i E$, $i = 1, 2, 3$, so that

$$\widetilde{A}^1 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}, \widetilde{A}^2 = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \widetilde{A}^3 = \begin{pmatrix} 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}.$$

Notice that

$$\sum_{i=1}^{3} \widetilde{A}^{i\dagger} \widetilde{A}^i = \frac{3}{4} \mathrm{Id}_2;$$

therefore the matrices $\widetilde{A}^i$ may be rescaled by $\frac{2}{\sqrt{3}}$, $A^i := \frac{2}{\sqrt{3}} \widetilde{A}^i$, $i = 1, 2, 3$. We obtain

$$\psi = \sum_{\alpha,\beta=0,1} \sum_{\substack{\alpha_i,\beta_i=0,1 \\ \gamma_i=1,2,3 \\ i=1\ldots,d}} \left(A^{\gamma_1}_{\alpha,\beta_1} A^{\gamma_2}_{\beta_1\alpha_2} \cdots A^{\gamma_{d-1}}_{\beta_{d-1}\alpha_d} P^{\gamma_d}_{\alpha_d\beta} \right) e_{\gamma_1\ldots\gamma_d} \in (\mathbb{C}^3)^{\otimes d}.$$

Fixing indices α and β we have the tensor

$$\psi_{\alpha\beta} = \sum_{\substack{\alpha_i,\beta_i=0,1 \\ \gamma_i=1,2,3 \\ i=1\ldots,d}} \left(A^{\gamma_1}_{\alpha,\beta_1} A^{\gamma_2}_{\beta_1\alpha_2} \cdots A^{\gamma_{d-1}}_{\beta_{d-1}\alpha_d} P^{\gamma_d}_{\alpha_d\beta} \right) e_{\gamma_1\ldots\gamma_d} \in (\mathbb{C}^3)^{\otimes d};$$

In [AKLT88] it is proved that $\psi_{00}, \psi_{01}, \psi_{10}, \psi_{11}$ are linearly independent and that consequently they span a 4-dimensional vector space of ground states. Therefore, for every choice of vectors $v_L, \widetilde{v}_R \in \mathbb{C}^2$ an element of their span can be written as

$$\psi_{\text{aklt}} = \sum_{\substack{\gamma_j=1,2,3 \\ j=1\ldots,d}} \left({v_L}^\dagger A^{\gamma_1} A^{\gamma_2} \cdots P^{\gamma_d} \widetilde{v}_R \right) e_{\gamma_1} \otimes \cdots \otimes e_{\gamma_d}.$$

If we multiply by $(\frac{2}{\sqrt{3}} E)(\frac{\sqrt{3}}{2} E^{-1})$ and we denote $\frac{\sqrt{3}}{2} E^{-1} \widetilde{v}_R = v_R$ we have

$$\psi_{\text{aklt}} = \sum_{\substack{\gamma_j=1,2,3 \\ j=1\ldots,d}} \left({v_L}^\dagger A^{\gamma_1} A^{\gamma_2} \cdots A^{\gamma_d} v_R \right) e_{\gamma_1} \otimes \cdots \otimes e_{\gamma_d}$$

with the choice of matrices

$$A^1 = \begin{pmatrix} 0 & \sqrt{\frac{2}{3}} \\ 0 & 0 \end{pmatrix}, A^2 = \begin{pmatrix} -\frac{1}{\sqrt{3}} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix}, A^3 = \begin{pmatrix} 0 & 0 \\ -\sqrt{\frac{2}{3}} & 0 \end{pmatrix}.$$

We obtain a 4-dimensional eigenspace associated to the lowest eigenvalue of the Hamiltonians $\widetilde{H}$ and H. The tensor ψ_{aklt} is a ground state for every $v_L, v_R \in \mathbb{C}^2$ for construction. Moreover $\psi_{\text{aklt}} \in \text{MPS}(2,3,d)$, see Definition 5.2.2; in particular $\psi_{\text{aklt}} \in \text{hMPS}(2,3,d)$, see Definition 5.2.7.

5.4 Matrix product operators

A Matrix Product Operators are generalizations of matrix product states to the space of operators. We give the definition of matrix product operator for $O \in \text{Hom}((\mathbb{C}^n)^{\otimes d}, (\mathbb{C}^n)^{\otimes d})$ (associated to the path graph), but the definition can be generalized to the case of operators in $\text{Hom}(\bigotimes_{i=1}^d \mathbb{C}^{n_i}, \bigotimes_{i=1}^d \mathbb{C}^{n_i})$ and to other graphs.

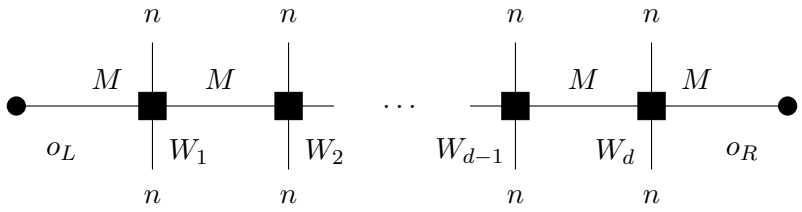


Figure 5.11: MPO representation associated to the path graph on d vertices, with open boundary conditions. Analogously to matrix product states, a tensor W_k is associated to each vertex, for every $k = 1, \dots, d$ and two vectors are associated to the boundaries. The tensors are pairwise contracted along the edges of the graph.

Consider $O \in \text{Hom}((\mathbb{C}^n)^{\otimes d}, (\mathbb{C}^n)^{\otimes d})$. Let $\{e_j\}_{j=1}^n$ and $\{e^{j'}\}_{j'=1}^n$ be chosen basis of $\mathbb{C}^n$ and $\mathbb{C}^{n*}$ respectively. Let $M \in \mathbb{N}$ and let $\{v_j\}_{j=1}^M$ and $\{v^{j'}\}_{j'=1}^M$ fixed basis of $\mathbb{C}^M$ and its dual $\mathbb{C}^{M*}$ respectively.

We define, for every $k = 1, \dots, d$, a tensor (of order four) $W_k \in \mathbb{C}^n \otimes \mathbb{C}^{n*} \otimes \mathbb{C}^M \otimes \mathbb{C}^{M*}$. In particular, for every $j_k, j'_k = 1, \dots, n$ we have

$$W_k{}_{j'_k}^{j_k} \in \mathbb{C}^{M \times M},$$

which is an $M \times M$ matrix in the given bases. Let $o_L, o_R \in \mathbb{C}^M$ be boundary vectors (analogously to the case of matrix product states with open boundary conditions).

Definition 5.4.1. An operator $O \in \text{Hom}((\mathbb{C}^n)^{\otimes d}, (\mathbb{C}^n)^{\otimes d})$ is said to have an MPO representation if it can be written as

$$O = \sum_{\substack{j_1,\dots,j_d=1 \\ j'_1,\dots,j'_d=1}}^{n} \left[o_L^\dagger \, W_1{}_{j'_1}^{j_1} \dots W_d{}_{j'_d}^{j_d} \, o_R \right] (e_{j_1} \otimes \dots \otimes e_{j_d}) \otimes (e^{j'_1} \otimes \dots \otimes e^{j'_d}).$$

A local Hamiltonian, c.f. Definition 1.2.8, can be *exactly* represented in the MPO format, with bond dimensions M small enough, c.f. [Sch11, HP18].

Consider the local Hamiltonian $H : (\mathbb{C}^n)^{\otimes d} \to (\mathbb{C}^n)^{\otimes d}$

$$H = \sum_{j=1}^{d-1} h_j = \sum_{j=1}^{d-1} \mathrm{Id}^{j-1} \boxtimes h \boxtimes \mathrm{Id}^{d-j-1}, \tag{5.24}$$

with $h : \mathbb{C}^n \otimes \mathbb{C}^n \to \mathbb{C}^n \otimes \mathbb{C}^n$ local operator acting on two nearby sites.

If the Hamiltonian is a sum over the number of sites of the same local term, as in expression (5.24), then $W_k = W \in (\mathbb{C}^{n*} \otimes \mathbb{C}^n) \otimes (\mathbb{C}^M \otimes \mathbb{C}^{M*})$, for every $k = 1, \dots, d$. Moreover, for $\alpha, \alpha' = 1, \dots, M$, the matrix $W_\alpha^{\alpha'} \in \mathbb{C}^{n \times n}$ has a form that is related to the standard writing of the Hamiltonian H, i.e. it is a matrix whose entries are the operators appearing in the local operator h. We give the example of the AKLT Hamiltonian.

Example 5.4.2. The AKLT Hamiltonian admits an exact MPO representation. We recall the Hamiltonian (5.12), defined in Section 5.3, where $\vec{S}_j$ is given in Equation (5.11)

$$\begin{aligned}
H =& \sum_{j=1}^{d-1} \Big(\vec{S}_j \cdot \vec{S}_{j+1} + \frac{1}{3} (\vec{S}_j \cdot \vec{S}_{j+1})^2 \Big) \\
=& \sum_{j=1}^{d-1} S_j^1 S_{j+1}^1 + \sum_{j=1}^{d-1} S_j^2 S_{j+1}^2 + \sum_{j=1}^{d-1} S_j^3 S_{j+1}^3 + \frac{1}{3} \Big(\sum_{j=1}^{d-1} (S^1)_j^2 (S^1)_{j+1}^2 + \cdots + \sum_{j=1}^{d-1} (S^3)_j^2 (S^3)_{j+1}^2 \Big) + \\
&+ \frac{1}{3} \Big(\sum_{j=1}^{d-1} S_j^1 S_{j+1}^2 + \sum_{j=1}^{d-1} S_j^1 S_{j+1}^3 + \cdots + \sum_{j=1}^{d-1} S_j^3 S_{j+1}^2 + \sum_{j=1}^{d-1} S_j^3 S_{j+1}^3 \Big) \\
=& (S^1 \boxtimes S^1 \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^3 \boxtimes S^3) + \\
&+ \frac{1}{3} ((S^1)^2 \boxtimes (S^1)^2 \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes (S^3)^2 \boxtimes (S^3)^2) + \\
&+ \frac{1}{3} (S^1 S^2 \boxtimes S^1 S^2 \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^3 S^2 \boxtimes S^3 S^2).
\end{aligned}$$

Notice that the spin operators

$$S^1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, S^2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, S^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

can be written in terms of products the $S^j S^k$, using the commutation relations

$$S^1 = -iS^2 S^3 + iS^3 S^2, \quad S^2 = -iS^3 S^1 + iS^1 S^3, \quad S^3 = -iS^1 S^2 + iS^2 S^1.$$

With these substitutions in the Hamiltonian, we get

$$
\begin{aligned}
H = & \frac{1}{3}\left(\sum_{j=1}^{d-1} S_j^1 S_{j+1}^1 + \sum_{j=1}^{d-1} S_j^2 S_{j+1}^2 + \sum_{j=1}^{d-1} S_j^3 S_{j+1}^3\right) + \\
& - \frac{2}{3}\left(\sum_{j=1}^{d-1} S_j^1 S_{j+1}^2 + \sum_{j=1}^{d-1} S_j^1 S_{j+1}^3 + \cdots + \sum_{j=1}^{d-1} S_j^3 S_{j+1}^2 + \sum_{j=1}^{d-1} S_j^3 S_{j+1}^3\right) + \\
& + \ S^1S^2 \boxtimes S^2S^1 \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^1S^3 \boxtimes S^3S^1 + \\
& + S^2S^1 \boxtimes S^1S^2 \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^2S^3 \boxtimes S^3S^2 + \\
& + S^3S^1 \boxtimes S^1S^3 \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^3S^2 \boxtimes S^2S^3 = \\
= & \frac{1}{3}\left((S^1)^2 \boxtimes (S^1)^2 \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes (S^3)^2 \boxtimes (S^3)^2\right) + \\
& + S^1S^2 \boxtimes (-\frac{2}{3}S^1S^2 + S^2S^1) \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^1S^3 \boxtimes (-\frac{2}{3}S^1S^3 + S^3S^1) + \\
& + S^2S^1 \boxtimes (-\frac{2}{3}S^2S^1 + S^1S^2) \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^2S^3 \boxtimes (-\frac{2}{3}S^2S^3 + S^3S^2) + \\
& + S^3S^1 \boxtimes (-\frac{2}{3}S^3S^1 + S^1S^3) \boxtimes \mathrm{Id}_3^{\boxtimes d-2} + \cdots + \mathrm{Id}_3^{\boxtimes d-2} \boxtimes S^3S^2 \boxtimes (-\frac{2}{3}S^3S^2 + S^2S^3).
\end{aligned}
$$

Consider the following matrix $\widetilde{W} \in \mathbb{C}^{Mn \times Mn}$, with $M = 11$ and $n = 3$; whose entries are operators appearing in the Hamiltonian:

$$
\widetilde{W} = \left(\begin{array}{cccccccccc|c}
\mathrm{Id}_3 & (S^1)^2 & (S^2)^2 & (S^3)^2 & S^1S^2 & S^1S^3 & S^2S^1 & S^2S^3 & S^3S^1 & S^3S^2 & \mathbb{O}_{3\times 3} \\
\hline
 & & & & & & & & & & \frac{1}{3}(S^1)^2 \\
 & & & & & & & & & & \frac{1}{3}(S^2)^2 \\
 & & & & & & & & & & \frac{1}{3}(S^3)^2 \\
 & & & & & & & & & & (-\frac{2}{3}S^1S^2 + S^2S^1) \\
 & & & & & & & & & & (-\frac{2}{3}S^1S^3 + S^3S^1) \\
 & & & & \mathbb{O}_{30\times 30} & & & & & & (-\frac{2}{3}S^2S^1 + S^1S^2) \\
 & & & & & & & & & & (-\frac{2}{3}S^2S^3 + S^3S^1) \\
 & & & & & & & & & & (-\frac{2}{3}S^3S^1 + S^1S^3) \\
 & & & & & & & & & & (-\frac{2}{3}S^3S^2 + S^2S^3) \\
 & & & & & & & & & & \mathrm{Id}_3
\end{array}\right).
$$

The matrix $\widetilde{W} \in \mathbb{C}^{Mn \times Mn}$ is a matrix reshape of the tensor W we are looking for.

It is straightforward to check that this matrix (of matrices) gives the matrix product operator representation of the Hamiltonian H, that is:

$$
H = o_L^\dagger \left(\boxtimes_{i=1}^d \widetilde{W}\right) o_R,
$$

with $o_L^\dagger = (1, 0 \ldots, 0) \in \mathbb{C}^{M*}$ and $o_R = (0, \ldots, 0, 1)^t \in \mathbb{C}^M$.

5.5 Computation of the gradient

In order to solve Problem 5.0.2 and to implement the nonlinear conjugate gradient on the variational variety of matrix product states with open boundary conditions, we need to compute the gradient $\nabla f \in D_{\text{MPS}}$, where $f := \rho \circ \phi : D_{\text{MPS}} \to \mathbb{R}$ (5.1) is defined as follows:

$$f(A) = \rho \circ \phi(A) = \frac{\phi(A)^\dagger H \phi(A)}{\phi(A)^\dagger \phi(A)}.$$

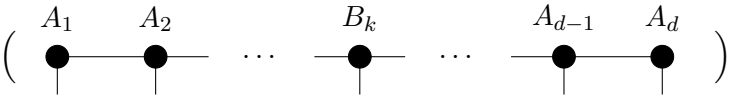
Choose a set of coordinates $\{z_1, \ldots, z_q\}$ for $\mathcal{H} = \mathbb{C}^q$. The gradient of the expectation value ρ in $z = (z_1, \ldots, z_q)$ is computed as the partial derivative with respect to the conjugate variables $\overline{z} = (\overline{z_1}, \ldots, \overline{z_q})$, c.f. [PP+08].

$$\begin{aligned}
\nabla \rho_z = 2\frac{\partial \rho(z)}{\partial \overline{z}} &= 2\left(\frac{Hz}{z^\dagger z} - \frac{z^\dagger H z}{(z^\dagger z)^2}\, z\right) \\
&= 2\frac{1}{z^\dagger z}(Hz - hz) \\
&= 2\frac{1}{z^\dagger z}(H - h\mathbb{1}_n)z \in \mathbb{C}^q, \quad \text{where } h = \rho(z).
\end{aligned}$$

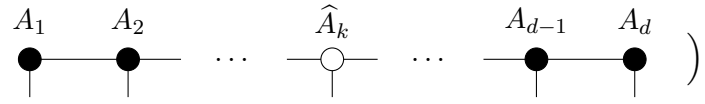
Consider $A = (A_1, \ldots, A_d) \in D_{\text{MPS}}$ and $B = (B_1, \ldots, B_d) \in T_A D_{\text{MPS}} \simeq D_{\text{MPS}}$. The differential map of ϕ at A, $d\phi_A : T_A D_{\text{MPS}} \simeq D_{\text{MPS}} \to T_{\phi(A)}\,\text{MPS}$, is given by

$$d\phi_A(B) = \frac{d\phi(A + tB)}{dt}\Big|_{t=0} = \sum_{k=1}^{d} \sum_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d} \left(A_1^{i_1} \cdots B_k^{i_k} \cdots A_d^{i_d}\right) e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_d}^{(d)}.$$

Remark 5.5.1. We can pictorially represent the vector $d\phi_A(B) \in T_{\phi(A)}\,\text{MPS}$:

$d\phi_A(B) = \sum_{k=1}^{d}$ 

Analogously, the linear operator $d\phi_A$ is represented as a sum of d tensors with the same format of $\phi(A)$, each one with the k-th tensor removed:

$d\phi_A = \sum_{k=1}^{d}$ 

Denote by $\overline{A}$ the element of D_{MPS} with coordinates that are complex-conjugated to the coordinates of A. The function $\rho(\phi(\cdot))$ is a real valued function depending on the independent variables A and $\overline{A}$. Notice that, if $A \in D_{\text{MPS}}$, then

$$\phi(A)^\dagger = \phi(\overline{A})^t.$$

Denote the vector of partial derivatives with respect to the formal conjugated variables of D_{MPS} by

$$\frac{\partial}{\partial \overline{A}} = \left(\frac{\partial}{\partial \overline{A}_k}\right)_{k=1,\dots,d} = \left(\frac{\partial}{\partial \overline{a}_{k l_k}}\right)_{\substack{l_k=l_1,\dots,l_d \\ k=1,\dots,d}} = \left(\frac{\partial}{\partial \overline{a}_j}\right)_{j=l} = \left(\frac{\partial}{\partial \overline{a}_{k\,\gamma_k\delta_k}^{\,i_k}}\right)_{k i_k \gamma_k \delta_k},$$

where l_k, $k = 1, \dots, d$ and l are indices associated to bijections L_k, $k = 1, \dots, d$ and L given in Remark 5.2.5.

Given $A \in D_{\mathrm{MPS}}$, the gradient $\nabla(\rho \circ \phi)_A$ is computed by the chain rule, differentiating over the variable $\overline{A}$ of D_{MPS} and it is the vector reshape of the following tensor

$$\begin{aligned} G_A &:= 2\,\frac{\partial(\rho\circ\phi(A))}{\partial\overline{A}} = 2\,\frac{\partial}{\partial\overline{A}}\left(\frac{\phi(\overline{A})^t H\phi(A)}{\phi(\overline{A})^t\phi(A)}\right) = \qquad (5.25)\\ &\frac{2}{(\phi(A)^\dagger\phi(A))^2}\left(\phi(A)^\dagger\phi(A)\frac{\partial}{\partial\overline{A}}\left(\phi(\overline{A})^t H\phi(A)\right) - \left(\phi(A)^\dagger H\phi(A)\right)\frac{\partial}{\partial\overline{A}}\left(\phi(\overline{A})^t\phi(A)\right)\right)\\ &= \frac{2}{\phi(A)^\dagger\phi(A)}\left(\frac{\partial}{\partial\overline{A}}(\phi(\overline{A})^t)\left(H - h\cdot\mathbb{1}\right)\phi(A)\right), \quad \text{where } h = \rho(\phi(A)). \end{aligned}$$

Denote the MPO representation of $(H - h \cdot \mathbb{1})$ by $\widetilde{H}$, c.f. Section 5.4, and denote

$$(G_A)_k := \frac{2}{\phi(A)^\dagger\phi(A)}\left(\frac{\partial}{\partial\overline{A}_k}(\phi(\overline{A})^t)\widetilde{H}\phi(A)\right), \quad \text{for } k = 1, \dots, d.$$

The gradient can be compactly written as $G_A = \sum_{k=1}^d (G_A)_k$ and its pictorial representation is given in Figure 5.12.

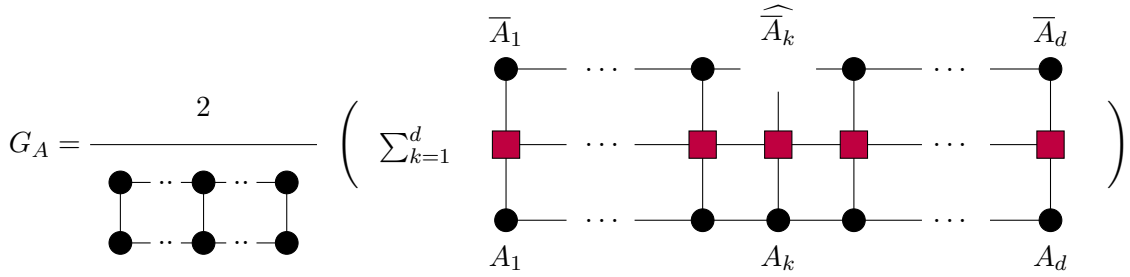


Figure 5.12: The pictorial representation of the gradient $G_A = \nabla(\rho \circ \phi)_A$.

Algorithm. Algorithm 5 is the pseudo-code of the NLCG method applied to $f = \rho \circ \phi$. With abuse of notation, we always consider the vector reshape of tensors under consideration, i.e. we denote $A = L(A)$, $G_A = L(G_A)$. In the pseudo-code, we denote elements of D_{MPS} by $A_{(j)}$ with $j \in \mathbb{N}$.

Remark 5.5.2. Notice that the NLCG method computes the tangent vector corresponding to the steepest descent direction in the Hilbert space, but then updates the tensors by simply adding them in the parameter space. Other methods could in principle do a line search along geodetic paths through the variety, which would involve integrating the geodesic equation or exploiting the concepts of retraction and vector transport, which are relaxations of the classical geometric concepts of motion along geodesics and parallel transport [AMS09, HVDH21]. Moreover, p_{j+1} is obtained by adding vectors $G_{A_{(j+1)}}$ and p_j that belong to tangent spaces at different points $A_{(j+1)}$ and $A_{(j)}$. This would require the parallel transport of p_j. Instead, we are approximating $A_{(j+1)} = A_{(j)} + \alpha_j p_j$, with $p_j = -G_{A_{(j+1)}} + \beta^{\mathrm{FR}}_{j+1}\, p_j$. In any case, the line search ensures the decrease of the functional value after every step and the only disadvantage consists in a slower convergence [MHO13].

Algorithm 5: NLCG on MPS

Input: Initial point $A_{(0)} \in D_{\mathrm{MPS}}$, tolerance $t \in \mathbb{R}$
Output: $A^* \in D_{\mathrm{MPS}}$ such that $\phi(A^*) \in \mathrm{MPS}$ approximate ground state

```
Compute N = dim D_MPS;
Evaluate f_{A_(0)} = f(A_(0));
p_0 ← −G_{A_(0)};
j ← 0;
while |G_{A_(j)}| > t do
    α_j = arg min_α f(A_(j) + α p_j);
    Set A_(j+1) = A_(j) + α_j p_j;
    Compute G_{A_(j+1)};
    if j = N then
        β_j^FR ← 0;
        j ← 0;
    else
        β_{j+1}^FR ← (G†_{A_(j+1)} G_{A_(j+1)}) / (G†_{A_(j)} G_{A_(j)});
    end
    p_{j+1} ← −G_{A_(j+1)} + β_{j+1}^FR p_j;
    j ← j + 1;
end
```

5.6 Section of the domain

In this section we describe the selection of a linear subspace of the domain of the matrix product state map. The subspace of the parameter space that we will construct, will contain a finite number of points of each fiber of the map. We briefly explain how we used the theory to propose a first variation of the NLCG and why the method was not successful. In the following, we recall again the Theorem of Dimension of the Fiber [Sha94, Thm. 1.25] of a morphism and the Bertini Theorem [Har13, Thm. 8.18]. The theorems allow us to give a reparametrization of the matrix product state map.

Proposition 5.6.1 (Fiber of a morphism)**.** *Let $\phi : W \to V$ be a dominant regular map of irreducible varieties. Then*

1. $\dim(W) \geq \dim(V)$.
2. *If $Q \in \phi(W)$, then $\dim(\phi^{-1}(Q)) \geq \dim(W) - \dim(V)$ for every $Q \in V$, with equality holding exactly on a nonempty open subset U of V.*
3. *The sets $V_i = \{Q \in V | \dim(\phi^{-1}(Q)) \geq i\}$ are closed in $\phi(W)$.*

Theorem 5.6.2 (Bertini)**.** *Let $\mathcal{V}$ be a nonsingular closed subvariety of $\mathbb{P}^q_K$, where K is an algebraically closed field. Then there exists a hyperplane $Y \subset \mathbb{P}^q_K$, not containing $\mathcal{V}$, and such that the scheme $Y \cap \mathcal{V}$ is regular at every point. Furthermore, the set of hyperplanes with this property forms an open dense subset of the complete linear system $|Y|$, considered as a projective space.*

Let $\phi : D_{\text{MPS}} \to \text{Im } (\phi) = \text{MPS}^\circ(\mathbf{m}, \mathbf{n}, d)$ be the parametrization of matrix product states with open boundary conditions (5.7), with

$$D_{\text{MPS}} = \bigtimes_{k=1}^{d} \mathbb{C}^{m_{k-1} \times n_k \times m_k}, \quad m_0 = m_d = 1.$$

Denote the dimension of the domain of ϕ by $N_{\text{MPS}} = \sum_{i=1}^{d} n_i m_{i-1} m_i$, $(m_0 = m_d = 1)$ and assume $s_{\text{MPS}} = \dim \text{MPS}(\mathbf{m}, \mathbf{n}, d)$.

The set $\text{MPS}^\circ(\mathbf{m}, \mathbf{n}, d) = \text{Im } (\phi)$ is a cone [CMS19]: if $Q \in \text{MPS}^\circ(\mathbf{m}, \mathbf{n}, d)$ then $\lambda Q \in \text{MPS}^\circ(\mathbf{m}, \mathbf{n}, d)$ for every $\lambda \in \mathbb{C}$. Consider the associated map between projective spaces

$$\phi_p : \mathbb{P}^{n_1 m_1 - 1} \times \left(\bigtimes_{i=2}^{n-1} \mathbb{P}^{n_i m_{i-1} m_i - 1} \right) \times \mathbb{P}^{n_d m_{d-1}} \to \mathbb{P}(\text{MPS}^\circ(\mathbf{m}, \mathbf{n}, d)) \subset \mathbb{P}(\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}),$$

which is polynomial and dominant. Define $\overline{\text{Im } (\phi_p)} =: \mathcal{M}ps(\mathbf{m}, \mathbf{n}, d)$. Denote the domain of ϕ_p by $D_{\mathcal{M}ps}$, its projective dimension by $N_{\mathcal{M}ps} = N_{\text{MPS}} - 1$ and the projective dimension of $\mathcal{M}ps(\mathbf{m}, \mathbf{n}, d)$ by $s_{\mathcal{M}ps} = s_{\text{MPS}} - 1$.

Given $[Q] \in \mathcal{M}ps(\mathbf{m}, \mathbf{n}, d)$, by Proposition 5.6.1 the dimension of the fiber ${\phi_p}^{-1}([Q])$ is bounded by

$$\dim({\phi_p}^{-1}([Q])) \geq N_{\mathcal{M}ps} - s_{\mathcal{M}ps} =: e_{\mathcal{M}ps}.$$

Moreover, by Bertini's Theorem 5.6.2 the intersection of ${\phi_p}^{-1}([Q])$ with $e_{\mathcal{M}ps}$ general hyperplanes Y_j, for $j = 1, \dots, e_{\mathcal{M}ps}$:

$$Y_1 \cap \cdots \cap Y_{e_{\mathcal{M}ps}} \cap {\phi_p}^{-1}([Q])$$

consists of a finite number of reduced points that coincides with the degree of the fiber. Denote $Y := Y_1 \cap \cdots \cap Y_{e_{\mathcal{M}ps}}$. Therefore, ${\phi_p}^{-1}|_Y : Y \subset D_{\mathcal{M}ps} \to \mathcal{M}ps(\mathbf{m}, \mathbf{n}, d)$ gives a reparametrization of the variety $\mathcal{M}ps(\mathbf{m}, \mathbf{n}, d)$.

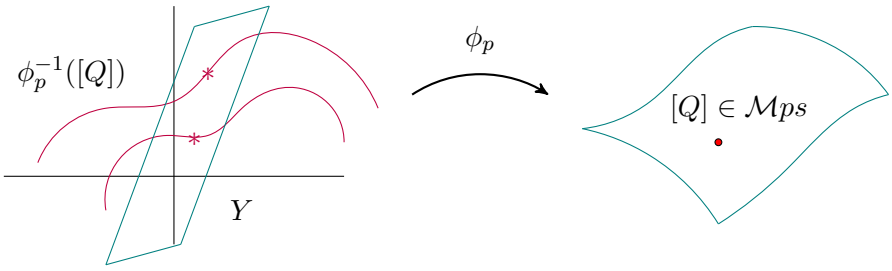


Figure 5.13: Pictorial representation of the matrix product state maps. Given $[Q] \in \mathcal{M}ps(\mathbf{m}, \mathbf{n}, d)$, the fiber $\phi_p^{-1}([Q])$ (in red) intersects $Y = Y_1 \cap \cdots \cap Y_{e_{\mathcal{M}ps}}$ in a finite number of reduced points. Therefore $\overline{\phi_p(Y)} = \mathcal{M}ps(\mathbf{m}, \mathbf{n}, d)$.

The first attempt we did was based on this approach: in the affine setting, we fixed an embedding $i : \mathbb{C}^{s_{\text{MPS}}} \hookrightarrow \widehat{Y} \subset D_{\text{MPS}}$ in order to set a general $\widehat{Y} \simeq \mathbb{C}^{s_{\text{MPS}}} \subset D_{\text{MPS}}$ such that

$$\widehat{Y} = \{A \in D_{\text{MPS}} : A = i(v), v \in \mathbb{C}^{s_{\text{MPS}}}\}.$$

We applied the NLCG to the following minimization problem

$$\begin{aligned} \min_{Q \in \text{MPS}^\circ} \rho(Q) &= \min_{A \in \widehat{Y}} \ \rho \circ \phi(A) = \min \{\rho \circ \phi(i(v)) : v \in \mathbb{C}^{s_{\text{MPS}}}\} \\ &= \min \left\{ \frac{\phi(i(v))^\dagger H \phi(i(v))}{\phi(i(v))^\dagger \phi(i(v))} : v \in \mathbb{C}^{s_{\text{MPS}}} \right\}, \end{aligned}$$

in order to find a solution of Problem 5.0.2.

Notice that, the domain of the reparametrization is a fixed vector space of the minimal number of parameters needed to parametrize the variety. This implies that the *gradient* of the functional $\rho \circ \phi \circ i$, which is an element of $\widehat{Y}$, has the minimal number of parameters. For this reason, implementing the NLCG with this method, the computations of the reparametrized gradient turned out to be faster compared to the standard one. On the other hand, the variation of the algorithm was not competitive due to the substantially higher number of iterations performed to reach convergence. The number was so higher to always make longer the runtime to convergence, compared to the standard algorithm. We can heuristically explain this behavior. We expect that the gradient, constrained to be a vector of the fixed subspace $\widehat{Y}$, finds a sequence of points that are certainly contained in different fibers of ϕ. However, even if we expect that the gradient is a vector that

points to the minimum at every step, it is not the deepest descent direction anymore and the algorithm takes too many more steps to reach convergence compared to the original gradient.

5.7 Decomposition of the domain

In this section, we describe a natural pointwise decomposition of the domain of the matrix product state map. We start from the general setting of tensor network varieties and then, in Subsection 5.7.1, we focus on the case of matrix product states. Finally, we study the open boundary case for both matrix product states and homogeneous matrix product states, c.f. Subsections 5.7.2 and 5.7.3 respectively.

Coming back to the general setting, we describe a natural pointwise decomposition of the domain of the parametrization of tensor network varieties.

Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network and let $\mathcal{V} = \mathcal{TNS}^{\Gamma}_{\mathbf{m},\mathbf{n}}$ be the associated tensor network variety defined as either the Euclidean or Zariski closure of the map

$$\begin{aligned}\Phi : \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d) &\to \mathcal{V} \subset V_1 \otimes \cdots \otimes V_d,\\ X = (X_1 \otimes \cdots \otimes X_d) &\mapsto (X_1 \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}).\end{aligned}$$

Denote $\mathcal{D} = \mathrm{Hom}(W_1, \dots, W_d; V_1, \dots, V_d)$ and $\mathcal{H} = V_1 \otimes \cdots \otimes V_d$.

The differential map of Φ at the point $X \in \mathcal{D}$ is

$$\begin{aligned}&d\Phi_X : T_X\mathcal{D} \simeq \mathcal{D} \to T_{\Phi(X)}\mathcal{V},\\ &Y = (Y_1, \dots, Y_d) \mapsto \sum_{k=1}^{d} (X_1 \otimes \cdots \otimes Y_k \otimes \cdots \otimes X_d) \cdot T(\Gamma, \mathbf{m}).\end{aligned}$$

Denote the dimension of the domain $\mathcal{D}$ by

$$N =: \dim \mathcal{D} = \sum_{v \in \mathbf{v}(\Gamma)} N_v n_v - d + 1.$$

The gauge subgroup, defined in Section 2.2, is the group

$$\mathcal{G}_{\Gamma,\mathbf{m}} \simeq \bigtimes_{e \in \mathbf{e}(\Gamma)} PGL_{m_e} \subseteq G(W_v : v \in \mathbf{v}(\Gamma)),$$

of dimension

$$g := \dim \mathcal{G}_{\Gamma,\mathbf{m}} = \sum_{e \in \mathbf{e}(\Gamma)} (m_e^2 - 1).$$

The action of $\mathcal{G}_{\Gamma,\mathbf{m}}$ on $\mathcal{D}$ is given in Definition 2.2.1. We denote the gauge orbit of $X \in \mathcal{D}$ by $\mathcal{O}_{\mathcal{G}(X)} = \mathcal{G}_{\Gamma,\mathbf{m}} \cdot X$. Denote $f = \dim \mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$, then the dimension of the orbit is

$\dim \mathcal{O}_{\mathcal{G}(X)} = g - f$. By Corollary 3.3.2, we can assume that the dimension of the variety $\mathcal{V}$ is bounded by

$$\dim \mathcal{V} = s \leq N - (g - f).$$

Since the domain $\mathcal{D}$ of Φ is isomorphic to $\mathbb{C}^N$, we endow $\mathcal{D}$ with the standard inner product of $\mathbb{C}^N$

$$\begin{aligned} h : \mathbb{C}^N \times \mathbb{C}^N &\to \mathbb{C} \\ (v, w) &\mapsto h(v, w) := w^\dagger v. \end{aligned}$$

The inner product allows the identification between $\mathbb{C}^N$ and its dual $\mathbb{C}^{N*}$ via the isomorphism

$$\begin{aligned} R : \mathbb{C}^N &\to \mathbb{C}^{N*} \\ w &\mapsto \left(w^\dagger : v \mapsto w^\dagger v \right). \end{aligned}$$

Let $T \in \mathcal{V}$ be a smooth point of the variety. If $X \in \Phi^{-1}(T) \subset \mathcal{D}$ is a point in the fiber, then $\mathrm{rk}(d\Phi_X) = s$. We denote by $\widetilde{\mathcal{N}}_X := \ker(d\Phi_X) \subset T_X \Phi^{-1}(T) \subset T_X D$ the vector space of tangent directions to the fiber at X. We define

$$\widetilde{\mathcal{B}}_X := \mathcal{N}_X^\perp = \{v \in \mathcal{D} \ : \ v^\dagger n = 0, \text{ for every } n \in \widetilde{\mathcal{N}}_X\} \subset \mathcal{D}.$$

Then

$$\begin{aligned} \dim \widetilde{\mathcal{N}}_X &= \dim \Phi^{-1}(T) = \mathrm{codim}_{\mathcal{D}} \mathcal{V} = N - s, \\ \dim \widetilde{\mathcal{B}}_X &= N - \dim \widetilde{\mathcal{N}}_X = s = \dim \mathcal{V}. \end{aligned}$$

The tangent space of $\mathcal{D}$ at the point X, as a complex vector space, splits into the following direct sum

$$T_X \mathcal{D} \simeq \mathcal{D} = \widetilde{\mathcal{N}}_X \oplus \widetilde{\mathcal{B}}_X,$$

and we have the following isomorphism of vector spaces

$$d\Phi_X|_{\mathcal{B}_X} : \widetilde{\mathcal{B}}_X \simeq T_X \mathcal{D} / \widetilde{\mathcal{N}}_X \to T_{\Phi(X)} \mathcal{V}.$$

Therefore, every tangent vector in $v \in T_{\Phi(X)} \mathcal{V}$ admits a unique representative $w \in \widetilde{\mathcal{B}}_X$ such that $d\Phi_X(w) = v$.

As we can see in the proof of Theorem 3.3.1, we have that $\mathcal{O}_{\mathcal{G}(X)} \subseteq \Phi^{-1}(T)$ and therefore

$$T_X \mathcal{O}_{\mathcal{G}(X)} \subseteq \widetilde{\mathcal{N}}_X,$$

whose dimension is bounded by

$$\dim T_X \mathcal{O}_{\mathcal{G}(X)} = \dim \mathcal{G}_{\Gamma, \mathbf{m}} - \dim \mathrm{Stab}_{\mathcal{G}_{\Gamma, \mathbf{m}}}(X) = g - f \leq N - s.$$

Moreover, for every $X \in \mathcal{D}$, $T_X \mathcal{O}_{\mathcal{G}(X)}$ is naturally isomorphic to the Lie algebra of $\mathcal{G}_{\Gamma,\mathbf{m}}$ quotiented by the Lie algebra of the stabilizer $\mathrm{Stab}_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)$, that we denote by $\mathfrak{s}_X$:

$$T_X \mathcal{O}_{\mathcal{G}(X)} \simeq \bigoplus_{e \in \mathbf{e}(\Gamma)} \mathfrak{sl}_{m_e} / \mathfrak{s}_X.$$

Remark 5.7.1. Based on results in Chapter 3, Section 3.4, we expect $f = 0$ in a wide range of cases and therefore $\dim T_X \mathcal{O}_{\mathcal{G}(X)} = g$. In this case, Theorem 3.3.2 implies that $s \leq N - g$. If moreover we assume that the orbit fills the fiber of Φ then

$$\mathcal{V} = s = N - g = \dim \mathcal{D} - \dim \mathcal{G}_{\Gamma,m},$$

is the expected dimension of $\mathcal{V}$, c.f. Equation (3.3), and therefore of $\widetilde{\mathcal{B}}_X$. Under these assumptions we have

$$T_X D = T_X \mathcal{O}_{\mathcal{G}(X)} \oplus \widetilde{\mathcal{B}}_X.$$

with isomorphism

$$T_X \mathcal{O}_{\mathcal{G}(X)} \simeq \bigoplus_{e \in \mathbf{e}(\Gamma)} \mathfrak{sl}_{m_e}.$$

By Remark 5.2.3, this holds for matrix product states with open boundary conditions.

5.7.1 Gauge orbit of matrix product states

We specialize the study of the gauge orbit in the case of matrix product states. Recall the map defining matrix product states $\mathcal{TNS}^{C_d}_{\mathbf{m},\mathbf{n}}$ given in Definition 5.2.1

$$\overline{\phi}(A_1, \ldots, A_d) = \sum_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d} \mathrm{Tr}\left(A_1^{i_1} \cdots A_d^{i_d}\right) e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_d}^{(d)}.$$

Denote the domain of the map by $D = \bigtimes_{k=1}^{d} \mathbb{C}^{m_{k-1} \times m_k \times n_k}$, with $m_0 = m_d$.

Consider $A = (A_1, \ldots, A_d)$ and $B = (B_1, \ldots, B_d) \in D$. The differential map of $\overline{\phi}$ at the point $A \in D$ is the linear map

$$\begin{aligned} d\overline{\phi}_A : T_A D \simeq D &\to T_{\overline{\phi}(A)} \mathcal{TNS}^{C_d}_{\mathbf{m},\mathbf{n}} \\ B &\mapsto d\overline{\phi}_A(B). \end{aligned}$$

given by

$$d\overline{\phi}_A(B) = \frac{d\overline{\phi}(A + tB)}{dt}\Big|_{t=0} = \sum_{k=1}^{d} \sum_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d} \mathrm{Tr}\left(A_1^{i_1} \cdots B_k^{i_k} \cdots A_d^{i_d}\right) e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_d}^{(d)}. \tag{5.26}$$

Gauge action on matrix product states. Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network. The gauge subgroup associated to $(\Gamma, \mathbf{m}, \mathbf{n})$, c.f. Section 2.2, is the group

$$\mathcal{G}_{\Gamma,\mathbf{m}} \simeq \bigtimes_{e\in\mathbf{e}(\Gamma)} PGL_{m_e} \subseteq G(W_v : v \in \mathbf{v}(\Gamma)),$$

that is the image of the homomorphism (2.9)

$$\Psi : \bigtimes_{e\in\mathbf{e}(\Gamma)} (GL(U_e) \times GL(U_e^*))/Z_e \to G(W_k : k \in \mathbf{v}(\Gamma)),$$

restricted to

$$G_{\Gamma,\mathbf{m}} := \bigtimes_{e\in\mathbf{e}(\Gamma)} GL_{m_e}^{\Delta} \simeq \bigtimes_{e\in\mathbf{e}(\Gamma)} GL_{m_e}.$$

In words, the gauge subgroup is obtained by taking, for every edge $e \in \mathbf{e}(\Gamma)$, the quotient of GL_{m_e} with the central subgroup $Z_e = \{(\lambda \mathrm{Id}_{U_e}, \lambda^{-1}\mathrm{Id}_{U_e^*}) : \lambda \in \mathbb{C}^*\}$.

We call $G_{\Gamma,\mathbf{m}}$ gauge group.

Considering matrix product states, the gauge group is

$$G_{C_d,\mathbf{m}} = \bigtimes_{e=1}^{d} GL_{m_e}.$$

Definition 5.7.2. Given $M = (M_1, \ldots, M_d) \in G_{C_d,\mathbf{m}}$ and $A \in D$, the gauge action $G_{C_d,\mathbf{m}} \times D \to D$ is given by

$$\begin{aligned} M \cdot A &= \left((M_{k-1})^{-1} A_k M_k\right)_{k=1,\ldots,d} \\ &= \left((M_{k-1})^{-1} A_k^{i_k} M_k\right)_{\substack{k=1,\ldots,d \\ i_k=1,\ldots,n_k}}, \end{aligned}$$

with convention $M_0 = M_d$. Denote the orbit of $A \in D$ under the action of $G_{C_d,\mathbf{m}}$ by

$$\mathcal{O}_{G(A)} = \{M \cdot A, \text{ for } M \in G_{C_d,\mathbf{m}}\} \subseteq D.$$

Lemma 5.7.3. *Given a point of the variety $Q \in \mathcal{TNS}_{\mathbf{m},\mathbf{n}}^{C_d}$ and an element $A \in D$ such that $A \in \overline{\phi}^{-1}(Q)$, then $\mathcal{O}_{G(A)}$ is contained in the fiber $\overline{\phi}^{-1}(Q)$.*

Proof. The result follows from Corollary 2.3.2 but it is easy to see that

$$\begin{aligned} \overline{\phi}(M \cdot A) &= \overline{\phi}(M_d^{-1} A_1 M_1, \ldots, M_{d-1}^{-1} A_d M_d) \\ &= \sum_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d} \mathrm{Tr}\left((M_d M_d^{-1}) A_1^{i_1} (M_1 M_1^{-1}) \cdots A_d^{i_d}\right) e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_d}^{(d)} \\ &= \sum_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d} \mathrm{Tr}\left(A_1^{i_1} \cdots A_d^{i_d}\right) e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_d}^{(d)} = \overline{\phi}(A), \end{aligned}$$

for every $M = (M_1, \ldots, M_d) \in G_{C_d,\mathbf{m}}$. □

Tangent space to the gauge orbit. We now study the tangent space to the gauge orbit and we characterize its elements in terms of the Lie algebra of the gauge group.

Let $(\Gamma, \mathbf{m}, \mathbf{n})$ be a tensor network. In Section 2.3 we denoted the Lie algebra of $\mathcal{G}_{\Gamma,\mathbf{m}}$ by $\mathbf{g}_{\Gamma,\mathbf{m}}$. In the case of matrix product states it is

$$\mathbf{g}_{C_d,\mathbf{m}} = \bigoplus_{i=1}^{d} \mathfrak{sl}_{m_i},$$

where $\mathfrak{sl}_m$ is the Lie algebra of $m \times m$ complex matrices with trace zero [HOV13].

The Lie algebra of the gauge group $G_{C_d,\mathbf{m}}$ is denoted by $\mathfrak{g}_{C_d,\mathbf{m}}$ and is

$$\mathfrak{g}_{C_d,\mathbf{m}} = \bigoplus_{i=1}^{d} \mathrm{End}_{m_i}\,. \tag{5.27}$$

Consider the action of $G_{C_d,\mathbf{m}}$ on D, given in Definition 5.7.2. We define the orbit map $\vartheta^{(A)} : G_{C_d,\mathbf{m}} \to D$ by $\vartheta^{(A)}(M) = M \cdot A$. The image of the map is clearly the orbit $\mathcal{O}_{G(A)}$ of $A \in D$ under the given action. Denote the identity element of the group by $e = \mathrm{Id}_{G_{C_d,\mathbf{m}}}$. The surjective map $\vartheta^{(A)} : G_{C_d,\mathbf{m}} \to \mathcal{O}_{G(A)}$ induces the map

$$d(\vartheta^{(A)})_e : T_e G_{C_d,\mathbf{m}} \simeq \mathfrak{g}_{C_d,\mathbf{m}} \to T_A \mathcal{O}_{G(A)},$$

where $\mathfrak{g}_{C_d,\mathbf{m}}$ is the Lie algebra of $G_{C_d,\mathbf{m}}$, given in Equation (5.27).

Let $\mathfrak{m} = (\mathfrak{m}_1, \ldots, \mathfrak{m}_d) \in \mathfrak{g}_{C_d,\mathbf{m}}$. For every component $\mathfrak{m}_k \in \mathrm{End}_{m_k}$ consider the 1-parametric subgroup

$$\begin{aligned} g_k : \mathbb{R} &\to GL_{m_k} \\ t &\mapsto \exp(t\mathfrak{m}_k). \end{aligned}$$

For every $k = 1, \ldots, d$, $g_k(t) = \exp(t\mathfrak{m}_k)$ is a smooth curve in GL_{m_k} passing through the identity $g_k(0) = \mathrm{Id}_{m_k}$ of End_{m_k}, and such that $g_k'(0) = \mathfrak{m}_k$.

Denote $\gamma(t) = (g_1(t), \ldots, g_d(t))$. It is a smooth curve in $G_{C_d,\mathbf{m}}$ passing through the identity $e \in G_{C_d,\mathbf{m}}$, and such that $\gamma'(0) = \mathfrak{m}$. For each $A \in D$ we have [Lee13]

$$d\vartheta_e^{(A)}(\mathfrak{m}) = (\vartheta^{(A)} \circ \gamma)'(0) = \frac{d}{dt}_{|t=0} \gamma(t) \cdot A := v_A \in T_A \mathcal{O}_{G(A)}.$$

For every $A \in D$, $k = 1, \ldots, d$ and $s = 1, \ldots, n_k$, we compute the action:

$$\begin{aligned} g_{k-1}(t)^{-1} A_k^s g_k(t) &= \exp(t\mathfrak{m}_{k-1})^{-1} A_k^s \exp(t\mathfrak{m}_k) \\ &= (\mathrm{Id}_{k-1} - t\mathfrak{m}_{k-1} + O(t^2)) A_k^s (\mathrm{Id}_k + t\mathfrak{m}_k + O(t^2)) \\ &= A_k^s - t\mathfrak{m}_{k-1} A_k^s + t A_k^s \mathfrak{m}_k + O(t^w), \quad w \geq 2 \\ &= A_k^s + t[A_k^s \mathfrak{m}_k - \mathfrak{m}_{k-1} A_k^s] + O(t^w), \quad w \geq 2. \end{aligned}$$

Then the infinitesimal generator of the action is

$$\frac{d}{dt}_{|t=0}(g_{k-1}(t)^{-1}A_k^s g_k(t)) = A_k^s \mathfrak{m}_k - \mathfrak{m}_{k-1}A_k^s.$$

This defines the action of $\mathfrak{m} = (\mathfrak{m}_1, \dots, \mathfrak{m}_d) \in \mathfrak{g}_{C_d,\mathbf{m}}$ on $A \in D$ given by

$$\mathfrak{m} \cdot A = (A_k^s \mathfrak{m}_k - \mathfrak{m}_{k-1}A_k^s)_{\substack{k=1,\dots,d \\ s=1,\dots,n_k}} \in T_A\mathcal{O}_{G(A)}, \tag{5.28}$$

and therefore the expression of the generic element of the tangent space to the gauge orbit at the point $A \in D$.

Define $\mathcal{N}_A = \ker(d\overline{\phi}_A)$ and $\mathcal{B}_A = (\mathcal{N}_A)^\perp$ and consider the following decomposition of the tangent space for matrix product states

$$T_A D = \mathcal{N}_A \oplus \mathcal{B}_A. \tag{5.29}$$

Lemma 5.7.4. *The tangent space $T_A\mathcal{O}_{G(A)} \subseteq D$ is contained in $\mathcal{N}_A$, for every $A \in D$.*

Proof. It is straightforward to see that

$$V = (A_k^s \mathfrak{m}_k - \mathfrak{m}_{k-1}A_k^s)_{\substack{k=1,\dots,d \\ s=1,\dots,n_k}} \in \ker(d\overline{\phi}_A),$$

substituting the expression of V in place of B in Equation (5.26):

$$\begin{aligned}
d\overline{\phi}_A(v)_{i_1,\dots,i_d} &= \sum_{k=1}^{d} \operatorname{Tr}\left(A_1^{i_1} \cdots \left(A_k^{i_k}\mathfrak{m}_k - \mathfrak{m}_{k-1}A_k^{i_k}\right) \cdots A_d^{i_d}\right) \\
&= \operatorname{Tr}\left((A_1^{i_1}\mathfrak{m}_1 - \mathfrak{m}_d A_1^{i_1})A_2^{i_2} \cdots A_d^{i_d}\right) + \operatorname{Tr}\left(A_1^{i_1}(A_2^{i_2}\mathfrak{m}_2 - \mathfrak{m}_1 A_2^{i_2}) \cdots A_d^{i_d}\right) + \\
&\quad + \cdots + \operatorname{Tr}\left(A_1^{i_1} \cdots A_{d-1}^{i_{d-1}}(A_d^{i_d}\mathfrak{m}_d - \mathfrak{m}_{d-1}A_d^{i_d})\right) \\
&= \operatorname{Tr}\Big(A_1^{i_1}\mathfrak{m}_1 A_2^{i_2} \cdots A_d^{i_d} - \mathfrak{m}_d A_1^{i_1} A_2^{i_2} \cdots A_d^{i_d} + A_1^{i_1} A_2^{i_2}\mathfrak{m}_2 \cdots A_d^{i_d} - A_1^{i_1}\mathfrak{m}_1 A_2^{i_2} \cdots A_d^{i_d} + \\
&\quad + \cdots + A_1^{i_1} \cdots A_{d-1}^{i_{d-1}} A_d^{i_d}\mathfrak{m}_d - A_1^{i_1} \cdots A_{d-1}^{i_{d-1}}\mathfrak{m}_{d-1}A_d^{i_d}\Big) \\
&= \operatorname{Tr}\Big(-\mathfrak{m}_d A_1^{i_1} A_2^{i_2} \cdots A_d^{i_d} + A_1^{i_1} \cdots A_{d-1}^{i_{d-1}} A_d^{i_d}\mathfrak{m}_d\Big) = 0;
\end{aligned}$$

since for every $k \in \{1, \dots, d\}$ the first term of the $(k-1)$-th tensor $A_{k-1}^{i_{k-1}}$ cancels out with the second one of the k-th tensor $A_k^{i_k}$. □

5.7.2 Open boundary conditions with bond dimension two

We compute the tangent space to the gauge orbit in the case of $\mathrm{MPS}(2, n, d)$ the variety of matrix product states with open boundary conditions and bond dimension two.

We start with a general consideration on the dimension of matrix product states with open boundary conditions. Fix $\Gamma = P_d$, the path graph on d vertices. Consider the variety of matrix product states with open boundary conditions, $\mathrm{MPS}(m, n, d)$, parametrized by the following map:

$$\begin{aligned}\phi :&\mathbb{C}^{n\times m} \times (\bigtimes_{i=1}^{d-2} \mathbb{C}^{n\times m\times m}) \times \mathbb{C}^{m\times n} \to (\mathbb{C}^n)^{\otimes d}\\ &(A_1, A_2, \ldots, A_{d-1}, A_d) \mapsto \sum_{i_1,\ldots,i_d=1}^{n} A_1^{i_1}\cdots A_d^{i_d}\ e_{i_1}\otimes\cdots\otimes e_{i_d},\end{aligned}$$

where $\{e_{i_j} : j = 1, \ldots, d, i_j = 1, \ldots, n\}$ is the canonical basis of the d copies of $\mathbb{C}^n$. Denote the domain by $D_{\mathrm{MPS}} \simeq \mathbb{C}^{2nm+(d-2)nm^2}$ and its dimension by $N_{\mathrm{MPS}} := 2nm+(d-2)nm^2$. The gauge group is $G_{\Gamma,m} = \bigtimes_{k=1}^{d-1} GL_m$ and its Lie algebra $\mathfrak{g}_{\Gamma,m} = \bigoplus_{k=1}^{d-1} \mathrm{End}_m$.

As we have already highlighted in Remark 5.2.3, our formula of the dimension of the variety coincides with Equation (5.8), given in [HMOV14]. More precisely:

$$\begin{aligned}&\dim \mathrm{MPS}(\mathbf{m}, \mathbf{n}, d) = \dim \mathcal{D}_{\mathrm{MPS}(\mathbf{m},\mathbf{n},d)} - \dim \mathcal{G}_{\Gamma,m}\\ &= \dim \mathbb{P}^{n_1m_1-1} + \dim \mathbb{P}^{n_dm_{d-1}-1} + \dim \mathbb{P}^{n_im_{i-1}m_i-1} + 1 - \sum_{i=2}^{d-1} \dim \mathbb{P}(GL_{m_i})\\ &= (n_1m_1) + (n_dm_{d-1}) + \sum_{i=2}^{d-1}(n_im_{i-1}m_i) - (d-1) - \sum_{i=1}^{d-1} m_i^2 + (d-1)\\ &= \sum_{i=1}^{d} n_im_{i-1}m_i - \sum_{i=1}^{d-1} m_i^2 = N_{\mathrm{MPS}(\mathbf{m},\mathbf{n},d)} - G_{\Gamma,m},\end{aligned}$$

where we are considering all affine dimensions. By a dimension count, for matrix product states with open boundary conditions, decomposition (5.29) can be written as

$$T_A D_{\mathrm{MPS}(\mathbf{m},\mathbf{n},d)} = T_A\mathcal{O}_{G(A)} \oplus \mathcal{B}_A,$$

with the following isomorphism:

$$T_A\mathcal{O}_{G(A)} \simeq \mathfrak{g}_{P_d,\mathbf{m}}, \tag{5.30}$$

where $\mathcal{B}_A = (T_A\mathcal{O}_{G(A)})^\perp = (\ker(\overline{\phi}_A))^\perp$ and $\dim \mathcal{B}_A = \dim \mathrm{MPS}(\mathbf{m}, \mathbf{n}, d)$.

Bond dimensions two. Now, consider the parametrization of matrix product states with open boundary conditions, $\mathrm{MPS}(2,n,d)$, with $m=2$ fixed:

$$\begin{aligned}\phi :&\mathbb{C}^{n\times 2}\times (\bigtimes_{i=1}^{d-2}\mathbb{C}^{n\times 2\times 2})\times \mathbb{C}^{2\times n}\to (\mathbb{C}^n)^{\otimes d}\\ &(A_1,A_2,\dots,A_{d-1},A_d)\mapsto \sum_{i_1,\dots,i_d=1}^{n} A_1^{i_1}\cdots A_d^{i_d}\ e_{i_1}\otimes\cdots\otimes e_{i_d}.\end{aligned}$$

The domain is denoted by $D_{\mathrm{MPS}}\simeq\mathbb{C}^{4n(d-1)}$ and its dimension is defined as $N_{\mathrm{MPS}}=4n(d-1)$. The gauge group is $G_{\Gamma,2}=\bigtimes_{k=1}^{d-1}GL_2$ and its Lie algebra $\mathfrak{g}_{\Gamma,2}=\bigoplus_{k=1}^{d-1}\mathrm{End}_2$.

The dimension of the variety is given in Example 5.2.4 and it is

$$\dim \mathrm{MPS}(2,n,d)=4(nd-d-n+1)=4(n-1)(d-1)=N_{\mathrm{MPS}}-\dim G_{\Gamma,2}.$$

Let $\mathfrak{m}=(\mathfrak{m}_1,\dots,\mathfrak{m}_d)\in\mathfrak{g}_{\Gamma,2}$ and $A\in D_{\mathrm{MPS}}$. For every $\mathfrak{m}_i\in\mathrm{End}_2$, $i=1,\dots,d-1$, we fix the canonical basis and write $\mathfrak{m}_i=(x^i_{\alpha\beta})$ for $i=1,\dots,d-1$ and $\alpha,\beta=1,2$. An element $A\in D_{\mathrm{MPS}}$ is a collection of two boundary tensors $A_1=(a_{1\,j}^{\,s})$ and $A_d=(a_{d\,s}^{\,j})$, for $s=1,\dots,n$ and $j=1,2$; and the inner tensors $A_i=(a_{i\,\alpha\beta}^{\,s})$, for $i=2,\dots,d-1$, $s=1,\dots,n$ and $\alpha,\beta=1,2$.

We compute the coordinates of $\mathfrak{m}\cdot A\in T_A\mathcal{O}_{G(A)}$. More precisely, by Equation (5.28), $\mathfrak{m}\cdot A=(V_k^s)$ for $k=1,\dots,d$ and $s=1,\dots,n$ is given by

$$\begin{aligned}V_1^s&=A_1^s\mathfrak{m}_1=\mathfrak{m}_1^tA_1^s,\\ V_i^s&=A_i^s\mathfrak{m}_i-\mathfrak{m}_{i-1}A_i^s,\quad i=2,\dots,d-1,\\ V_d^s&=-\mathfrak{m}_{d-1}A_d^s.\end{aligned}$$

We start computing V_i^s, for every $i=2,\dots,d-1$ and $s=1,\dots,n$

$$\begin{aligned}V_i^s&=\begin{pmatrix}a_{i\,11}^{\,s}&a_{i\,12}^{\,s}\\ a_{i\,21}^{\,s}&a_{i\,22}^{\,s}\end{pmatrix}\cdot\begin{pmatrix}x_{11}^i&x_{12}^i\\ x_{21}^i&x_{22}^i\end{pmatrix}-\begin{pmatrix}x_{11}^{i-1}&x_{12}^{i-1}\\ x_{21}^{i-1}&x_{22}^{i-1}\end{pmatrix}\cdot\begin{pmatrix}a_{i\,11}^{\,s}&a_{i\,12}^{\,s}\\ a_{i\,21}^{\,s}&a_{i\,22}^{\,s}\end{pmatrix}\\ &=\begin{pmatrix}(x_{11}^i-x_{11}^{i-1})a_{i\,11}^{\,s}-x_{12}^{i-1}a_{i\,21}^{\,s}+x_{21}^ia_{i\,12}^{\,s}&(x_{22}^i-x_{11}^{i-1})a_{i\,12}^{\,s}-x_{12}^{i-1}a_{i\,22}^{\,s}+x_{12}^ia_{i\,11}^{\,s}\\ -x_{21}^{i-1}a_{i\,11}^{\,s}+(x_{11}^i-x_{22}^{i-1})a_{i\,21}^{\,s}+x_{21}^ia_{i\,22}^{\,s}&-x_{21}^{i-1}a_{i\,12}^{\,s}+(x_{22}^i-x_{22}^{i-1})a_{i\,22}^{\,s}+x_{12}^ia_{i\,21}^{\,s}\end{pmatrix}.\end{aligned}$$

Reshaping the matrix $V_i^s\in\mathbb{C}^{2\times 2}$ into a vector, we obtain $t_i^s\in\mathbb{C}^4$:

$$t_i^s=\begin{pmatrix}(x_{11}^i-x_{11}^{i-1})a_{i\,11}^{\,s}-x_{12}^{i-1}a_{i\,21}^{\,s}+x_{21}^ia_{i\,12}^{\,s}\\ (x_{22}^i-x_{11}^{i-1})a_{i\,12}^{\,s}-x_{12}^{i-1}a_{i\,22}^{\,s}+x_{12}^ia_{i\,11}^{\,s}\\ -x_{21}^{i-1}a_{i\,11}^{\,s}+(x_{11}^i-x_{22}^{i-1})a_{i\,21}^{\,s}+x_{21}^ia_{i\,22}^{\,s}\\ -x_{21}^{i-1}a_{i\,12}^{\,s}+(x_{22}^i-x_{22}^{i-1})a_{i\,22}^{\,s}+x_{12}^ia_{i\,21}^{\,s}\end{pmatrix}.$$

Define

$$K_i(\mathfrak{m}_{i-1},\mathfrak{m}_i):=\begin{pmatrix} x_{11}^i-x_{11}^{i-1} & x_{21}^i & -x_{12}^{i-1} & 0\\ x_{12}^i & x_{22}^i-x_{11}^{i-1} & 0 & -x_{12}^{i-1}\\ -x_{21}^{i-1} & 0 & x_{11}^i-x_{22}^{i-1} & x_{21}^i\\ 0 & -x_{21}^{i-1} & x_{12}^i & x_{22}^i-x_{22}^{i-1}\end{pmatrix}\in\mathbb{C}^{2\times 2}$$

and denote the vector of entries of A_i^s, $i=2,\dots,d-1$ and $s=1,\dots,n$ by

$$a_i{}^s=(a_i{}_{11}^s,a_i{}_{12}^s,a_i{}_{21}^s,a_i{}_{22}^s)^t\in\mathbb{C}^4.$$

Then $t_i^s=K_i(\mathfrak{m}_{i-1},\mathfrak{m}_i)a_i{}^s$, for every $i=2,\dots,d-1$ and $s=1,\dots,n$.

We compute now the coordinates of the action on the boundary tensors, i.e. the vectors V_1^s and V_d^s, for every $s=1,\dots,n$:

$$V_1^s=A_1^s\mathfrak{m}_1=\begin{pmatrix}a_1{}_1^s & a_1{}_2^s\end{pmatrix}\begin{pmatrix}x_{11}^1 & x_{12}^1\\ x_{21}^1 & x_{22}^1\end{pmatrix}=\begin{pmatrix}x_{11}^1 & x_{21}^1\\ x_{12}^1 & x_{22}^1\end{pmatrix}\begin{pmatrix}a_1{}_1^s\\ a_1{}_2^s\end{pmatrix},$$

$$V_d^s=-\mathfrak{m}_{d-1}A_d^s=\begin{pmatrix}-x_{11}^{d-1} & -x_{12}^{d-1}\\ -x_{21}^{d-1} & -x_{22}^{d-1}\end{pmatrix}\begin{pmatrix}a_d{}_1^s\\ a_d{}_2^s\end{pmatrix},$$

where $a_i^s=(a_i{}_1^s,a_i{}_2^s)^t\in\mathbb{C}^2$, for $i=1,d-1$. Define the matrices

$$K_1(\mathfrak{m}_1):=\begin{pmatrix}x_{11}^1 & x_{21}^1\\ x_{12}^1 & x_{22}^1\end{pmatrix},\quad K_d(\mathfrak{m}_{d-1}):=\begin{pmatrix}-x_{11}^{d-1} & -x_{12}^{d-1}\\ -x_{21}^{d-1} & -x_{22}^{d-1}\end{pmatrix}.$$

Then $V_1^s=K_1(\mathfrak{m}_1)a_1{}^s$ and $V_d^s=K_d(\mathfrak{m}_{d-1})a_d{}^s$.

We define a block diagonal matrix of order N_{MPS}:

$$\begin{aligned}\mathcal{M}(\mathfrak{m}):=\mathrm{diag}\{&\underbrace{K_1(\mathfrak{m}_1),\dots,K_1(\mathfrak{m}_1)}_{n\text{-times}},\underbrace{K_2(\mathfrak{m}_1,\mathfrak{m}_2),\dots,K_2(\mathfrak{m}_1,\mathfrak{m}_2)}_{n\text{-times}}\dots\\ &\underbrace{K_{d-1}(\mathfrak{m}_{d-2},\mathfrak{m}_{d-1}),\dots,K_{d-1}(\mathfrak{m}_{d-2},\mathfrak{m}_{d-1})}_{n\text{-times}},\underbrace{K_d(\mathfrak{m}_{d-1}),\dots,K_d(\mathfrak{m}_{d-1})}_{n\text{-times}}\},\end{aligned}$$

that is the matrix representation of the action of $\mathfrak{m}\in\mathfrak{g}_{\Gamma,2}$ on $A\in D_{\mathrm{MPS}}$. If we denote by $a\in\mathbb{C}^{N_{\mathrm{MPS}}}$ the column vector of all the entries of A, then

$$\begin{aligned}\mathcal{M}(\mathfrak{m}):D_{\mathrm{MPS}}\simeq T_AD_{\mathrm{MPS}}&\to T_A\mathcal{O}_{G(A)}\subset T_AD_{\mathrm{MPS}}\\ a&\mapsto\mathcal{M}(\mathfrak{m})a.\end{aligned}$$

Since the tangent space of the gauge orbit at the point $A\in D_{\mathrm{MPS}}$ is parametrized by the Lie algebra $\mathfrak{g}_{\Gamma,2}$ via isomorphism (5.30), we choose a basis $\mathcal{B}$ for the $\mathfrak{g}_{\Gamma,2}$ in order to find a basis of $T_A\mathcal{O}_{G(A)}$, for every $A\in D_{\mathrm{MPS}}$.

Denote by $E_{i,j} \in \mathrm{End}_2$ the matrix with 1 in the (i,j)-th entry and zero otherwise. The set $\mathcal{E} = \{E_{\alpha\beta}\}_{\alpha,\beta=1,2}$ gives a basis of End_2. The tangent space to the gauge orbit is spanned by the image of $\mathcal{M}(\mathfrak{m})$, for $\mathfrak{m}$ varying in $\mathcal{B} = \bigtimes_{i=1}^{d} \mathcal{E}$:

$$T_A(\mathcal{O}_{G(A)}) = \langle \mathcal{M}(\mathfrak{m})a \rangle_{\mathfrak{m}\in\mathcal{B}}.$$

For $\mathfrak{m}_{i-1} = E_{11}, E_{12}, E_{21}, E_{22}$ we have

$$K_i(E_{11},0){a_i}^s = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} -a_{i11}^s \\ -a_{i12}^s \\ 0 \\ 0 \end{pmatrix},$$

$$K_i(E_{12},0){a_i}^s = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} -a_{i21}^s \\ -a_{i22}^s \\ 0 \\ 0 \end{pmatrix},$$

$$K_i(E_{21},0){a_i}^s = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -a_{i11}^s \\ -a_{i12}^s \end{pmatrix},$$

$$K_i(E_{22},0){a_i}^s = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -a_{i21}^s \\ -a_{i22}^s \end{pmatrix},$$

respectively. For $\mathfrak{m}_i = E_{11}, E_{12}, E_{21}, E_{22}$ we have

$$K_i(0,E_{11}){a_i}^s = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} a_{i11}^s \\ 0 \\ a_{i21}^s \\ 0 \end{pmatrix},$$

$$K_i(0,E_{12}){a_i}^s = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} 0 \\ a_{i11}^s \\ 0 \\ a_{i21}^s \end{pmatrix},$$

$$K_i(0,E_{21}){a_i}^s = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} a_{i12}^s \\ 0 \\ a_{i22}^s \\ 0 \end{pmatrix},$$

$$K_i(0,E_{22}){a_i}^s = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{i11}^s \\ a_{i12}^s \\ a_{i21}^s \\ a_{i22}^s \end{pmatrix} = \begin{pmatrix} 0 \\ a_{i12}^s \\ 0 \\ a_{i22}^s \end{pmatrix},$$

respectively. For the boundary vectors, for $\mathfrak{m}_1 = E_{11}, E_{12}, E_{21}, E_{22}$ we have

$$K_1(E_{11}){a_1}^s = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} {a_1}_1^s \\ {a_1}_2^s \end{pmatrix} = \begin{pmatrix} {a_1}_1^s \\ 0 \end{pmatrix},$$

$$K_1(E_{12}){a_1}^s = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} {a_1}_1^s \\ {a_1}_2^s \end{pmatrix} = \begin{pmatrix} 0 \\ {a_1}_1^s \end{pmatrix},$$

$$K_1(E_{21}){a_1}^s = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} {a_1}_1^s \\ {a_1}_2^s \end{pmatrix} = \begin{pmatrix} {a_1}_2^s \\ 0 \end{pmatrix},$$

$$K_1(E_{22}){a_1}^s = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} {a_1}_1^s \\ {a_1}_2^s \end{pmatrix} = \begin{pmatrix} 0 \\ {a_1}_2^s \end{pmatrix},$$

respectively, and for $\mathfrak{m}_{d-1} = E_{11}, E_{12}, E_{21}, E_{22}$ we have

$$K_d(E_{11}){a_d}^s = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} {a_d}_1^s \\ {a_d}_2^s \end{pmatrix} = \begin{pmatrix} -{a_d}_1^s \\ 0 \end{pmatrix},$$

$$K_d(E_{12}){a_d}^s = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} {a_d}_1^s \\ {a_d}_2^s \end{pmatrix} = \begin{pmatrix} -{a_d}_2^s \\ 0 \end{pmatrix},$$

$$K_d(E_{21}){a_d}^s = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} {a_d}_1^s \\ {a_d}_2^s \end{pmatrix} = \begin{pmatrix} 0 \\ -{a_d}_1^s \end{pmatrix},$$

$$K_d(E_{22}){a_d}^s = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} {a_d}_1^s \\ {a_d}_2^s \end{pmatrix} = \begin{pmatrix} 0 \\ -{a_d}_2^s \end{pmatrix},$$

respectively. We can define the following matrices

$$\begin{aligned} m_L^s(1) &:= \begin{pmatrix} {a_1}_1^s & 0 & {a_1}_2^s & 0 \\ 0 & {a_1}_1^s & 0 & {a_1}_2^s \end{pmatrix} = {a_1}^s \boxtimes \mathrm{Id}_2, \\ m^s(i) &:= \begin{pmatrix} -{a_i}_{11}^s & -{a_i}_{21}^s & 0 & 0 & {a_i}_{11}^s & 0 & {a_i}_{12}^s & 0 \\ -{a_i}_{12}^s & -{a_i}_{22}^s & 0 & 0 & 0 & {a_i}_{11}^s & 0 & {a_i}_{12}^s \\ 0 & 0 & -{a_i}_{11}^s & -{a_i}_{21}^s & {a_i}_{21}^s & 0 & {a_i}_{22}^s & 0 \\ 0 & 0 & -{a_i}_{12}^s & -{a_i}_{22}^s & 0 & {a_i}_{21}^s & 0 & {a_i}_{22}^s \end{pmatrix} \\ &= \big(-\mathrm{Id}_2 \boxtimes (A_i^s)^t \quad (A_i^s) \boxtimes \mathrm{Id}_2 \big), \quad i \in \{2, \dots, d-1\} \\ m_R^s(d) &:= \begin{pmatrix} -{a_d}_1^s & -{a_d}_2^s & 0 & 0 \\ 0 & 0 & -{a_d}_1^s & -{a_d}_2^s \end{pmatrix} = -\mathrm{Id}_2 \boxtimes {a_d}^s, \end{aligned}$$

that are the building blocks that constitute the matrix, that we will denote by $\mathbb{T}_A$, whose columns generate (and provide a basis of) $T_A(\mathcal{O}_{G(A)})$, for every $A \in D_{\mathrm{MPS}}$. In particular, $\mathbb{T}_A$ is a $N_{\mathrm{MPS}} \times \dim G_{\Gamma,m}$ complex matrix.

Example 5.7.5. The tangent space to the gauge orbit of MPS$(2,3,3)$ is

$$\mathbb{T}_A = \begin{pmatrix} {a_1}^1 \boxtimes \mathrm{Id}_2 & \mathbb{O}_{2\times 4} \\ {a_1}^2 \boxtimes \mathrm{Id}_2 & \mathbb{O}_{2\times 4} \\ {a_1}^3 \boxtimes \mathrm{Id}_2 & \mathbb{O}_{2\times 4} \\ -\mathrm{Id}_2 \boxtimes (A_2^1)^t & A_2^1 \boxtimes \mathrm{Id}_2 \\ -\mathrm{Id}_2 \boxtimes (A_2^2)^t & A_2^2 \boxtimes \mathrm{Id}_2 \\ -\mathrm{Id}_2 \boxtimes (A_2^3)^t & A_2^3 \boxtimes \mathrm{Id}_2 \\ \mathbb{O}_{2\times 4} & -\mathrm{Id}_2 \boxtimes {a_3}^1 \\ \mathbb{O}_{2\times 4} & -\mathrm{Id}_2 \boxtimes {a_3}^2 \\ \mathbb{O}_{2\times 4} & -\mathrm{Id}_2 \boxtimes {a_3}^3, \end{pmatrix}$$

which is a 24×8 matrix.

Remark 5.7.6 (Orthonormal complement)**.** In the next chapter, we will need a orthonormal basis of the vector space $\mathcal{B}_A = (T_A(\mathcal{O}_{G(A)}))^\perp$, for every $A \in D_{\mathrm{MPS}}$.
We will numerically compute the QR decomposition of the matrix $\mathbb{T}_A$, i.e. $\mathbb{T}_A = Q_A R_A$. The columns of Q_A give an orthonormal basis of $T_A(\mathcal{O}_{G(A)})$. Therefore, the orthonormal completion of a basis of $T_A D_{\mathrm{MPS}} \simeq D_{\mathrm{MPS}}$ provides the vectors which span $\mathcal{B}_A$. In order to obtain the completion, we will compute $P_A = \ker(Q_A^\dagger)$ whose columns span the space $\mathcal{B}_A$ for $A \in D_{\mathrm{MPS}}$.

5.7.3 Homogeneous case

Analogous computations are done in the case of hMPS$(2,n,d)$ homogeneous matrix product states with open boundary conditions and bond dimension two. However, in this case, the use of *affine* coordinates gives rise to redundant scalar degrees of freedom. This will be made clear in Remark 5.7.7. As a consequence, the kernel of the differential of the parametrization contains at least one vector space which is not generated by the gauge Lie algebra action.

Fix $\Gamma = P_{d+2}$, the path graph on $d+2$ vertices. Consider the parametrization (5.9) of homogeneous matrix product states with open boundary conditions hMPS(m,n,d):

$$\begin{aligned} \psi : \mathbb{C}^m \times (\mathbb{C}^{m\times m})^{\times n} \times \mathbb{C}^m &\to (\mathbb{C}^n)^{\otimes d} \\ p = (v_L, A, v_R) &\mapsto \sum_{\substack{i_j=1,\\ j=1,\dots,d}}^{n} \Big(v_L^\dagger A_{i_1} \cdots A_{i_d} v_R\Big) e_{i_1} \otimes \cdots \otimes e_{i_d}, \end{aligned}$$

Denote the domain by

$$D_{\mathrm{hMPS}} = \mathbb{C}^m \times (\mathbb{C}^{m\times m})^{\times n} \times \mathbb{C}^m \simeq \mathbb{C}^{2m+m^2 n}.$$

The gauge group is $G_{\Gamma,m} = GL_m$ and its Lie algebra $\mathfrak{g}_{\Gamma,m} = \mathrm{End}_m$. Given $\mathfrak{m} \in \mathfrak{g}_{\Gamma,m}$ and $p = (v_L, A, v_R) \in D_{\mathrm{hMPS}}$, the action of the Lie algebra (5.28) specializes to

$$\mathfrak{m} \cdot (v_L, A, v_R) = (v_L^\dagger \mathfrak{m}, A^s \mathfrak{m} - \mathfrak{m} A^s, -\mathfrak{m} v_R) \quad \text{for } s = 1, \dots, n. \tag{5.31}$$

Remark 5.7.7. The upper bound on the dimension of the variety, given in Equation (5.10), is the following:

$$\begin{aligned}\dim \mathrm{hMPS}(m,n,d) &\leq 2\dim \mathbb{P}^{m-1} + \dim \mathbb{P}^{m^2n-1} + 1 - \dim \mathbb{P}^{m^2} \\ &= 2(m-1) + (m^2n-1) + 1 - (m^2-1) \\ &= 2m + m^2(n-1) - 1.\end{aligned}$$

Define the upper bound on the dimension by

$$\mathrm{expdim}\,\mathrm{hMPS}(m,n,d) := 2m + m^2(n-1) - 1.$$

On the other hand, when working in the affine setting we have $N_{\mathrm{hMPS}} = \dim D_{\mathrm{hMPS}} = 2m + m^2 n$ and $\dim G_{\Gamma,m} = m^2$. Therefore the vector space $(T_p\mathcal{O}_{G(p)})^\perp$ for $p \in D_{\mathrm{hMPS}}$ has dimension

$$\begin{aligned}\dim(T_p\mathcal{O}_{G(p)})^\perp &= N_{\mathrm{hMPS}} - \dim G_{\Gamma,m} = 2m + m^2n - m^2 = 2m + m^2(n-1) \\ &= \mathrm{expdim}\,\mathrm{hMPS}(m,n,d) + 1.\end{aligned}$$

The 1-dimensional gap is due to the overabundant affine coordinates that we are considering in the domain of the parametrization. The kernel of the differential of the map, $\ker(d\psi_p)$ for $p \in D_{\mathrm{hMPS}(m,n,d)}$, contains at least one vector space which is not generated by the gauge Lie algebra action. We denote this vector space by $S_p \simeq \mathbb{C}$ and we have $S_p \subset \ker(d\psi_p) \setminus T_p\mathcal{O}_{G(p)}$. We will provide this vector space order to obtain

$$T_p D_{\mathrm{hMPS}} = \left(T_p\mathcal{O}_{G(p)} \oplus S_p\right) \oplus \mathcal{B}_p,$$

with $\dim \mathcal{B}_p = \mathrm{expdim}\,\mathrm{hMPS}(m,n,d)$.

Bond dimensions two. We fix now $m = 2$. Let $\mathfrak{m} \in \mathfrak{g}_{\Gamma,2} = \mathrm{End}_2$ and $p \in D_{\mathrm{hMPS}} \simeq \mathbb{C}^{4(n+1)}$. We fix the canonical basis and write $\mathfrak{m} = (x_{\alpha\beta})_{\alpha,\beta=1,2}$, $A = (a^s_{\alpha\beta})^{s=1,\dots,n}_{\alpha,\beta=1,2}$, and the boundary tensors $v_L = (\overline{v}^1_L, \overline{v}^2_L)^t$, $v_R = (v^1_R, v^2_R)^t$.

We compute the coordinates of $\mathfrak{m}\cdot p \in T_p\mathcal{O}_{G(p)}$ which, by Equation (5.31) is given by $\mathfrak{m}\cdot p = (V_L, V_I, V_R)$ is given by

$$\begin{aligned}V_L &= v_L^\dagger \mathfrak{m} = \mathfrak{m}^t \overline{v}_L, \\ V_I^s &= A^s\mathfrak{m} - \mathfrak{m}A^s, \quad s = 1,\dots,n \\ V_R &= -\mathfrak{m}v_R,\end{aligned}$$

where $\overline{v}_L = (v^1_L, v^2_L)^t$.

We start computing V_I^s, for every $s = 1,\dots,n$

$$\begin{aligned}V_I^s &= \begin{pmatrix} a^s_{11} & a^s_{12} \\ a^s_{21} & a^s_{22}\end{pmatrix} \cdot \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22}\end{pmatrix} - \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22}\end{pmatrix} \cdot \begin{pmatrix} a^s_{11} & a^s_{12} \\ a^s_{21} & a^s_{22}\end{pmatrix} \\ &= \begin{pmatrix} -x_{12}a^s_{21} + x_{21}a^s_{12} & (x_{22}-x_{11})a^s_{12} - x_{12}a^s_{22} + x_{12}a^s_{11} \\ -x_{21}a^s_{11} + (x_{11}-x_{22})a^s_{21} + x_{21}a^s_{22} & -x_{21}a^s_{12} + x_{12}a^s_{21}\end{pmatrix}\end{aligned}$$

Reshaping the matrix $V_I^s \in \mathbb{C}^{2\times 2}$ into a vector, we obtain $t^s \in \mathbb{C}^4$:

$$t^s = \begin{pmatrix} -x_{12}a_{21}^s + x_{21}a_{12}^s \\ (x_{22}-x_{11})a_{12}^s - x_{12}a_{22}^s + x_{12}a_{11}^s \\ -x_{21}a_{11}^s + (x_{11}-x_{22})a_{21}^s + x_{21}a_{22}^s \\ -x_{21}a_{12}^s + x_{12}a_{21}^s \end{pmatrix}.$$

Define

$$K(\mathfrak{m}) := \begin{pmatrix} 0 & x_{21} & -x_{12} & 0 \\ x_{12} & x_{22}-x_{11} & 0 & -x_{12} \\ -x_{21} & 0 & x_{11}-x_{22} & x_{21} \\ 0 & -x_{21} & x_{12} & 0 \end{pmatrix},$$

and denote the vector reshape of $A^s \in \mathbb{C}^{2\times 2}$ for $s = 1, \ldots, n$ by

$$a^s = (a_{11}^s, a_{12}^s, a_{21}^s, a_{22}^s)^t \in \mathbb{C}^4.$$

Then $t^s = K(\mathfrak{m})a^s$ for every $s = 1, \ldots, n$.

The action on the boundary tensors is given by:

$$V_L = v_L^\dagger \mathfrak{m} = \begin{pmatrix} v_L^1 & v_L^2 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \end{pmatrix} \begin{pmatrix} v_L^1 \\ v_L^2 \end{pmatrix},$$

$$V_R = -\mathfrak{m} v_R = \begin{pmatrix} -x_{11} & -x_{12} \\ -x_{21} & -x_{22} \end{pmatrix} \begin{pmatrix} v_R^1 \\ v_R^2 \end{pmatrix},$$

where we recall that $v_L = (v_L^1, v_L^2)^t, v_R = (v_R^1, v_R^2)^t \in \mathbb{C}^2$. Define

$$K_L(\mathfrak{m}) := \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \end{pmatrix}, \quad K_R(\mathfrak{m}) := \begin{pmatrix} -x_{11} & -x_{12} \\ -x_{21} & -x_{22} \end{pmatrix}$$

then $V_L = K_L(\mathfrak{m})v_L$ and $V_R = K_R(\mathfrak{m})v_R$.

We define a block diagonal matrix of order N_{hMPS}

$$\mathcal{M}(\mathfrak{m}) = \mathrm{diag}\{K_L(\mathfrak{m}), \underbrace{K(\mathfrak{m}), \ldots, K(\mathfrak{m})}_{n\text{-times}}, K_R(\mathfrak{m})\},$$

which is the matrix representation of the action of $\mathfrak{m} \in \mathfrak{g}_{\Gamma,2} = \mathrm{End}_2$ on $p \in D_{\mathrm{hMPS}}$. If we denote the vector reshape of $A^s \in \mathbb{C}^{2\times 2}$ for $s = 1, \ldots, n$, by

$$a^s = (a_{11}^s, a_{12}^s, a_{21}^s, a_{22}^s)^t \in \mathbb{C}^4$$

and the vector reshape of $p = (v_L, A, v_R) \in D_{\mathrm{hMPS}}$ by

$$a = (v_L^1, v_L^2, a^1 \ \ldots \ a^n, v_R^1, v_R^2)^t \in \mathbb{C}^{4(n+1)},$$

then we can write

$$\begin{aligned}\mathcal{M}(\mathfrak{m}) : D \simeq T_pD &\to T_p\mathcal{O}_{G(p)} \subset T_pD \\ a &\mapsto \mathcal{M}(\mathfrak{m})a.\end{aligned}$$

Again, since the tangent space of the gauge orbit at the point $p \in D_{\text{hMPS}}$ is parametrized by the Lie algebra $\mathfrak{g}_{\Gamma,2}$, we fix a basis $\mathcal{B}$ of $\mathfrak{g}_{\Gamma,2}$ in order to determine a basis of $T_p\mathcal{O}_{G(p)}$, for every $p \in D$:

$$\mathcal{B} = \{\mathrm{Id}_2 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, H = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, X = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, Y = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\}.$$

The tangent space to the gauge orbit is spanned by the image of $\mathcal{M}(\mathfrak{m})$, for $\mathfrak{m}$ varying in $\mathcal{B}$

$$T_p(\mathcal{O}_{G(p)}) = \langle \mathcal{M}(\mathfrak{m}) \cdot a \rangle_{\mathfrak{m}\in\mathcal{B}}.$$

We compute the blocks associated to the internal tensor

$$\begin{aligned}
K(\mathrm{Id}_2)a^s &= \begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}a^s_{11}\\ a^s_{12}\\ a^s_{21}\\ a^s_{22}\end{pmatrix} = \begin{pmatrix}0\\ 0\\ 0\\ 0\end{pmatrix},\\
K(H)a^s &= \begin{pmatrix}0 & 0 & 0 & 0\\ 0 & -2 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}a^s_{11}\\ a^s_{12}\\ a^s_{21}\\ a^s_{22}\end{pmatrix} = \begin{pmatrix}0\\ -2a^s_{12}\\ 2a^s_{21}\\ 0\end{pmatrix},\\
K(X)a^s &= \begin{pmatrix}0 & 0 & -1 & 0\\ 1 & 0 & 0 & -1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}a^s_{11}\\ a^s_{12}\\ a^s_{21}\\ a^s_{22}\end{pmatrix} = \begin{pmatrix}-a^s_{21}\\ a^s_{11}-a^s_{22}\\ 0\\ a^s_{21}\end{pmatrix},\\
K(Y)a^s &= \begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ -1 & 0 & 0 & 1\\ 0 & -1 & 0 & 0\end{pmatrix}\begin{pmatrix}a^s_{11}\\ a^s_{12}\\ a^s_{21}\\ a^s_{22}\end{pmatrix} = \begin{pmatrix}a^s_{12}\\ 0\\ a^s_{22}-a^s_{11}\\ -a^s_{12}\end{pmatrix}.
\end{aligned}$$

We compute the blocks associated to the boundary vectors

$$K_L(\mathrm{Id}_2)v_L = \begin{pmatrix}v^1_L\\ v^2_L\end{pmatrix},\ K_L(H)v_L = \begin{pmatrix}v^1_L\\ -v^2_L\end{pmatrix},\ K_L(X)v_L = \begin{pmatrix}0\\ v^1_L\end{pmatrix},\ K_L(Y)v_L = \begin{pmatrix}v^2_L\\ 0\end{pmatrix},$$

$$K_R(\mathrm{Id}_2)v_R = \begin{pmatrix}-v^1_R\\ -v^2_R\end{pmatrix},\ K_R(H)v_R = \begin{pmatrix}-v^1_R\\ v^2_R\end{pmatrix},\ K_R(X)v_R = \begin{pmatrix}-v^2_R\\ 0\end{pmatrix},\ K_R(Y)v_R = \begin{pmatrix}0\\ -v^1_R\end{pmatrix}.$$

We define the building blocks

$$m_A^s := \begin{pmatrix} 0 & 0 & -a_{21}^s & a_{12}^s \\ 0 & -2a_{12}^s & a_{11}^s - a_{22}^s & 0 \\ 0 & 2a_{21}^s & 0 & a_{22}^s - a_{11}^s \\ 0 & 0 & a_{21}^s & -a_{12}^s \end{pmatrix}, \quad s = 1, \dots, n,$$

$$m_L := \begin{pmatrix} v_L^1 & v_L^1 & 0 & v_L^2 \\ v_L^2 & -v_L^2 & v_L^1 & 0 \end{pmatrix},$$

$$m_R := \begin{pmatrix} -v_R^1 & -v_R^1 & -v_R^2 & 0 \\ -v_R^2 & v_R^2 & 0 & -v_R^1 \end{pmatrix},$$

and, finally, we obtain the matrix whose columns give a basis of $T_p\mathcal{O}_{G(p)}$

$$\widetilde{\mathbb{T}}_p = \begin{pmatrix} m_L \\ m_A^1 \\ \vdots \\ m_A^n \\ m_R \end{pmatrix} = \begin{pmatrix} v_L^1 & v_L^1 & 0 & v_L^2 \\ v_L^2 & -v_L^2 & v_L^1 & 0 \\ 0 & 0 & -a_{21}^1 & a_{12}^1 \\ 0 & -2a_{12}^1 & a_{11}^1 - a_{22}^1 & 0 \\ 0 & 2a_{21}^1 & 0 & a_{22}^1 - a_{11}^1 \\ 0 & 0 & a_{21}^1 & -a_{12}^1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & -a_{21}^n & a_{12}^n \\ 0 & -2a_{12}^n & a_{11}^n - a_{22}^n & 0 \\ 0 & 2a_{21}^n & 0 & a_{22}^n - a_{11}^n \\ 0 & 0 & a_{21}^n & -a_{12}^n \\ -v_R^1 & -v_R^1 & -v_R^2 & 0 \\ -v_R^2 & v_R^2 & 0 & -v_R^1 \end{pmatrix} \in \mathbb{C}^{N_{\mathrm{hMPS}} \times 4}.$$

In order to find the 1-dimensional vector space $S_p \simeq \mathbb{C} \subset \ker(d\phi_p) \setminus T_p\mathcal{O}_{G(p)}$ for every $p \in D_{\mathrm{hMPS}}$, c.f. Remark 5.7.7, we notice that there is an action of $\mathbb{C}^*$ on the domain of the parametrization which leaves invariant the image of the map but that is not provided by the action of the gauge group $G_{\Gamma,2}$.

Consider a point $p \in D_{\mathrm{hMPS}}$ and $\alpha, \beta, \gamma \in \mathbb{C}^*$. Then

$$\psi(\alpha v_L, \beta A, \gamma v_R) = \alpha\beta^d\gamma \; \psi(v_L, A, v_R) = \psi(v_L, A, v_R)$$

if and only if $\alpha\beta^d\gamma = 1$. Fix $\alpha = \beta^{-d}$ and $\gamma = 1$, which satisfy the condition $\alpha\beta^d\gamma = 1$. We define an action of $\mathbb{C}^*$ on the domain D_{hMPS} of ψ as follows: given $p = (v_L, A, v_R) \in D_{\mathrm{hMPS}}$ and $\beta \in \mathbb{C}^*$, the action is given by

$$\beta \cdot p = (\beta^{-d} v_L, \beta A, v_R).$$

Let $\widetilde{\beta} \in \mathbb{C}$ such that $\beta(t) = \exp(t\widetilde{\beta})$ is a smooth curve in $\mathbb{C}^*$ passing through the identity of $\mathbb{C}$, $\beta(0) = 1$, and such that $\beta'(0) = \widetilde{\beta}$. The action of $\widetilde{\beta} \in \mathbb{C}$ on $p \in D_{\mathrm{hMPS}}$ is given by

$\widetilde{\beta}\cdot p = (V_L, V_I^1, V_I^n, \dots, V_R) \in \ker(d\psi_p)$, with

$$\begin{aligned} V_L &= \frac{d}{dt}_{|_{t=0}} \exp(t\widetilde{\beta})^{-d} v_L^\dagger = -d\widetilde{\beta}^{-d-1} v_L^\dagger, \\ V_I^s &= \frac{d}{dt}_{|_{t=0}} \exp(t\widetilde{\beta}) A^s = (\widetilde{\beta}\exp(t\widetilde{\beta})A^s)_{|_{t=0}} = \widetilde{\beta}A^s, \quad s = 1, \dots, n, \\ V_R &= \frac{d}{dt}_{|_{t=0}} \exp(0) v_R = 0. \end{aligned}$$

Fixing the generator $\widetilde{\beta} = 1 \in \mathbb{C}$ we obtain the vector

$$v_1(p) := (-dv_L^1, -dv_L^2, a_{11}^1, a_{12}^2, \dots, a_{22}^n, 0, 0)\ker(d\phi_p) \setminus T_p\mathcal{O}_{G(p)}.$$

and finally we define the matrix

$$\mathbb{T}_p = \begin{pmatrix} -dv_L^1 & v_L^1 & v_L^1 & 0 & v_L^2 \\ -dv_L^2 & v_L^2 & -v_L^2 & v_L^1 & 0 \\ a_{11}^1 & 0 & 0 & -a_{21}^1 & a_{12}^1 \\ \vdots & 0 & -2a_{12}^1 & a_{11}^1 - a_{22}^1 & 0 \\ a_{22}^1 & 0 & 2a_{21}^1 & 0 & a_{22}^1 - a_{11}^1 \\ a_{11}^2 & 0 & 0 & a_{21}^1 & -a_{12}^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{22}^{n-1} & 0 & 0 & -a_{21}^n & a_{12}^n \\ a_{11}^n & 0 & -2a_{12}^n & a_{11}^n - a_{22}^n & 0 \\ \vdots & 0 & 2a_{21}^n & 0 & a_{22}^n - a_{11}^n \\ a_{22}^n & 0 & 0 & a_{21}^n & -a_{12}^n \\ 0 & -v_R^1 & -v_R^1 & -v_R^2 & 0 \\ 0 & -v_R^2 & v_R^2 & 0 & -v_R^1 \end{pmatrix} \in \mathbb{C}^{N_{\mathrm{hMPS}} \times 5}. \tag{5.32}$$

whose columns span $S_p \oplus T_p\mathcal{O}_{G(p)} \subset D_{\mathrm{hMPS}}$.

Remark 5.7.8. Let $p \in D_{\mathrm{hMPS}}$. Denote the second column vector of the matrix $\mathbb{T}_p$ (5.32) by $v_2(p)$. The vector can be compactly written as $v_2(p) = (v_L, \mathbb{0}, -v_R)^t$. Notice that, fixing $\gamma = \alpha^{-1}$ and $\beta = 1$ is equivalent to the gauge action of the matrix $\alpha\mathrm{Id}_2$:

$$\psi(v_L \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}, \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha^{-1} \end{pmatrix} A \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}, \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha^{-1} \end{pmatrix} v_R) = \psi(\alpha v_L, A, \alpha^{-1} v_R),$$

and actually provides $v_2(p)$, which is the second column vector of $\mathbb{T}_p$ (5.32). This is the only scalar multiplication which derives from the *affine* gauge group $G_{\Gamma,2}$.

Moreover, if we consider the remaining choice, $\alpha = 1$ and $\gamma = \beta^{-d}$ (which is symmetric to $\alpha = \beta^{-d}$, $\gamma = 1$), we obtain the vector

$$s(p) = (0, A, -dv_R)\,,$$

which is a linear combination of $v_1(p)$ and $v_2(p)$, precisely $s(p) = v_1(p) + dv_2(p)$.

Conclusions. In this chapter, we have collected all the ingredients to implement a variational conjugate gradient method on matrix product states (with open boundary conditions). In the last part of the chapter, we analyze further the geometry of the fiber of the variety of matrix product states; in particular we study the role of the gauge orbit and of its tangent space. We describe a pointwise decomposition of the tangent space to the fiber that we specialize to the case of open boundary conditions. The decomposition will be used in the following chapter for finding a good representative of the gradient of the functional we want to minimize.

Chapter 6

Preliminary numerical calculations

In this chapter, we present preliminary numerical results on the performance of the nonlinear conjugate gradient method on the variational class of matrix product states with open boundary conditions. The physical model we use is the AKLT model, c.f. Sections 5.3 and 5.4. The code, implemented in MATLAB [Mat20] and available at `https://github.com/claudia-dela/NLCG_MPS_open-boundaries/`, makes use of the general tensor network contraction routine **ncon()** [PESV14]. The analysis is done with real random starting points but the code works also in the complex field.
In Section 6.1, we analyze the runtime of the algorithm implemented using all the ingredients exposed in Chapter 5. Minimizing the expectation value functional, the NLCG properly approximates the lowest eigenvalue of the AKLT Hamiltonian and an associated ground state. More precisely, given $\phi : D_{\mathrm{MPS}} \to \mathrm{MPS}^{\circ}(\mathbf{m}, \mathbf{n}, d)$, the parametrization of matrix product states with open boundary conditions, and H the AKLT Hamiltonian, the method solves Problem 5.0.2

$$\min_{A \in D_{\mathrm{MPS}}} f(A) = \min_{A \in D_{\mathrm{MPS}}} \frac{\phi(A)^{\dagger} H \phi(A)}{\phi(A)^{\dagger}\phi(A)}.$$

Meanwhile, we compare the method to a first small variation. The variation is a slight modification of the algorithm which consists in restarting the NLCG after less steps and it is based on dimensional considerations. The variation will lead to a small gain in terms of time of convergence, for a small number of sites.
In Section 6.2 we combine the restart reduction with a variation of the *line search*, which is the most expensive routine invoked by the NLCG algorithm. The modification of the line search is based on the decomposition of the domain described in Chapter 5, Section 5.6. In particular, we reparametrize the gradient $\nabla f : D_{\mathrm{MPS}} \to D_{\mathrm{MPS}}$, computed in Section 5.5, reducing the number of its coordinates from $\dim D_{\mathrm{MPS}}$ to the minimal number of parameters needed to parametrize the variety, i.e. the dimension of matrix product states with open boundary conditions and the "expected dimension" of homogeneous matrix product state with open boundary conditions.

6.1 Restart in less iterations

In this section, we show the performances of the algorithm compared to a slightly modified version. By Theorem 3.3.1 in Chapter 3, the matrix product states parametrization (5.6) is overparametrized since its generic fiber is a high dimensional subspace of the domain. Moreover, by Remark 5.2.3, in the case of matrix product states *with open boundary conditions*, the generic fiber of the parametrization (5.7) is isomorphic to the gauge orbit and the dimension of the variety, that we denote by $s_{\text{MPS}} = \dim \text{MPS}(\mathbf{m}, \mathbf{n}, d)$, is given by Equation (5.8). Therefore, we have a strict inequality $s_{\text{MPS}} < \dim D_{\text{MPS}}$. The first small variation of the algorithm that we compare to the standard one consists in *restarting* the iterations of the NLCG after s_{MPS} steps.

6.1.1 Results on $\text{MPS}(2, 3, d)$

Fix $m_i = m = 2$ and $n_i = n = 3$ for every $i = 1, \ldots, d$. Under these assumptions, the matrix product state map (5.7) is given by

$$\begin{aligned} \phi : \mathbb{C}^{3\times 2} \times \bigtimes_{i=2}^{d-1} \mathbb{C}^{3\times 2\times 2} \times \mathbb{C}^{2\times 3} &\to (\mathbb{C}^3)^{\otimes 3} \\ (A_1, \ldots, A_d) &\mapsto \sum_{i_1,\ldots,i_d=1}^{3} A_1^{i_1} \cdots A_d^{i_d}\ e_{i_1} \otimes \cdots \otimes e_{i_d}, \end{aligned} \tag{6.1}$$

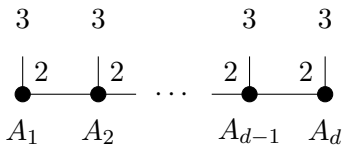


Figure 6.1: Matrix product state with open boundary conditions associated to the path graph with d vertices. The first and last tensors $A_1 \in \mathbb{C}^{3\times 2}$ and $A_d \in \mathbb{C}^{2\times 3}$ are matrices; the inner tensors are $A_k \in \mathbb{C}^{3\times 2\times 2}$, for $k = 2, \ldots, d-1$ and can be seen as a collection of 3 matrices of order 2, i.e. $A_k^{i_k} \in \mathbb{C}^{2\times 2}$ is a matrix for $i_k = 1, 2, 3$.

Denote the domain of the map (6.1) by

$$D_{\text{MPS}} := \mathbb{C}^{3\times 2} \times \left(\bigtimes_{k=2}^{d-1} \mathbb{C}^{3\times 2\times 2} \right) \times \mathbb{C}^{2\times 3} \simeq \mathbb{C}^{12(d-1)},$$

and its dimension by $N_{\text{MPS}} := \dim D_{\text{MPS}} = 12(d-1)$. Let $A = (A_1, \ldots, A_d) \in D_{\text{MPS}}$ be a point of the domain. Then, tensors A_1 and A_d have order 2 and they can be thought of as collections of 3 vectors of $\mathbb{C}^2$, and the inner tensors $A_2, \ldots, A_{d-1}$ have order 3 and each one can be thought of as a collection of 3 matrices of $\mathbb{C}^{2\times 2}$. In this case, the dimension of $\text{MPS}(2, 3, d)$, given in Example 5.2.4 and plotted in Figure 6.2, is

$$s_{\text{MPS}} := \dim \text{MPS}(2, 3, d) = 12(d-1) - 4(d-1) = 8(d-1).$$

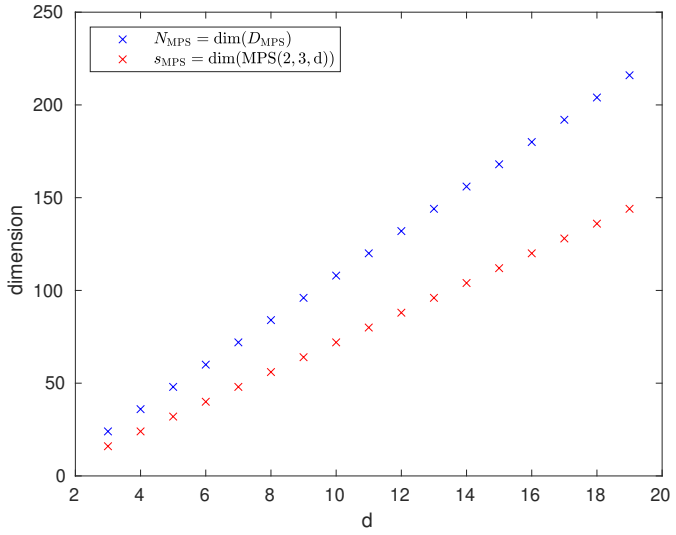


Figure 6.2: Comparison, for $d = 3, \dots, 19$, between the dimension of the domain of the MPS map $N_{\mathrm{MPS}} = 12(d-1)$ and the dimension of the variety $s_{\mathrm{MPS}} = \dim \mathrm{MPS}(2,3,d) = 8(d-1)$.

We consider Algorithm 5 on $\mathrm{MPS}(2,3,d)$, with different restarts: N_{MPS} for the standard NLCG, and s_{MPS} for the variation. For $d = 3, \dots, 19$ number of sites, we run 50 times the algorithms; each run starts from the same random point of D_{MPS} for both variants of Algorithm 5. The gradient's tolerance is set to $t = 10^{-8}$ and, for every d and in both variants, it allows to accomplish a precision of 10^{-15} with respect to the exact minimum $\lambda_{0,d} = -\frac{2}{3}(d-1)$ (5.17) of the functional for the Hamiltonian of the AKLT model, see Figure 6.3.

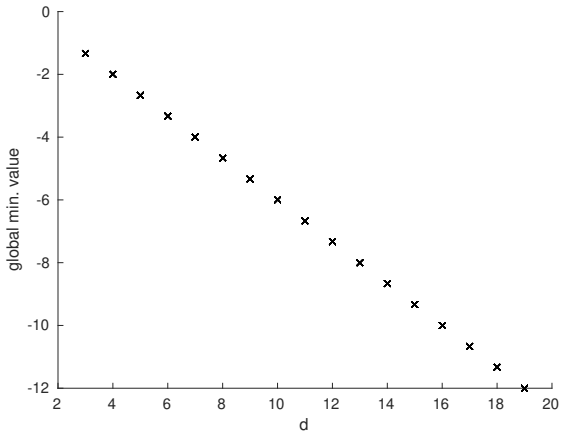


Figure 6.3: Value of global minimum computed by Algorithm 5 with restart $N_{\mathrm{MPS}} = \dim D_{\mathrm{MPS}}$ (standard NLCG), and restart $s_{\mathrm{MPS}} = \dim \mathrm{MPS}(2,3,d)$ (variation), for $d = 3, \dots, 19$. Both variants accomplish a precision of 10^{-15} with respect to the exact minimum $\lambda_{0,d} = -\frac{2}{3}(d-1)$.

The two variants of Algorithm 5 have a comparable time of the single line search on average, but they perform a different number of line searches in order to converge, c.f. Figure 6.4. The variation of the algorithm performs less line searches, i.e. less iterations of the NLCG, in the range $[3, \dots, 12]$. This provides a slightly faster time of convergence, around the 8%, which can be seen in Figure 6.5.

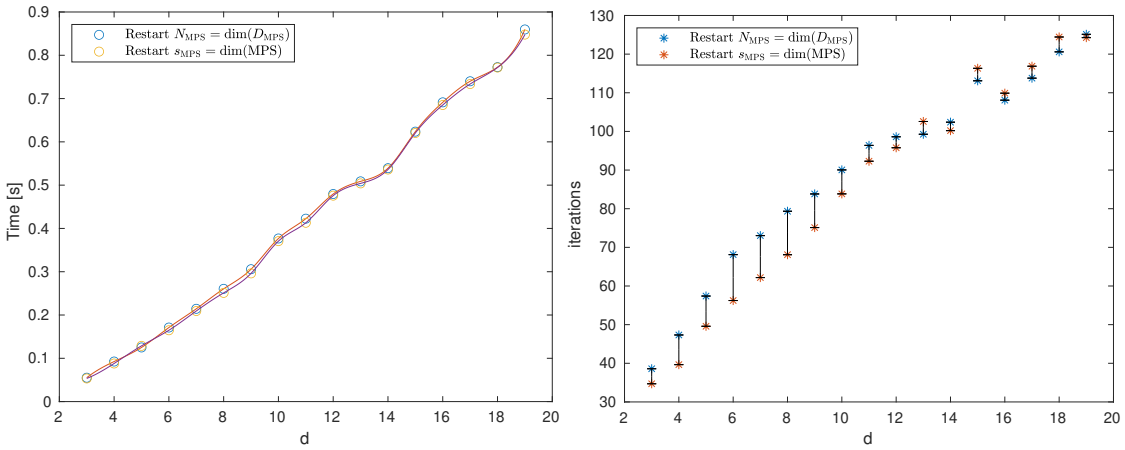


Figure 6.4: Comparison, for $d = 1, \dots, 19$, between Algorithm 5 with restart after $N_{\mathrm{MPS}} = \dim D_{\mathrm{MPS}}$ iterations (standard NLCG) and $s_{\mathrm{MPS}} = \dim \mathrm{MPS}$ iterations (variation), respectively. Left: comparison of the time of one line search; the time is comparable. Right: Comparison of the number of line searches needed by the two versions of the algorithm to reach convergence: the variation with restart s_{MPS} needs less iterations in the range $[3, \dots, 12]$ of sites. This leads to a moderate gain in time of convergence.

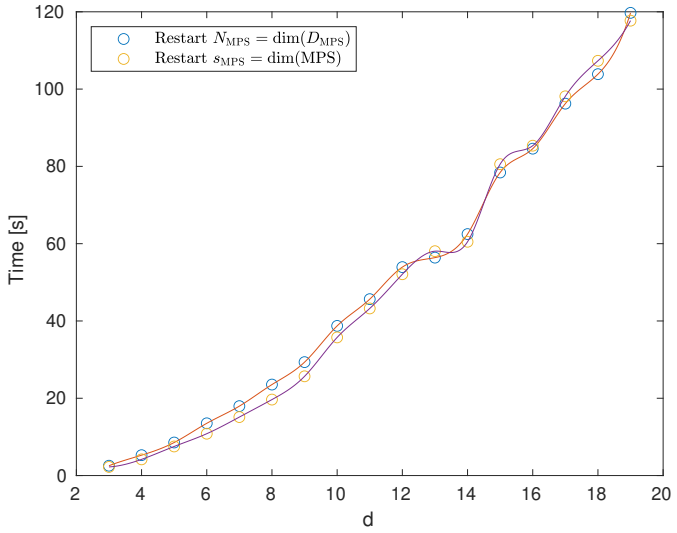


Figure 6.5: Time comparison, for $d = 1, \dots, 19$, between Algorithm 5 with restart after $N_{\mathrm{MPS}} = \dim D_{\mathrm{MPS}}$ iterations (standard NLCG) and $s_{\mathrm{MPS}} = \dim \mathrm{MPS}$ iterations (variation), respectively.

Plots 6.4 and 6.5 refer to *convergent* runs for both algorithms. Indeed, both variants of the algorithm have a small number of runs which do not reach convergence. The presence of non convergent runs is due to an iteration bound that we decide to impose to $3s_{\text{MPS}}$, within which the algorithms essentially converge. The role of the bound will be made more clear in the next section, precisely in Remark 6.1.1. However, the inclusion of non convergent runs in the MPS$(2, 3, d)$ model analysis is not much informative: in the case of non convergent runs, the runtime of Algorithm 5 is essentially the same for both variants, since, at the same runtime of line search, both variants reach the iteration bound. The comparison of the number of non convergent runs is shown in Figure 6.6 for the sake of completeness.

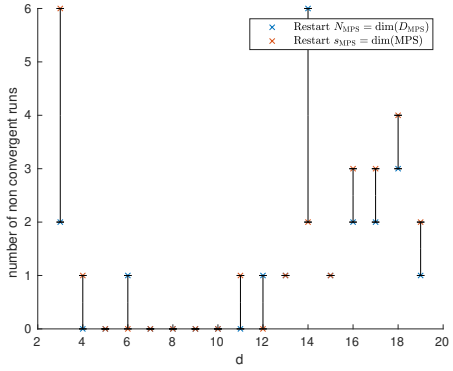


Figure 6.6: Comparison, for $d = 1, \ldots, 19$, of the number of non convergent runs (over 50 runs) due to the imposed iteration bound of $3s_{\text{MPS}}$ iterations.

6.1.2 Results on hMPS$(2, 3, d)$

We restrict our attention to homogeneous matrix product states with open boundary conditions hMPS$(2, 3, d)$, c.f. Definition 5.2.7. When $m = 2$ and $n = 3$ the homogeneous matrix product state map (5.9) is the following

$$\begin{aligned} \psi : \mathbb{C}^2 \times (\mathbb{C}^{2\times 2})^{\times 3} \times \mathbb{C}^2 &\to (\mathbb{C}^3)^{\otimes 3} \\ p = (v_L, A, v_R) &\mapsto \sum_{i_1,\ldots,i_d=1}^{3} \Big(v_L^\dagger A_{i_1} \cdots A_{i_d} v_R \Big) e_{i_1} \otimes \cdots \otimes e_{i_d}. \end{aligned} \tag{6.2}$$

The domain of the hMPS map (6.2) is $D_{\text{hMPS}} := \mathbb{C}^{16}$, of dimension $N_{\text{hMPS}} := 16$. An element $(v_L, A, v_R) \in D_{\text{hMPS}}$ is such that $v_L, v_R \in \mathbb{C}^2$ and $A = (A_1, A_2, A_3) \in (\mathbb{C}^{2\times 2})^{\times 3}$ is a tensor of order 3, identified with a set of 3 matrices of order 2.

A bound on the dimension of the variety is given in Equation (5.10) and in this case, it is $\dim \text{hMPS}(2, 3, d) \leq 11$. We consider the upper bound on the dimension of the variety as its "expected dimension" and we define

$$s_{\text{hMPS}} := \operatorname{expdim} \text{hMPS}(2, 3, d) = 11.$$

Figure 6.7: Graphical representation of a homogeneous matrix product state with open boundary conditions, with bond dimensions equal to 2 and local dimensions equal to 3. Each inner vertex is associated with the same tensor $A \in (\mathbb{C}^{2\times 2})^{\times 3}$, with $A_i \in \mathbb{C}^2$ for every $i = 1, 2, 3$. The boundaries are vectors $v_L, v_R \in \mathbb{C}^2$.

We consider Algorithm 5 with different restarts: N_{MPS} for the standard NLCG, and s_{MPS} for the variation. The two variants of Algorithm 5 have comparable performances again. More precisely, for $d = 3, \ldots, 40$ number of sites, we run 100 times the algorithms; each run starts from the same random point of D_{hMPS} for both variants. The gradient's tolerance is set to $t = 10^{-6}$. In Figure 6.8 we display the functional values reached. Unlike the case of $\mathrm{MPS}(2, 3, d)$, minimizing on the subvariety $\mathrm{hMPS}(2, 3, d)$, both the NLCG and the variant converge to two different kinds of points, for every d: one point that approximates a ground state and that accomplishes a precision of 10^{-11} with respect to the exact minimum of the functional $\lambda_{0,d} = -\frac{2}{3}(d-1)$ (5.17), and another point that satisfies the tolerance condition that we have fixed to $t = 10^{-6}$. These second kind of points approximately lie on a line of slope $-1/3$ that can be approximately described by the expression $-(q + (1/3)(d-1))$, with $q \sim 0.05$.

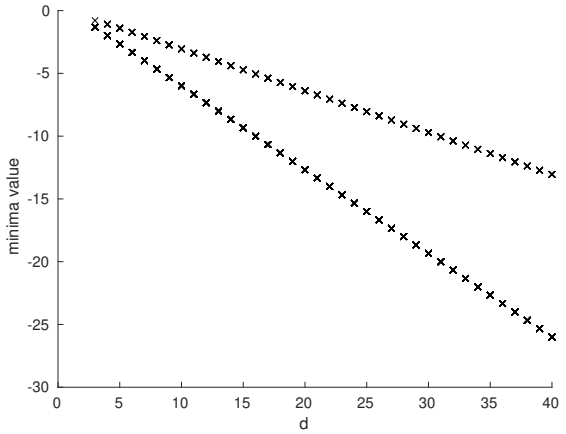


Figure 6.8: Values of global minima (lower data) and another point (upper data), for $d = 3, \ldots, 40$, computed by Algorithm 5 with restart $N_{\mathrm{hMPS}} = \dim D_{\mathrm{hMPS}}$ (standard NLCG), and restart $s_{\mathrm{hMPS}} = \mathrm{expdim}\, \mathrm{hMPS}(2, 3, d)$ (variation). The global minimum, in both variants, has a precision of (the order of) 10^{-11} with respect to the exact minimum $\lambda_{0,d} = -\frac{2}{3}(d-1)$. The other point lies on the line $-(q + (1/3)(d-1))$, with $q \sim 0.05$.

Non global minimum. The gradient of $f : D_{\text{hMPS}} \to \mathbb{R}$, computed in Section 5.5, Equation (5.25), is given by

$$G_A = \frac{2}{\phi(A)^\dagger \phi(A)} \left(\frac{\partial}{\partial \overline{A}} (\phi(\overline{A})^t) \, (H - h \cdot \mathbb{1}) \, \phi(A) \right), \quad \text{where } h = \rho(\phi(A)).$$

When the gradient is zero, i.e. $G_A = 0$, for $A \in D_{\text{hMPS}}$ then either $\phi(A)$ is an eigenvector of H with associated eigenvalue $h = \rho(\phi(A))$, or $d\phi_{\overline{A}} = \frac{\partial}{\partial \overline{A}}(\phi(\overline{A})^t) = \mathbb{0}$, or $(H - h \cdot \mathbb{1})\phi(A) \in \ker(d\phi_A)$. With a direct computation in the range of sites $[3, \ldots, 14]$ (computationally possible), we verified that the non-global "stationary" points reached do not satisfy the eigenvalue equation, i.e. $|H\phi(A) - h\phi(A)|_F > 10^3$, with $|\cdot|_F$ the Frobenius norm. The convergent points which are not points of global minimum are not even points of local minimum. We therefore *assume* that the algorithm finds one global minimum for each d. However, within the tolerance threshold imposed, the algorithms converge to these points. In the following analysis, we take them into account since, a priori, we can only see that the algorithms are converging to them: we will consider their runtime, which is longer than that of the global minima, and their contribution to the total time of the algorithms. We will refer to them as "false local minima".

Remark 6.1.1 (Iteration bound)**.** We decide to bound the number of iterations by $3s_{\text{hMPS}} = 33$, i.e. 3 complete rounds of NLCG with restart after $s_{\text{hMPS}} = 11$ iterations. We think that the bound is a natural threshold to impose: it ensures convergence to the global minimum on average, for every d and both algorithms. Indeed, in the range of sites, the global minimum is reached, on average, in 18 iterations by the standard Algorithm 5, and in 16 iterations with the lower restart s_{hMPS}, that is the variation needs less iterations to satisfy the tolerance condition. The number of global and non-global minima reached (in 100 runs) is compared in Figure 6.9 (left) and (right), respectively.

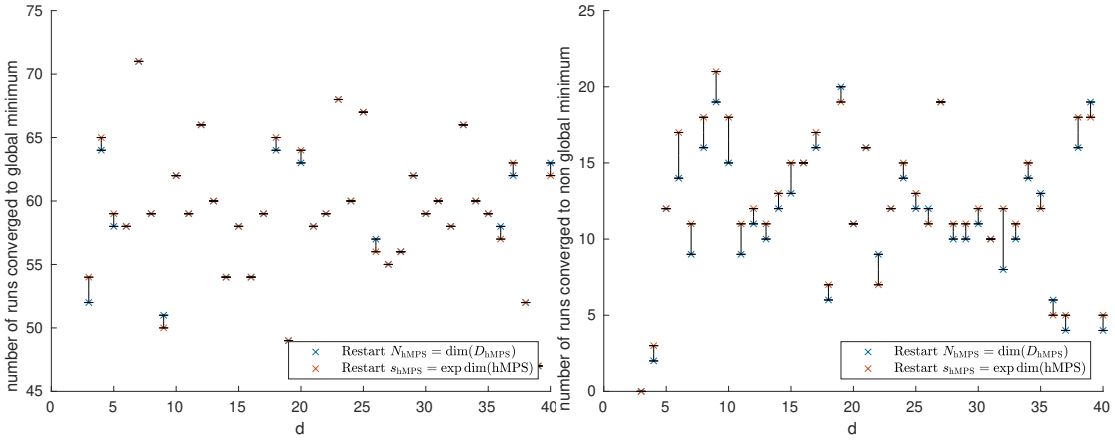


Figure 6.9: Left: Comparison, for $d = 3, \ldots, 40$, of the number of runs converged to the global minimum (over 100 runs). Right: Comparison, for $d = 3, \ldots, 40$, of the number of runs converged to the false local minimum (over 100 runs). Both refer to convergent runs reached by Algorithm 5 with restart $N_{\text{hMPS}} = \dim D_{\text{hMPS}}$ (standard NLCG), and restart $s_{\text{hMPS}} = \text{expdim hMPS}(2, 3, d)$ (variation).

The false local minimum needs more iterations for convergence (26 and 24 on average) and therefore it has a slower runtime to convergence compared to the global minimum. The iteration bound causes a loss of convergent runs in both variants. The 100 runs divide into the number of runs converged to the global minimum, 59%, the number of runs converged to the false local minimum, 12%, and the number of runs that do not reach convergence. The latter runs are compared in Figure 6.10 (around 29% of runs).

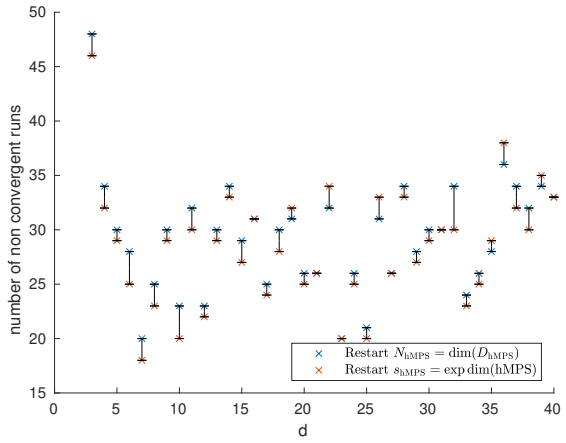


Figure 6.10: Comparison, for $d = 3, \ldots, 40$, of the number of non convergent runs (over 100 runs) within the iteration bound described in Remark 6.1.1.

In general, the two algorithms have a comparable time of line search, c.f. Figure 6.11.

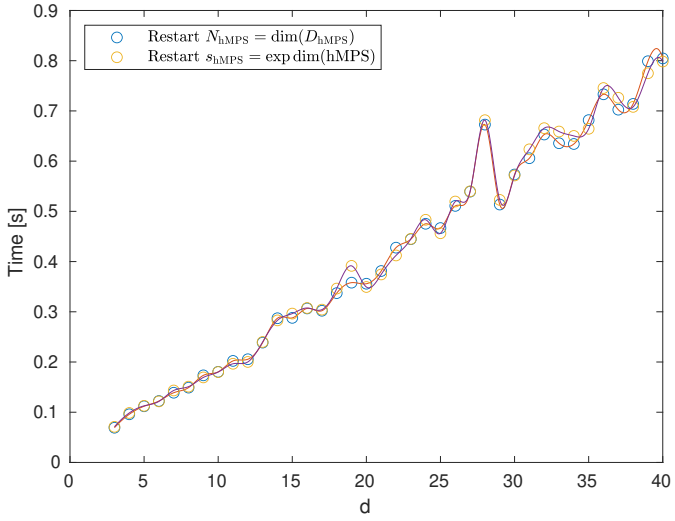


Figure 6.11: Comparison, for $d = 3, \ldots, 40$, of the time of one line search, on average, between Algorithm 5 with restart after $N_{\mathrm{hMPS}} = \dim D_{\mathrm{hMPS}}$ iterations (standard NLCG) and $s_{\mathrm{hMPS}} = \exp\dim \mathrm{hMPS}$ iterations (variation). The time is comparable.

On the other hand, Algorithm 5 with restart after s_{hMPS} iterations performs less line searches to reach convergence. Since the runtime to convergence is roughly the time of

the line search multiplied by the number of line searches (iterations), the latter determines the runtime comparison, retaining the comparable time of line search.

Considering the overall performance, i.e. the time of *all runs* (convergent and non convergent), Algorithm 5 implemented with restart s_{hMPS} performs, on average, around 23 against 22 line searches over the whole interval of sites, c.f. Figure 6.12 (left). In terms of time, this leads to a gain of 7% of runtime that can be seen in Figure 6.12 (right).

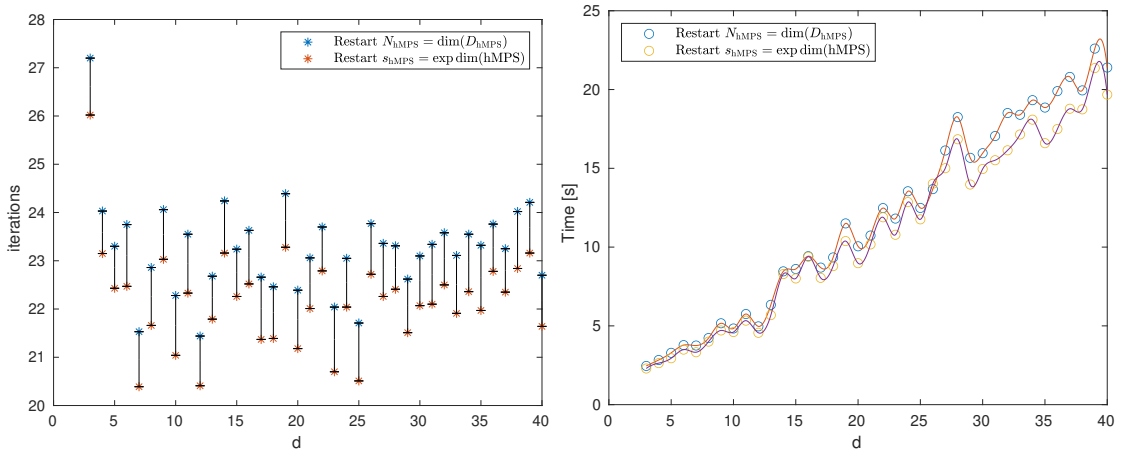


Figure 6.12: Left: Comparison, for $d = 3, \ldots, 40$, of the number of line searches between Algorithm 5 with restart $N_{\mathrm{hMPS}} = \dim D_{\mathrm{hMPS}}$ (standard NLCG) and $s_{\mathrm{hMPS}} = \exp\dim \mathrm{hMPS}$ (variation); the variation needs less iterations leading to a gain in total runtime. Right: Overall time comparison.

Comparing the *convergent runs* for both variants of the algorithm, the iterations needed to converge are, on average, 19 against 18 and the variation gains around 8% of runtime, c.f. Figure 6.13.

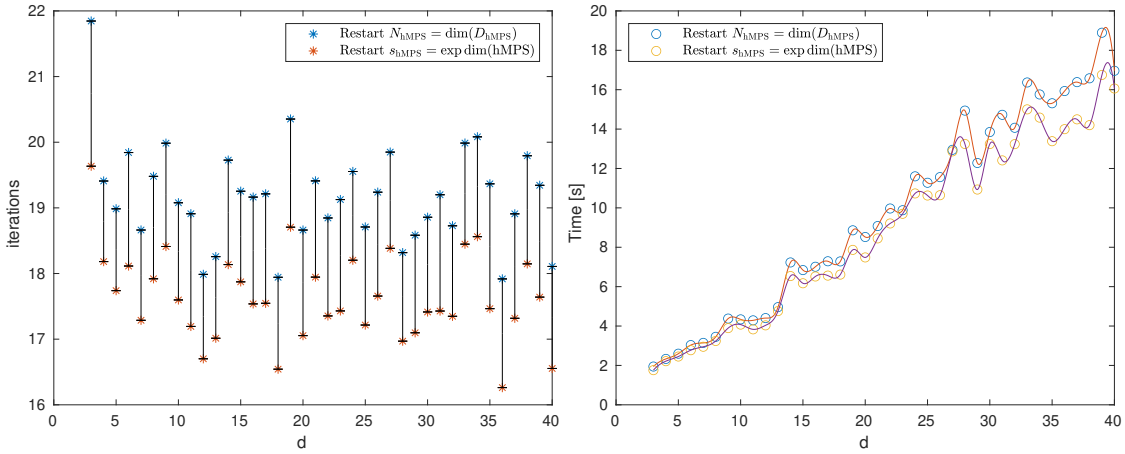


Figure 6.13: Comparison, for $d = 3, \ldots, 40$, of the number of line searches needed to reach convergence: Algorithm 5 with restart $s_{\mathrm{hMPS}} = \exp\dim \mathrm{hMPS}$ needs less iterations, leading to a gain in time of convergence. Right: Runtime comparison to convergent runs.

Comparing the *convergent runs to the global minimum* for both the variants of the algorithm, the iterations needed to converge are, on average, 18 against 16 and the variation gains around 8% of runtime, c.f. Figure 6.14.

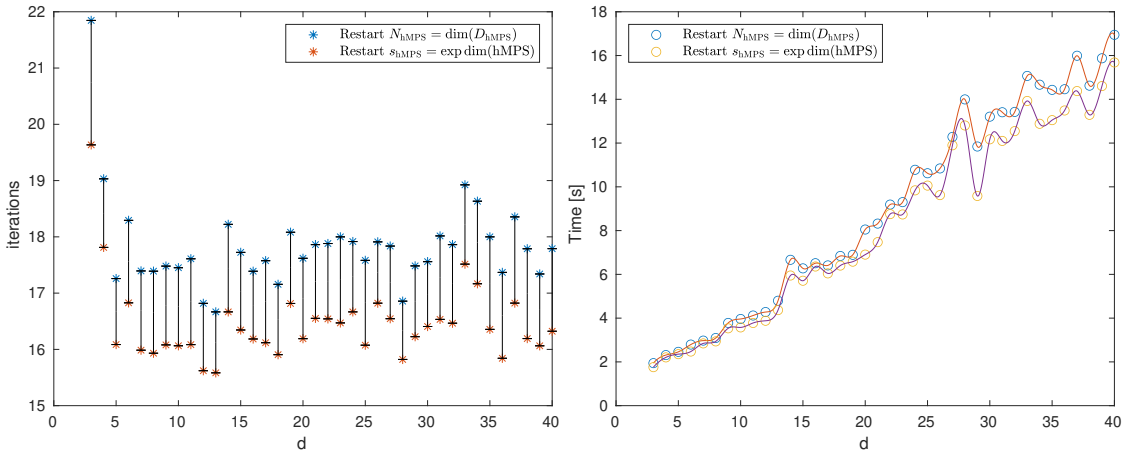


Figure 6.14: Left: Comparison, for $d = 3, \ldots, 40$, of the number of line searches needed by the two variants of Algorithm 5 to reach the global minimum: the variation needs less iterations leading to a gain in time of convergence. Right: Runtime comparison to the global minimum.

In conclusion, the NLCG properly finds an approximation of the ground state of the AKLT Hamiltonian: a gradient's tolerance $t = 10^{-8}$ allows for a precision of 10^{-15} on $\mathrm{MPS}(2, 3, d)$, and a tolerance $t = 10^{-6}$ allows for a precision of 10^{-11} on $\mathrm{MPS}(2, 3, d)$. Moreover, the first small variation of Algorithm 5 with restart after less iterations, corresponding to the dimension of the respective variety, either $\mathrm{MPS}(2, 3, d)$ or $\mathrm{hMPS}(2, 3, d)$, gains around 8% of runtime to convergence in the respective interval of sites. The speed-up is due to the fact that less iterations of the line search are needed for convergence, retaining a comparable time of each line search. We want to highlight that this is generally the reason why one algorithm performs better than the other: the principle bottleneck of the NLCG runtime is the number and the expensiveness of line searches performed. On the variety $\mathrm{MPS}(2, 3, d)$, both variants of Algorithm 5 find a global minimum and they fail few times to converge within the fixed bound on the number of iterations. On $\mathrm{hMPS}(2, 3, d)$, instead, they both find one value of global minimum but they fail several times to converge to it, within the bound. Indeed, both variants can also converge to a non global minimum (fixed the tolerance to $t = 10^{-6}$), which is assumed to be a "false local minimum". We have a probability of 59% of finding the global minimum in the runtime given in Figure 6.14, and 12% the false local minimum. In general, the runtime to convergence is clearly higher in the case of $\mathrm{MPS}(2, 3, d)$, since the dimension of the variety increases in the system size, contrary to the homogeneous case $\mathrm{hMPS}(2, 3, d)$ in which the dimension is independent of the number of sites.

6.2 Variation of the algorithm

We finally present our work for a matrix product state algorithm to approximate ground states of systems with open boundary conditions. Our algorithm is a variation of the nonlinear conjugate gradient method (c.f. Section 5.1), based on computations of Section 5.7 in Chapter 5.

Let $\phi : D_{\text{MPS}} \to \text{MPS}(\mathbf{m}, \mathbf{n}, d)$ be the parametrization of the matrix product states with open boundary conditions, given in Definition 5.2.2. Let $Q \in \text{MPS}(\mathbf{m}, \mathbf{n}, d)$ be a smooth point of the variety and let $A \in \phi^{-1}(Q)$ be a point in the fiber. In the case of matrix product states with open boundary conditions, the fiber $\phi^{-1}(Q)$ of the map coincides exactly with the orbit of $A \in D_{\text{MPS}}$ under the action of the gauge group $\mathcal{O}_{G(A)}$. Equivalently, the kernel of the differential of the map $\ker(d\phi_A)$ coincides with the tangent space to the gauge orbit $T_A\mathcal{O}_{G(A)}$. Denote $\mathcal{N}_A = \ker(d\phi_A) = T_A\mathcal{O}_{G(A)}$ and the complementary vector space by $\mathcal{B}_A = \mathcal{N}_A^{\perp}$. The domain of the map admits a natural pointwise decomposition given by the tangent space to the gauge orbit and its complementary vector space:

$$D_{\text{MPS}} \simeq T_A D_{\text{MPS}} = \mathcal{N}_A \oplus \mathcal{B}_A, \tag{6.3}$$

where, in particular, $s_{\text{MPS}} = \dim \mathcal{B}_A = \dim \text{MPS}(\mathbf{m}, \mathbf{n}, d) < \dim D_{\text{MPS}}$. Every vector $w \in T_Q \text{MPS}$ therefore admits a representative $v \in \mathcal{B}_A \subset T_A D_{\text{MPS}}$ such that $w = d\phi_A(v)$.

Consider now the functional to minimize $f = \rho \circ \phi : D_{\text{MPS}} \to \mathbb{R}$, given by

$$f(A) = \rho \circ \phi(A) = \frac{\phi(A)^\dagger H \phi(A)}{\phi(A)^\dagger \phi(A)}. \tag{6.4}$$

The gradient of the functional $G_A = \nabla f_A \in T_A D_{\text{MPS}} \simeq D_{\text{MPS}}$ has been computed in Chapter 5, Section 5.5. In this section, we show that the gradient $\nabla f_A \in D_{\text{MPS}}$ is a vector that naturally lives in the subspace $\mathcal{B}_A \subset T_A D_{\text{MPS}}$. Therefore, if we find a basis of $T_A D_{\text{MPS}}$, for every $A \in D_{\text{MPS}}$, given by the union of a basis of $\mathcal{B}_A$ and a basis of $\mathcal{N}_A$, the gradient ∇f_A, in this basis, has null components in $\mathcal{N}_A$ and can be represented in $s_{\text{MPS}} = \dim \mathcal{B}_A$ coordinates.

Assuming to have an orthonormal basis satisfying this condition, we can represent the gradient in this new basis. Indeed, the basis requested can be derived from computations done in Section 5.7 of Chapter 5. In Subsections 5.7.2 and 5.7.3 we exactly computed the matrices $\mathbb{T}_A$, whose columns span the space $\mathcal{N}_A = T_A\mathcal{O}_{G(A)}$ (not necessarily orthonormal), for (homogeneous) matrix product states with open boundary conditions. Then, in the case of matrix product states with open boundary conditions, for every $A \in D_{\text{MPS}}$, the orthogonal completion of the columns of $\mathbb{T}_A$ in D_{MPS} gives a set of vectors which span the space $\mathcal{B}_A$. These vectors are used to build a matrix, denoted by P_A, c.f. Remark 5.7.6. More precisely, in our algorithm the matrix P_A is found numerically with MATLAB [Mat20]: the matrix $\mathbb{T}_A$ is orthonormalized with a QR decomposition, $\mathbb{T}_A = Q_A R_A$, and then $P_A = \ker(Q_A^\dagger)$ is computed. For homogeneous matrix product

states with open boundary conditions we obtain instead a matrix whose columns span $\mathcal{B}_p$, for every $p \in D_{\text{hMPS}}$, which has dimension equal to the "expected dimension" of the variety. Based on this change of coordinates of the gradient, we propose a variation to the nonlinear conjugate gradient method.

6.2.1 Change of coordinates of the gradient

Let $\phi : D_{\text{MPS}} \to \text{MPS}(\mathbf{m}, \mathbf{n}, d)$ be the parametrization of the matrix product states with open boundary conditions, c.f. Definition 5.2.2. Let $L : D_{\text{MPS}} \to \mathbb{C}^{N_{\text{MPS}}}$ be the bijection given in Remark 5.2.5, which is a fixed reshape of tensors into vectors. Let $G_A = \nabla f_A \in T_A D_{\text{MPS}} \simeq D_{\text{MPS}}$ be the gradient of the functional f (6.4).

Lemma 6.2.1. *Let $Q \in \text{MPS}(\mathbf{m}, \mathbf{n}, d)$ and $A \in \phi^{-1}(Q) \subseteq D_{\text{MPS}}$. The gradient $G_A \in D_{\text{MPS}}$, is orthogonal to the tangent directions to the fiber $\phi^{-1}(Q)$, and therefore to the tangent directions of the gauge orbit.*

Proof. If $v \in T_A \phi^{-1}(Q) \subset T_A D_{\text{MPS}}$ then $v \in \ker(d\phi_A)$, and

$$L(G_A)^\dagger v = \nabla(\rho \circ \phi)_A^\dagger \, v = d(\rho \circ \phi)_A(v) = d\rho_z|_{z=\phi(A)} d\phi_A(v) = 0. \qquad \square$$

Corollary 6.2.2. *Let $A \in D_{\text{MPS}}$. The gradient $L(G_A) \in D_{\text{MPS}}$ is a vector of the orthogonal complement of the kernel of the differential in $A \in D_{\text{MPS}}$, i.e. $L(G_A) \in \mathcal{B}_A = \mathcal{N}_A^\perp$ in decomposition (6.3).*

Remark 6.2.3. The image of the matrix product state map is gauge invariant, c.f. Theorem 3.3.1. "Fixing" the gauge degrees of freedom refers to fixing a representative of the fiber of the map. It is a well-known technique in physics and, in the context of matrix product states, it usually consists in putting the tensors of the network in the so called *canonical form* which consists in imposing the conditions

$$\sum_{i=1}^{n_k} A_k^{i_k} {A_k^{i_k}}^\dagger = \text{Id}_{m_k},$$

for every $k = 1, \ldots, d$. We refer to [PGVWC07] for the definition and properties of this representation. After fixing a representative of the fiber of the map, gauge degrees of freedom are left in the representation of vectors of the tangent space to the variety [HMOV14], because, on the other side, the kernel of the differential of the map is non empty because it contains the tangent directions to the gauge orbits. Our method has precisely the goal to "fix" the gauge degrees of freedom of the tangent space of the variety, pointwise choosing a representative of the tangent space. This technique is used also in the TDVP [HCO+11] and the VUMPS algorithms [ZSVF+18]. We want to highlight that a motivation which made our first tentative interesting, c.f. in Section 5.6, is the fact that it fixes the gauge degrees of freedom of the fiber and of the tangent space, simultaneously. This is due to the fact that the section of the domain of dimension equal to the codimension of the variety provides a new domain of the parametrization.

The new domain has dimension equal to the dimension of the variety and it can be identified with the tangent space of the variety, since the kernel of the differential of the reparametrization is trivial for a dimensional count.

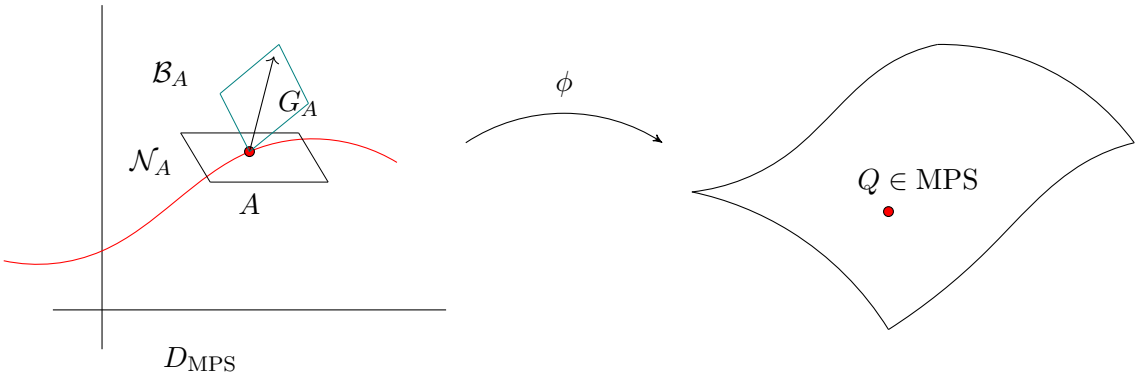


Figure 6.15: Graphical representation of the matrix product state map. The fiber $\phi^{-1}(Q)$ is represented as a red line in the domain of the map D_{MPS}. The vector space $T_A D_{\text{MPS}}$ decomposes into $\mathcal{N}_A \oplus \mathcal{B}_A$, with $\mathcal{N}_A = \ker(\phi_A)$ and $\mathcal{B}_A = \mathcal{N}_A^{\perp}$. In the case of open boundary conditions $\mathcal{N}_A = T_A \mathcal{O}_{G(A)}$ and $\dim \mathcal{B}_A = s_{\text{MPS}}$.

Assume that $s_{\text{MPS}} = \dim \text{MPS}(\mathbf{m}, \mathbf{n}, d)$ and that $\{w_1(A), \dots, w_{s_{\text{MPS}}}(A)\}$, $w_j(A) \in D_{\text{MPS}}$, is an orthonormal basis of $\mathcal{B}_A$. Denote

$$P_A = (w_1(A) \dots w_s(A)) \in \mathbb{C}^{\dim(D_{\text{MPS}}) \times s_{\text{MPS}}}. \tag{6.5}$$

The matrix $P_A = (w(A)^i_j)^{i=1,\dots,N}_{j=1,\dots,s_{\text{MPS}}}$ is reshaped into a tensor $B \in D_{\text{MPS}} \otimes \mathbb{C}^{s_{\text{MPS}}*}$ in the following way

$$B_{k\,\gamma_k \delta_k}^{\;\;i_k} = B_{l^{-1}(i)} = \sum_{j=1}^{s_{\text{MPS}}} w_j^i(A)\; e_j^* \in \mathbb{C}^{s_{\text{MPS}}*},$$

where $\{e_j^*\}_{j=1}^{s_{\text{MPS}}}$ is the canonical basis of $\mathbb{C}^{s_{\text{MPS}}*}$ and $(k i_k \gamma_k \delta_k) = l^{-1}(i)$, where l^{-1} is the inverse of the index bijection given in Remark 5.2.5. That is, every column of P_A, which is a vector in $\mathbb{C}^{D_{\text{MPS}}}$ is reshaped via L^{-1}.

Analogously, the dual operator $P_A^\dagger$ is reshaped into $\overline{B} \in D_{\text{MPS}}^* \otimes \mathbb{C}^{s_{\text{MPS}}}$

$$\overline{B}_{k\,\gamma_k \delta_k}^{\;\;i_k} = \overline{B}_{l^{-1}(i)} = \sum_{j=1}^{s_{\text{MPS}}} w_j^i(\overline{A})\; e_j \in \mathbb{C}^{s_{\text{MPS}}},$$

where $\{e_j\}_{j=1}^{s_{\text{MPS}}}$ is the canonical basis of $\mathbb{C}^{s_{\text{MPS}}}$.

Since, by Corollary 6.2.2, the gradient $L(G_A) \in \mathcal{B}_A$ then $g_A = P_A^\dagger L(G_A) \in \mathbb{C}^{s_{\text{MPS}}}$ is the gradient written in coordinates with respect to the basis $\{w_1(A) \dots w_s(A)\}$ of $\mathcal{B}_A$.

Equivalently, g_A is the vector

$$g_A = \overline{B}(G_A) = \sum_{k=1}^{d} \overline{B}_k (G_A)_k, \tag{6.6}$$

which is represented in Figure 6.16.

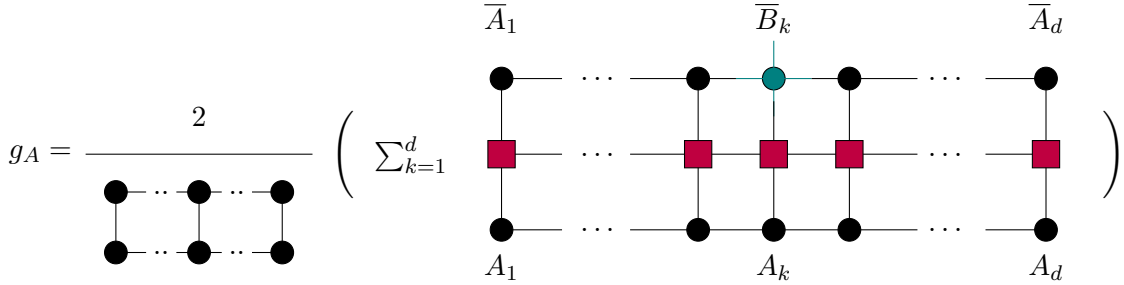


Figure 6.16: The pictorial representation of the projection of the gradient on the orthogonal complement of the tangent space to the fiber of the parametrization.

Remark 6.2.4. Notice that $P_A^\dagger P_A = \mathrm{Id}_{\mathcal{B}_A}$ and $P_A P_A^\dagger$ is the projector on $\mathcal{B}_A$. If $v \in \mathcal{B}_A \subset D_{\mathrm{MPS}}$ then $PP^\dagger v = v$. In particular, $L(G_A) = PP^\dagger L(G_A) = Pg_A$, equivalently

$$G_A = B(g_A) = \sum_{k=1}^{d} B_k (g_A)_k.$$

Moreover, notice that

$$L(G_A)^\dagger L(G_A) = (P_A g_A)^\dagger (P_A g_A) = g_A^\dagger P_A^\dagger P_A g_A = g_A^\dagger g_A,$$

therefore $|L(G_A)| = |g_A|$. This implies that every scalar product between gradients is computed between vectors written in the basis $\{w_1(A), \dots, w_s(A)\}$ of $\mathcal{B}_A$, i.e. vectors of $\mathbb{C}^{s_{\mathrm{MPS}}}$.

Notation 6.2.5. Recall that, with abuse of notation, we always consider the vector reshape of tensors, i.e. we denote $A = L(A)$, $G_A = L(G_A)$. In this notation, we have the following relations

$$g_A = P_A^\dagger G_A,$$
$$G_A = P_A g_A.$$

6.2.2 Variation of the line search

We describe the variation of the line search. In particular, we will highlight the difference between the two versions in the performance of the first and second steps of the NLCG. The line search method, Algorithm 3, solves

$$\alpha_j = \arg\min_{\alpha \in \mathbb{R}} f(A_{(j)} + \alpha p_j)$$

requiring several *evaluations of the functional* and several *computations of gradients*. In what follows, we compute these two quantities, highlighting the role of matrices $P_\bullet^\dagger$ and $P_\bullet$, that are the changes of basis of $\mathcal{B}_\bullet$ inside $T_\bullet D_{\text{MPS}}$.

First line search. Let $A \in D_{\text{MPS}}$ be the starting point. Compute the gradient $G_A \in D_{\text{MPS}}$ as prescribed in Equation (5.25). Let $p_A = -G_A = -P_A P_A^\dagger G_A$ be the first descent direction. The first line search solves

$$\alpha_0 = \arg\min_{\alpha \in \mathbb{R}} f(A - \alpha G_A).$$

Let $A' = A - \alpha G_A$, $\alpha \in \mathbb{R}$ be a point on the line $A - \alpha G_A$, and compute P_A. In the line search we have several computations of

1. Functional value:
$$f(A') = f(A - \alpha G_A) = f(A - \alpha P_A P_A^\dagger G_A) = f(A - \alpha P_A g_A). \tag{6.7}$$
2. Gradient:
$$G_{A'} = G_{A - \alpha G_A} = P_{A - \alpha G_A} g_{A - \alpha G_A} = P_{A'} g_{A'}. \tag{6.8}$$

First step comparison. The variation of the line search consists in the following procedure: for every iteration of the algorithm, we *fix* inside the line search routine the vector space $\mathcal{B}_A$ and therefore the matrix $P_A \in \mathbb{C}^{N_{\text{MPS}} \times s_{\text{MPS}}}$. The consequence is the definition of an *approximate gradient*, that will be given in Equation (6.10). This modification of the line search is based on two observations, whose graphical representation is given in Figure 6.17:

1. The gradient is a vector of $\mathcal{B}_A$ for every point $A \in D_{\text{MPS}}$, therefore the point $A' = A + \alpha p$, with $p \in D_{\text{MPS}}$ search direction is in a different gauge orbit (and different fiber of the parametrization), with respect to A.
2. In a neighborhood of the point $A \in D_{\text{MPS}}$ the vector space $\mathcal{B}_\bullet$ remains approximately parallel to $\mathcal{B}_A$. More precisely, if $A' = A + \alpha p \in D_{\text{MPS}}$ for α admissible step length, then the columns of $P_{A'}$ and P_A generates approximately the same vector space.

Definition 6.2.6. Given $A, A' \in D_{\text{MPS}}$ and $P_A, P_{A'} \in \mathbb{C}^{D_{\text{MPS}} \times s_{\text{MPS}}}$ as in equation (6.5); define

$$g_{A'}^A := P_A^\dagger G_{A'}, \qquad G_{A'}^A := P_A g_{A'}.$$

Notice that, for every $A \in D_{\text{MPS}}$, the gradients trivially satisfy: $g_A = P_A^\dagger G_A = g_A^A$ and $G_A = P_A g_A = G_A^A$.

Let $A \in D_{\text{MPS}}$ be the same starting point. Compute P_A and $g_A = P_A^\dagger G_A \in \mathbb{C}^{s_{\text{MPS}}}$ as in Equation (6.6). The first descent direction is $p_A = -g_A$ and the first line search solves

$$\alpha_0 = \arg\min_{\alpha \in \mathbb{R}} f(A - \alpha P_A g_A).$$

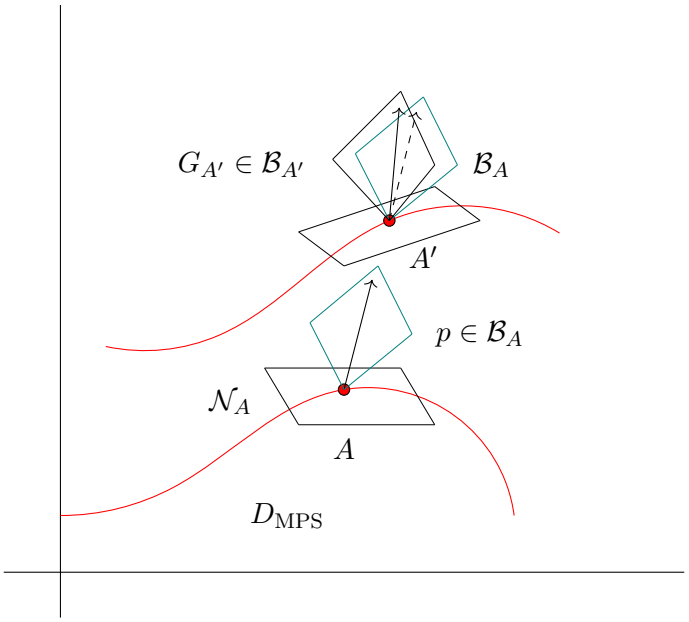


Figure 6.17: Graphical representation of the domain D_{MPS}. At each point $A \in D_{\text{MPS}}$ the domain admits the decomposition $D_{\text{MPS}} = \mathcal{N}_A \oplus \mathcal{B}_A$. The step from $A \in D_{\text{MPS}}$, along the direction $p \in \mathcal{B}_A$ brings to another point $A' \in D_{\text{MPS}}$ which is contained in another fiber of the map. The vector space $\mathcal{B}_{A'}$ is replaced by the previous $\mathcal{B}_A$.

Denote $A' = A - \alpha P_A g_A$, $\alpha \in \mathbb{R}$. In the line search, we impose the following rules for evaluating the functional values and computing of gradients:

1. Functional value:

$$f(A') = f(A - \alpha P_A g_A) (= f(A - \alpha G_A)). \tag{6.9}$$

2. Approximate gradient:

$$g_{A'}^{A} = P_A^{\dagger} G_{A'} \neq P_{A'}^{\dagger} G_{A'} = g_{A'}. \tag{6.10}$$

Notice that the computation of the functional value (6.9) is equivalent to (6.7) but the approximate gradient (6.10) is different from gradient (6.8).

Remark 6.2.7. Notation $g_{A'}^{A}$ indicates that the approximate gradient is obtained by the gradient $G_{A'}$ calculated in the point A', but multiplied by the matrix P_A. The columns of P_A give a basis of $\mathcal{B}_A$, that is a vector space containing G_A, the starting gradient computed in A.

Algorithm 6: Variation of NLCG on MPS

Input: Initial point $A_{(0)} \in D_{\text{MPS}}$, tolerance $t \in \mathbb{R}$
Output: $A^* \in D_{\text{MPS}}$ such that $\phi(A^*) \in$ MPS approximate ground state
Compute $s_{\text{MPS}} = \dim(\text{MPS}) < N = \dim D_{\text{MPS}}$;
Evaluate $f_{A_{(0)}} = f(A_{(0)})$;
Compute $P_{A_{(0)}}$ and $g_{A_{(0)}}$;
$p_0 \leftarrow -g_{A_{(0)}}$;
$j \leftarrow 0$;
while $|g_{A_{(j)}}| > t$ **do**
    $\boxed{\alpha_j = \arg\min_\alpha f(A_{(j)} + \alpha P_{A_{(j)}} p_j)}$;
    Set $A_{(j+1)} = A_{(j)} + \alpha_j P_{A_{(j)}}(p_j)$;
    Compute $P_{A_{(j+1)}}$ and $g_{A_{(j+1)}}$;
    if $\boxed{j = s_{\text{MPS}}}$ **then**
        $\beta_j^{\text{FR}} \leftarrow 0$;
        $j \leftarrow 0$;
    else
        $\beta_{j+1}^{\text{FR}} \leftarrow \frac{g_{A_{(j+1)}}^\dagger g_{A_{(j+1)}}}{g_{A_{(j)}}^\dagger g_{A_{(j)}}}$;
    end
    $p_{j+1} \leftarrow -g_{A_{(j+1)}} + \beta_{k+1}^{\text{FR}}\ P_{A_{(j)}}(p_j)$;
    $j \leftarrow j + 1$;
end

Second step comparison. Let $A_{(1)} \in D_{\text{MPS}}$ be the starting point and $A_{(0)}$ the previous point computed by both the NLCG and the variation. Denote the current search directions by $p = -G_{A_{(1)}} + \beta G_{A_{(0)}} \in D_{\text{MPS}}$ and $\widetilde{p} = -g_{A_{(1)}} + \beta g_{A_{(0)}} \in \mathbb{C}^{s_{\text{MPS}}}$, of the NLCG and the variation respectively.

Then for the NLCG we have

1. Functional value: $f(A')$, with

$$A' = A_{(1)} + \alpha p = A_{(1)} + \alpha(-G_{A_{(1)}} + \beta G_{A_{(0)}}) = A_{(1)} - \alpha G_{A_{(1)}} + \alpha\beta G_{A_0}.$$

2. Gradient:

$$G_{A'} = P_{A'}^\dagger g_{A'}.$$

For the variation of the algorithm we have:

1. Functional value: $f(A'') \neq f(A')$, since

$$\begin{aligned} A'' = A_{(1)} + \alpha P_{A_{(1)}} \widetilde{p} &= A_{(1)} - \alpha P_{A_{(1)}} g_{A_{(1)}} + \alpha\beta P_{A_1} g_{A_0} \\ &= A_{(1)} - \alpha G_{A_{(1)}} + \alpha\beta G_{A_{(0)}}^{A_{(1)}} \neq A'. \end{aligned}$$

2. Gradient:
$$g_{A'}^{A_{(1)}} = P_{A_{(1)}}^{\dagger} G_{A'} = P_{A_{(1)}}^{\dagger} P_{A'} g_{A'} \neq P_{A'}^{\dagger} P_{A'} g_{A'} = g_{A'}.$$

In general, from the second iteration on, we obtain two non equivalent algorithms. We give the pseudo-code of the NLCG variation in Algorithm 6.

We compare again the performances of Algorithms 5 and 6 on the varieties $\mathrm{MPS}(2,3,d)$ and $\mathrm{hMPS}(2,3,d)$, with open boundary conditions. We show that the algorithms are comparable and Algorithm 6 has a gain in terms of the time of convergence.

6.2.3 Results on $\mathrm{MPS}(2,3,d)$

Consider the variety of matrix product states with open boundary conditions, c.f. Definition 5.2.2. In particular, the variety $\mathrm{MPS}(2,3,d)$ is parametrized by the map

$$\begin{aligned} \phi : &\mathbb{C}^{3\times 2} \times \bigtimes_{i=2}^{d-1} \mathbb{C}^{3\times 2\times 2} \times \mathbb{C}^{2\times 3} \to (\mathbb{C}^3)^{\otimes 3} \\ &(A_1, A_2, \ldots, A_{d-1}, A_d) \mapsto \sum_{i_1,\ldots,i_d=1}^{3} A_1^{i_1} \cdots A_d^{i_d}\ e_{i_1} \otimes \cdots \otimes e_{i_d}, \end{aligned}$$

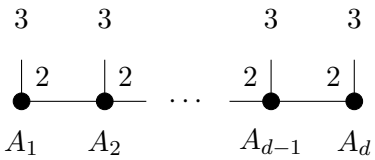


Figure 6.18: Example of a matrix product state with open boundary conditions associated to the path graph with d vertices, bond dimensions 2 and local dimensions 3. The first and last tensors $A_1 \in \mathbb{C}^{3\times 2}$ and $A_d \in \mathbb{C}^{2\times 3}$ are matrices; the inner tensors are $A_k \in \mathbb{C}^{3\times 2\times 2}$, for $k = 2, \ldots, d-1$ and $A_k^{i_k} \in \mathbb{C}^{2\times 2}$ is a matrix for every $i_k = 1, 2, 3$.

The domain of the parametrization is $D_{\mathrm{MPS}} := \mathbb{C}^{12(d-1)}$, the gauge group, $\bigtimes_{i=1}^{d-1} GL_2$, has dimension $4(d-1)$ and the variety has dimension $s_{\mathrm{MPS}} := 8(d-1)$.

In Section 5.7, Subsection 5.7.2, we compute the tangent space $T_A\mathcal{O}_{G(A)}$ in $A \in D_{\mathrm{MPS}}$ to the gauge orbit of $A \in D_{\mathrm{MPS}}$. The building blocks for writing the matrix $\mathbb{T}_A \in$

$\mathbb{C}^{12(d-1)\times 4(d-1)}$, whose columns generate $T_A\mathcal{O}_{G(A)}$, are:

$$\begin{aligned} m_L^s(1) &= {a_1}^s \boxtimes \mathrm{Id}_2, \\ m^s(i) &= \big(-\mathrm{Id}_2 \boxtimes (A_i^s)^t \ (A_i^s) \boxtimes \mathrm{Id}_2 \big), \quad i = 1, \dots, d, \\ m_R^s(d) &= -\mathrm{Id}_2 \boxtimes {a_d}^s(2), \end{aligned}$$

where the boundary tensors are $A_1 = (a_{1_j}^s) \in \mathbb{C}^{3\times 2}$ and $A_d = (a_{d_j}^s) \in \mathbb{C}^{3\times 2}$, and the inner tensors are $A_i = (a_{i_{jl}}^s) \in \mathbb{C}^{3\times 2\times 2}$, for $i = 2, \dots, d-1$.

Given the starting point, $A \in D_{\mathrm{MPS}}$, of a line search, the matrix $\mathbb{T}_A \in \mathbb{C}^{12(d-1)\times 4(d-1)}$ is orthonormalized with a QR decomposition, $\mathbb{T}_A = Q_A R_A$ and then $P_A = Q_A^\dagger \in \mathbb{C}^{12(d-1)\times 8(d-1)}$ is computed (Equation (6.5)). Both computations are done numerically in MATLAB [Mat20] and are available at https://github.com/claudia-dela/NLCG_MPS_open-boundaries/. In our variation of the NLCG, for every line search, the matrix P_A is computed at the starting point of the line search $A \in D_{\mathrm{MPS}}$, and it is *fixed* inside that line search routine.

We compare again convergent runs for both algorithms. For $d = 3, \dots, 19$ number of sites, we run 50 times the algorithms, each run starts from the same random point of D_{MPS}. The gradient's tolerance is set again to $t = 10^{-8}$ and it allows to accomplish a precision of 10^{-15} with respect to the exact minimum of the functional $\lambda_{0,d} = -\frac{2}{3}(d-1)$ (5.17), for every d also for Algorithm 6, c.f. Figure 6.19.

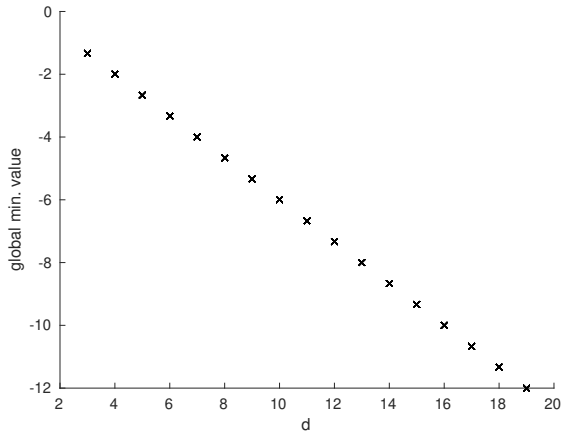


Figure 6.19: Value of global minimum computed Algorithms 5 and 6, for $d = 3, \dots, 19$. Both algorithms accomplish a precision of 10^{-15} with respect to the exact minimum $\lambda_{0,d} = -\frac{2}{3}(d-1)$.

We give a naive formula that models the time of the algorithm. For every $d = 3, \dots, 19$, denote the time of the algorithm by $T_{\mathrm{alg}}(d)$. The time of one line search is denoted by $S_{\mathrm{ls(alg)}}(d)$, that is the time to compute:

$$\alpha_j = \arg\min_{\alpha\in\mathbb{R}} f(A_{(j)} + \alpha P_{A_{(j)}} p_j).$$

The computational time *outside* the line search in one iteration of NLCG, is denoted by $O_{\mathrm{ls(alg)}}(d)$. For Algorithm 5, it is the cost of computing the following part:

Set $A_{(j+1)} = A_{(j)} + \alpha_j p_j \quad (\alpha_j = \arg\min_\alpha f(A_{(j)} + \alpha p_j))$;
Compute $\boxed{G_{A_{(j+1)}}}$;
if $j = N_{\mathrm{MPS}}$ **then**
  $\beta_j^{\mathrm{FR}} \leftarrow 0$;
  $j \leftarrow 0$;
else
  $\beta_{j+1}^{\mathrm{FR}} \leftarrow \frac{G^\dagger_{A_{(j+1)}} G_{A_{(j+1)}}}{G^\dagger_{A_{(j)}} G_{A_{(j)}}}$;
end
$p_{j+1} \leftarrow -G_{A_{(j+1)}} + \beta_{k+1}^{\mathrm{FR}}\, p_j$;

---

and for Algorithm 6, it is the cost of computing the following part:

Set $A_{(j+1)} = A_{(j)} + \alpha_j P_{A_{(j)}}(p_j) \quad (\alpha_j = \arg\min_\alpha f(A_{(j)} + \alpha P_{A_{(j)}} p_j))$;
Compute $\boxed{P_{A_{(j+1)}} \text{ and } g_{A_{(j+1)}}}$;
if $j = s_{\mathrm{MPS}}$ **then**
  $\beta_j^{\mathrm{FR}} \leftarrow 0$;
  $j \leftarrow 0$;
else
  $\beta_{j+1}^{\mathrm{FR}} \leftarrow \frac{g^\dagger_{A_{(j+1)}} g_{A_{(j+1)}}}{g^\dagger_{A_{(j)}} g_{A_{(j)}}}$;
end
$p_{j+1} \leftarrow -g_{A_{(j+1)}} + \beta_{k+1}^{\mathrm{FR}}\, P_{A_{(j)}}(p_j)$;

Denote the number of line searches by $N_{\mathrm{ls(alg)}}(d)$. The overall time of line search per algorithm is therefore $T_{\mathrm{ls(alg)}}(d) = S_{\mathrm{ls(alg)}}(d) N_{\mathrm{ls(alg)}}(d)$ and, analogously, the overall time outside the line search is $\Delta_{\mathrm{alg}}(d) = O_{\mathrm{ls(alg)}}(d) N_{\mathrm{ls(alg)}}(d)$. Therefore we have:

$$T_{\mathrm{alg}}(d) = T_{\mathrm{ls(alg)}}(d) + \Delta_{\mathrm{alg}}(d) \simeq \left(S_{\mathrm{ls(alg)}}(d) + O_{\mathrm{ls(alg)}}(d)\right) N_{\mathrm{ls(alg)}}(d).$$

In the plots that follow we decide to simulate data for $d = 20, \ldots, 25$. We simply perform a linear interpolation in MATLAB [Mat20] of the data for $d = 3, \ldots, 19$. The vertical line in Plots 6.20 and 6.21 divides the collected data from the simulated ones.

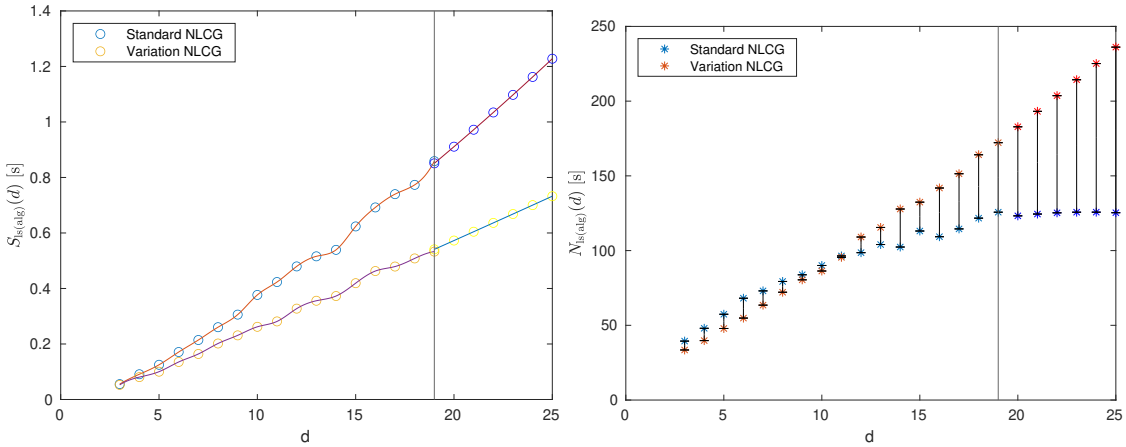


Figure 6.20: Left: Comparison of time of the single line search $S_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 25$: Algorithm 6 is faster in the line search. Right: Comparison of the number of line searches $N_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 25$: for $d > 11$, Algorithm 5 needs less iterations to reach convergence.

We can see in Figure 6.20 (left) that the line search of Algorithm 6 is faster than the line search of Algorithm 5, because computations depend on the number of sites d and therefore involve $s_{\mathrm{MPS}} = 8(d-1)$ coordinates instead of $N_{\mathrm{MPS}} = 12(d-1)$. On the other hand, the number of line searches of Algorithm 6 switches from lower to higher after $d = 12$, Figure 6.20 (right). Despite the latter observation, the overall time of the line search routine is lower for our Algorithm 6, in the range $[3, \ldots, 19]$, as can be seen in Figure 6.21.

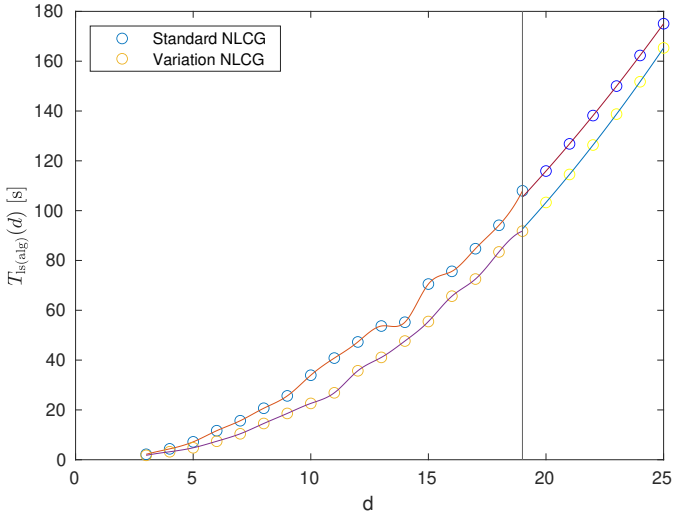


Figure 6.21: Comparison of the overall time of line search $T_{\mathrm{ls(alg)}}(d) = S_{\mathrm{ls(alg)}}(d) N_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 25$.

The runtime costs that do not depend on the line search routine are plotted in Figure 6.22 (left).

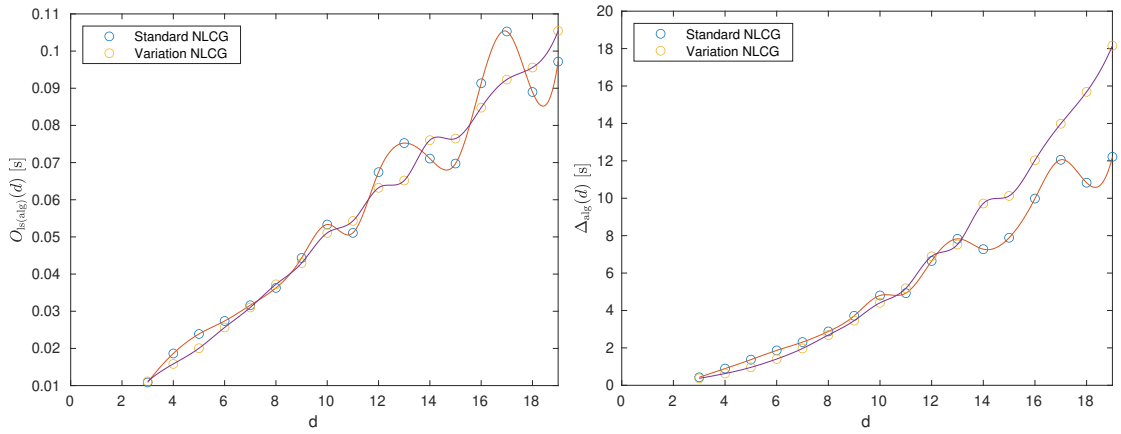


Figure 6.22: Left: Comparison of the time cost outside the line search per iteration $O_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 19$. Right: Comparison of the overall time outside the line search routine $\Delta_{\mathrm{alg}}(d) = O_{\mathrm{ls(alg)}}(d) N_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 19$.

Notice that, retaining the comparable runtime cost outside the line search per iteration, compared in Figure 6.22 (left), the higher number of line searches of Algorithm 6, c.f. Figure 6.20 (right) affects the total cost outside the line search, Figure 6.22 (right).We expect that increasing the number of sites, the rise of the number of line searches in Algorithm 6 (Figure 6.20, right), which weighs on the computational cost both inside and outside the line search routine, can make it lose the gain. In the simulation, after $d = 25$ sites, the algorithms become comparable in terms of time, c.f. Figure 6.23. However, despite the previous observation, in the interval under consideration $[3, \ldots, 19]$ we gain on average the 22% in terms of the time of convergence, see Figure 6.23.

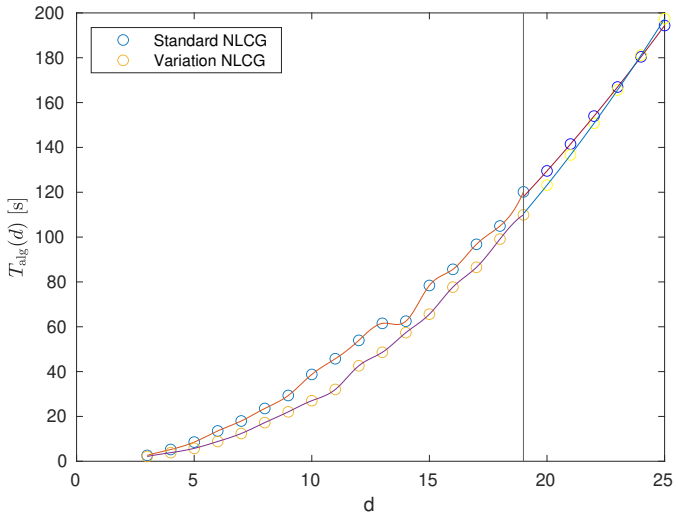


Figure 6.23: Overall time comparison between Algorithm 5 and Algorithm 6 on MPS$(2, 3, d)$, for $d = 3, \ldots, 25$.

Recall that we impose a bound of $3s_{\text{MPS}}$ on the number of iterations and that the two algorithms are compared within this bound. Also in Algorithm 6 the 50 runs divide into the number of runs converged to the global minimum and the number of runs that do not reach convergence within the bound. The latter are compared in Figure 6.6. We can notice that Algorithm 6 is faster to reach the iteration bound because, retaining the same number of iterations, its line search is faster. Moreover, we can notice that our Algorithm 6 (which has a faster line search) has few non convergent runs over the 50 runs, except for the cases $d = 3$ and $d = 4, 17$, in which the algorithm loses respectively 5 and 1 convergent runs. Algorithm 5 loses runs over the whole range of sites. However we do not consider the runtime of non convergent runs for the model $\text{MPS}(2, 3, d)$ as in the previous comparison: it is negligible due to the relatively low number of such runs.

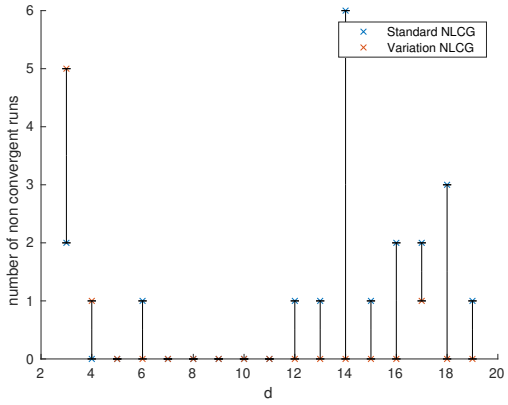


Figure 6.24: Comparison, for $d = 3, \ldots, 19$, of the number of non convergent runs (over 50 runs) within the iteration bound $3s_{\text{MPS}}$.

6.2.4 Results on $\text{hMPS}(2, 3, d)$

We restrict our attention to the homogeneous case with open boundary conditions $\text{hMPS}(2, 3, d)$, c.f. Definition 5.2.7, defined by the map

$$\psi : \mathbb{C}^2 \times (\mathbb{C}^{2\times 2})^{\times 3} \times \mathbb{C}^2 \to (\mathbb{C}^3)^{\otimes 3}$$
$$p = (v_L, A, v_R) \mapsto \sum_{i_1,\ldots,i_d=1}^{3} \left(v_L^\dagger A_{i_1} \cdots A_{i_d} v_R\right) e_{i_1} \otimes \cdots \otimes e_{i_d},$$

The domain of the parametrization is $D_{\text{hMPS}} := \mathbb{C}^{16}$, of dimension $N_{\text{hMPS}} := 16$ and the gauge group is GL_2 of dimension 4.

Denote again the upper bound on the dimension of the variety by

$$s_{\text{hMPS}} := \text{expdim}\, \text{hMPS}(2, 3, d) = 11.$$

Figure 6.25: Example of a homogeneous matrix product states with open boundary conditions with bond dimensions equal to 2 and local dimensions equal to 3. Each inner vertex is associated with the same tensor $A \in (\mathbb{C}^{2\times 2})^{\times 3}$, with $A_i \in \mathbb{C}^2$ for every $i = 1, 2, 3$. The boundaries are vectors $v_L, v_R \in \mathbb{C}^2$.

In Section 5.7, Subsection 5.7.3, we compute the tangent space to the gauge orbit $T_p\mathcal{O}_{G(p)}$ and the 1-dimensional vector space S_p, for $p = (v_L, A, v_R) \in D_{\mathrm{hMPS}}$, such that

$$\mathcal{B}_p = (S_p \oplus T_p\mathcal{O}_{G(p)})^{\perp},$$

with $\dim \mathcal{B}_p = \operatorname{expdim} \mathrm{hMPS}(m, n, d)$. Equation (5.32) gives the matrix whose columns generate $S_p \oplus T_p\mathcal{O}_{G(p)} \subset \ker(d\psi_p)$. For $m = 2$ and $n = 3$ we have:

$$\mathbb{T}_p = \begin{pmatrix}
-dv_L^1 & v_L^1 & v_L^1 & 0 & v_L^2 \\
-dv_L^2 & v_L^2 & -v_L^2 & v_L^1 & 0 \\
a_{11}^1 & 0 & 0 & -a_{21}^1 & a_{12}^1 \\
a_{12}^1 & 0 & -2a_{12}^1 & a_{11}^1 - a_{22}^1 & 0 \\
a_{21}^1 & 0 & 2a_{21}^1 & 0 & a_{22}^1 - a_{11}^1 \\
a_{22}^1 & 0 & 0 & a_{21}^1 & -a_{12}^1 \\
a_{11}^2 & 0 & 0 & -a_{21}^2 & a_{12}^2 \\
a_{12}^2 & 0 & -2a_{12}^2 & a_{11}^2 - a_{22}^2 & 0 \\
a_{21}^2 & 0 & 2a_{21}^2 & 0 & a_{22}^2 - a_{11}^2 \\
a_{22}^2 & 0 & 0 & a_{21}^2 & -a_{12}^2 \\
a_{11}^3 & 0 & 0 & -a_{21}^3 & a_{12}^3 \\
a_{12}^3 & 0 & -2a_{12}^3 & a_{11}^3 - a_{22}^3 & 0 \\
a_{21}^3 & 0 & 2a_{21}^3 & 0 & a_{22}^3 - a_{11}^3 \\
a_{22}^3 & 0 & 0 & a_{21}^3 & -a_{12}^3 \\
0 & -v_R^1 & -v_R^1 & -v_R^2 & 0 \\
0 & -v_R^2 & v_R^2 & 0 & -v_R^1
\end{pmatrix} \in \mathbb{C}^{N_{\mathrm{hMPS}} \times 5}.$$

The matrix P_p with $p \in D_{\mathrm{hMPS}}$ is constructed analogously to that of the previous model. We compute the QR decomposition $\mathbb{T}_p = Q_p R_p$, and then the kernel of the conjugate transpose of Q_p:

$$P_p = \ker(Q_p^{\dagger}), \quad p \in D_{\mathrm{hMPS}},$$

where $Q_p \in \mathbb{C}^{N_{\mathrm{hMPS}} \times s_{\mathrm{hMPS}}}$ is such that $\mathbb{T}_p = Q_p R_p$. We obtain a $N_{\mathrm{hMPS}} \times s_{\mathrm{hMPS}}$ matrix P_p for $p \in D_{\mathrm{hMPS}}$, whose columns span the s_{hMPS}-dimensional vector space $\mathcal{B}_p$.

We compare Algorithm 5 and Algorithm 6 implemented on $\mathrm{hMPS}(2, 3, d)$. In our variation, given $p \in D_{\mathrm{hMPS}}$ as a starting point of a line search, the matrix P_p is computed and it is *fixed* inside that line search routine. For $d = 3, \ldots, 40$, we run 100 times the

algorithms, each run starts from the same random point of D_{hMPS} for both algorithms. The gradient's tolerance is set again to $t = 10^{-6}$. Our Algorithm 6 also converges to two different kinds of points, for every d: one global minimum and one false local minimum, c.f. Figure 6.26. The point which approximates a ground state accomplishes a precision of 10^{-12} with respect to the exact minimum of the functional $\lambda_{0,d} = -\frac{2}{3}(d-1)$ (5.17), against the precision 10^{-11} of Algorithm 5. Again the false local minima lie on a line of slope $-1/3$ that can be approximately described by the expression $-(q + (1/3)(d-1))$, with $q \sim 0.05$.

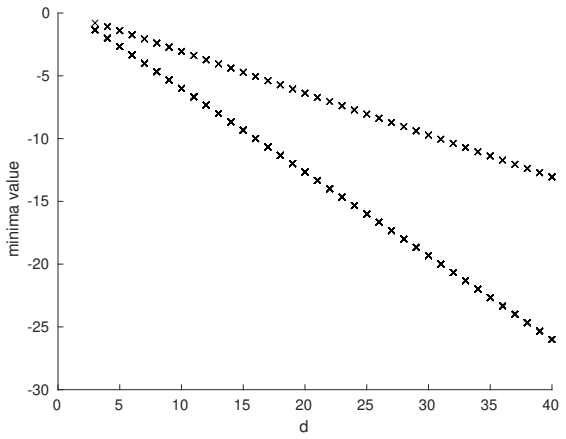


Figure 6.26: Values of global (lower data) and false local (upper data) minima, for $d = 3, \ldots, 40$, computed by both Algorithms 5 and 6. The variation of the algorithm allows for a precision of 10^{-12} with respect to the exact value of the global minimum $\lambda_{0,d} = -\frac{2}{3}(d-1)$, against the 10^{-11} of the standard algorithm (c.f. Subsection 6.1.2). The other point lies again on the line $-(q + (1/3)(d-1))$, with $q \sim 0.05$.

We recall that we set a bound on the number of iterations to $3s_{\text{hMPS}} = 33$, c.f. Remark 6.1.1 and that, unlike the general MPS case, the NLCG finds, for every d, the ground state and another point that satisfies the tolerance condition on the gradient, fixed to $t = 10^{-6}$. Algorithm 6 also converges to this point that we refer to as "false local minimum". Within the imposed bound $3s_{\text{hMPS}}$, the 100 runs divide into the number of runs converged to the global minimum (Figure 6.27 (left)), the false local minimum (Figure 6.27 (right)) and the number of runs that do not reach convergence (Figure 6.28). On average, Algorithm 6 needs 16 iterations in order to find the global minimum against the 18 needed by Algorithm 5. We have a probability of 59% of finding the global minimum, which coincides with that of Algorithm 5.

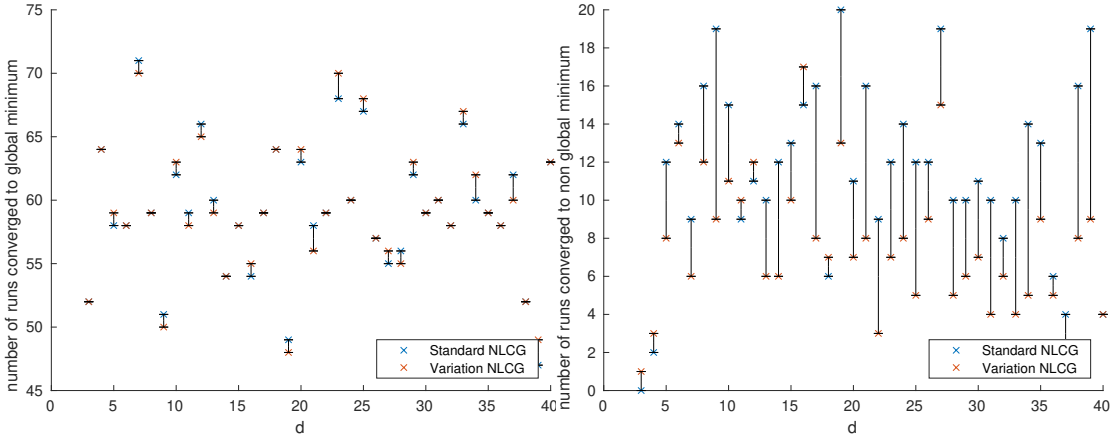


Figure 6.27: Left: Comparison, for $d = 3, \dots, 40$, of the number of runs converged to the global minimum (over 100 runs). Right: Comparison, for $d = 3, \dots, 40$, of the number of runs converged to the false local minimum (over 100 runs). Both plots refer to the comparison between Algorithms 5 and 6 on hMPS$(2, 3, d)$.

On the other hand, Algorithm 6 has less probability to reach the false local minimum, around 8%, compared to the standard algorithm (12%). This is because Algorithm 6 needs around 30 iterations of NLCG (against 26 of the standard algorithm) in order to converge to the false local minimum. Therefore, some runs which are converging to the false local minimum in Algorithm 5, do not converge in Algorithm 6. The number of non convergent runs (over 100 runs) is plotted in Figure 6.28 for the two algorithms (again around 33%).

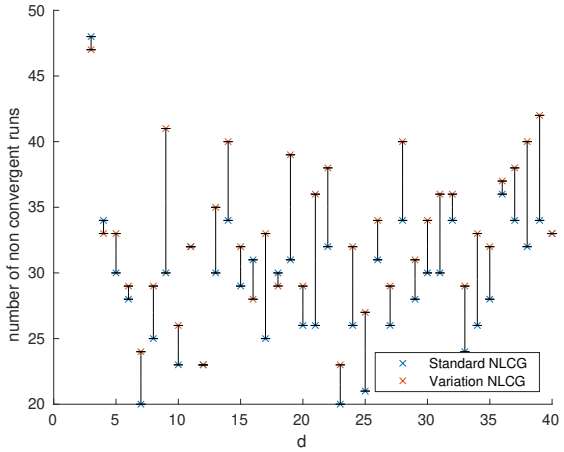


Figure 6.28: Comparison, for $d = 3, \dots, 40$, of the number of non convergent runs within the iteration bound (over 100 runs) given in Remark 6.1.1, between Algorithms 5 and 6 on hMPS$(2, 3, d)$

Despite the case of MPS$(2, 3, d)$, the runtime of the line search, Figure 6.29, is comparable since $N_{\text{hMPS}} - s_{\text{hMPS}} = 16 - 11 = 5$ is small. The difference between the two algorithms lies in the number of line searches, in particular in the case of convergence to *the global minimum*. Taking into account the overall runtime (convergent and non convergent runs), Algorithm 5 and Algorithm 6 perform an almost comparable number of iterations over the whole interval of sites, c.f. Figure 6.30 (left). For this reason, at a first sight, the two algorithms have the same performance in the interval $[3, \dots, 40]$, as shown in Figure 6.30 (right).

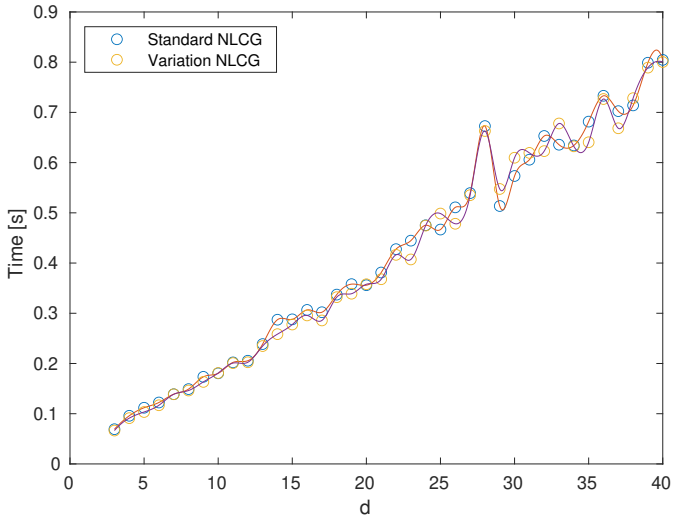


Figure 6.29: Comparison, for $d = 3, \dots, 40$, of the time of one line search, on average, between Algorithms 5 and 6 on hMPS$(2, 3, d)$. The time of the single line search is comparable over the whole interval of sites.

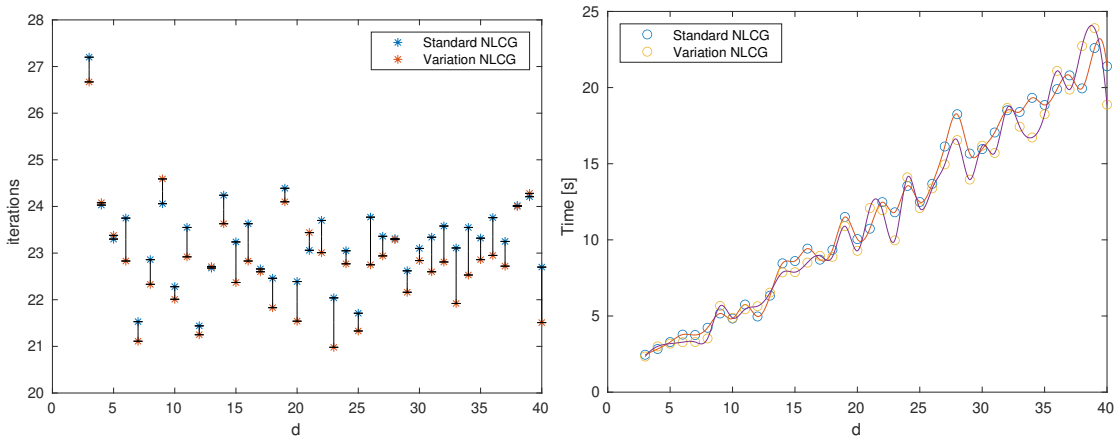


Figure 6.30: Left: Comparison of the number of line searches between Algorithms 5 and 6 on hMPS$(2, 3, d)$; the number is almost comparable. Right: Overall time comparison between Algorithms 5 and 6 on hMPS$(2, 3, d)$. The overall runtime is comparable.

The difference in runtime becomes more clear when we compare *convergent runs* for both the algorithms. Algorithm 6 needs, on average, 17 iterations against 18 (standard NLCG) in order to converge, c.f. Figure 6.31 (left), and the variation gains around 12% of runtime, c.f. Figure 6.31 (right).

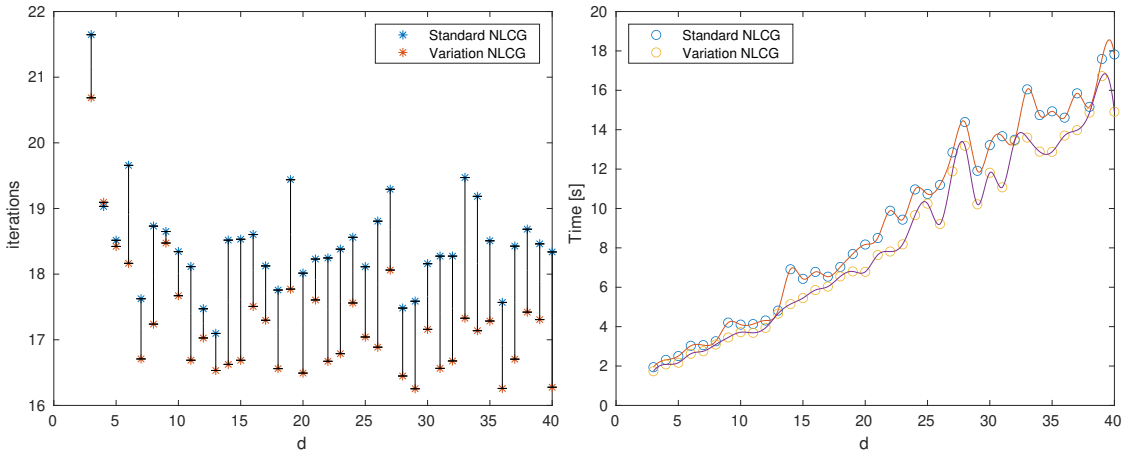


Figure 6.31: Left: Comparison of the number of line searches needed by the algorithms to reach convergence: the variation needs less iterations leading to a gain of 12% in time of convergence. Right: Runtime comparison to convergent runs.

Finally, we focus on the *convergent runs to the global minimum* for both algorithms. The iterations needed to converge in Algorithm 6 are on average 16 against 18 of the standard NLCG, c.f. Figure 6.32 (left), and therefore Algorithm 6 gains again around 12% of runtime, c.f. Figure 6.32 (right).

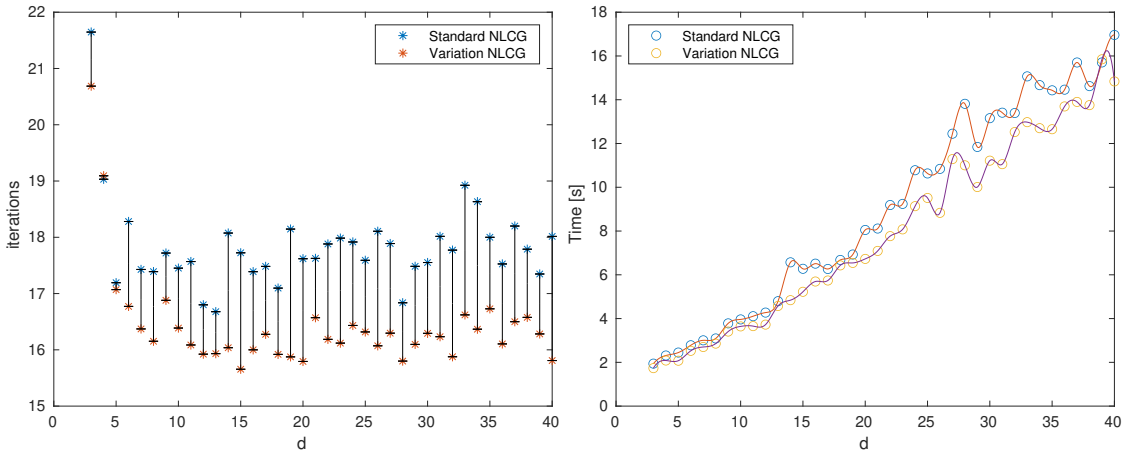


Figure 6.32: Left: Comparison of the number of line searches needed by the algorithms to reach global minima: the variation needs less iterations leading to a gain of 12% in time of convergence. Right: Runtime comparison to global minima.

Summing up the section, our variation of the NLCG correctly approximates the global minimum on both MPS$(2,3,d)$ and hMPS$(2,3,d)$. We notice that, in the case of hMPS$(2,3,d)$ the precision of the minimum is higher (10^{-12} against 10^{-11}), at the same gradient's tolerance. Our Algorithm 6 gains around 22% of runtime to convergence on MPS$(2,3,d)$, for $d \in \{3,\dots,12\}$ and 12% of runtime to *global minima* on hMPS$(2,3,d)$, for $d \in \{3,\dots,40\}$. In the case of MPS$(2,3,d)$ the speed-up is due to the *reduction of the runtime of each line search*, in the region in which less iterations of the line search are needed. However, we expect that by increasing the number of sites, the advantage is lost because of the opposite trend of the number of line searches. In the case of hMPS$(2,3,d)$ the speed-up is due to the *lower number of line searches*, retaining the comparable time of the line search. In this case, we expect that our variation is always faster than the standard NLCG. On hMPS$(2,3,d)$, our Algorithm 6 also can fail to reach convergence or it can converge to the "false local minima" that, in general, need more line searches (and time) to converge. In practice, a proper *restriction* of the iteration bound gives a probability of 59% of finding the ground state in the averaged runtime described in Figure 6.32.

Conclusions. In this chapter, we show a simple data analysis of our test of the nonlinear conjugate gradient method, for approximating ground states of the AKLT on matrix product states with open boundary conditions. We have proposed a variation of the algorithm, based on dimensional considerations, which modifies the line search, the most expensive routine of the algorithm. We moved our first steps towards the study of the NLCG performances and the comparison with our variation. A first analysis demonstrates that the variation has a slightly faster runtime to convergence compared to the standard NLCG.

The analysis of the algorithms of this last chapter of the thesis is preliminary and it would need further work; first of all, a direct comparison with the existing algorithm. In what follows we want to make some qualitative considerations, mainly concerning our algorithm implemented on homogeneous matrix product states.

Consider the well-known Density Matrix Renormalization Group (DMRG) algorithm. The algorithm sequentially solves minor diagonalization problems: all but one or two tensors of the MPS are fixed, an eigensolver is performed on the reduced eigenvalue problem and then the MPS form (that can have been lost) is restored through a Singular Value Decomposition. In terms of runtime to convergence, the DMRG applied to the AKLT model is extremely fast: for $d = 40$, bond dimension fixed to 2 and 4 complete sweeps of the chain, it takes around 5 seconds to converge, providing a precision of 10^{-9} with respect to the global minimum. Allowing the bond dimension to increase from 2 to 3, the precision is enhanced to 10^{-12}, essentially with the same runtime. The NLCG takes around 20 seconds to reach 10^{-12}. Even if the computational costs of the functional contraction and gradient contraction are comparable to those of the DMRG, we cannot control either how many computations of them are needed to complete each line search or how many iterations will be performed to reach convergence. This is

essentially the bottleneck of the algorithm that we tried to modify with our proposition. Therefore, in terms of runtime, the NLCG and our variation are not much competitive compared to the DMRG. One difference we highlight is that the DMRG cannot guarantee a monotonically decreasing energy and it needs an increase in the bond dimension to provide a more precise convergence, contrary to the gradient descent of the NLCG, which ensures convergence with fixed bond dimension equal to two, clearly at the cost of runtime. One of the further works would be a study of the runtime of the NLCG imposing different tolerances of the gradient. The other main difference which makes NLCG interesting is the fact that, contrary to the DMRG (and the majority of sequential algorithms), it always preserves the symmetries of the network. In particular, in the homogeneous matrix product states model, the NLCG updates the single inner tensors A along the whole chain. A very natural further work would be the implementation of the NLCG for matrix product states with periodic boundary conditions, exploiting this global and preserving update of the network. In general, the global method allows once for all fixing the tensor network format and its properties. From a geometric point of view and differently from the sequential methods, the search of ground states does not need to move outside the variety and subsequently project into it. As far as we know, the only non global algorithm that preserves the translation invariance of the MPS is the VUMPS algorithm [ZSVF+18]. However, the algorithm applies to translation invariant matrix product states in the thermodynamic limit, in contrast with the NLCG, which works in the homogeneous finite chain model (hMPS).

We tested the algorithm on the AKLT integrable model, whose ground states, in the finite open spin chain, are exactly tensors of $\mathrm{MPS}(2, 3, d)$ and $\mathrm{hMPS}(2, 3, d)$, that we used as the natural variational varieties for the methods. We have verified that the NLCG still finds the correct solution for ground states of the AKLT Hamiltonian increasing the bond dimensions. Therefore we expect that the algorithm could be applied to other 1-dimensional spin models, with different entanglement properties. However, the increasing of bond dimensions, and thus the number of parameters of the model, would inevitably cost in terms of runtime, which is already high compared for example to the runtime of the DMRG.

In a more general framework, our variation of the NLCG could in principle be tested in any case in which an overabundant parametrization is given, whose fiber structure is determined by the action of a group. In this case, the kernel of the differential of the parametrization coincides with the tangent space of the group orbit and it can be computed via the Lie algebra of the group. The complementary space, that we decide to fix in each line search, is the space that naturally contains the gradient of the functional we are minimizing on the parametrized variety.

Conclusions and further works

In this thesis, we studied tensor network varieties, which are varieties of tensors determined by the combinatorial structure of a given graph. In particular, we focused on the problem of determining their dimension, one of the most basic and fundamental geometric properties of the variety. In the framework of tensor network theory, many open problems concerning the dimension of tensor network varieties were and are still open.

In this thesis, we contribute to deepening the understanding of the dimension of tensor network varieties with a completely general result, which gives an upper bound on the dimension of the tensor network variety. With particular attention to cases relevant for application, we refine the bound for matrix product states and projected entangled pair states. Unexpectedly, we prove that matrix product states with bond dimension two can have a dimension that is not completely controlled by the gauge subgroup, as generically foreseen in literature. We characterize the dimension in the encountered cases and we conjecture that the value of the dimension of matrix product states of bond dimension two, is the "expected" one with the only exceptions we classify. There are several paths left open. First of all deepen the geometrical understanding of the "defective cases", since this phenomenon has not a general geometric characterization yet. On the other hand, it would be interesting to have a general understanding of lower bounds for the dimension of tensor network varieties, but this is a challenging problem. Determining lower bounds on the dimension can be reduced to determining lower bounds on the rank of the differential of the parametrization of the variety at a point; however, determining a suitable point, and computing such rank is non-trivial. An indirect method to determine lower bounds on the dimension of tensor network varieties consists in determining subvarieties of known dimension contained in it. We used this technique for characterizing the "defective cases" of matrix product states. This perspective has the power to link tensor network varieties to other classically studied varieties of tensors, such as determinantal varieties, Segre varieties and secant varieties to Segre varieties.

The second main contribution of the thesis concerns uniform matrix product states, which are translation invariant matrix product states. Matrix product states and uniform matrix product states earn their main relevance in the field of quantum many-body physics, where they represent particles placed on a ring and on an infinite chain, respectively. They are effectively used as variational ansatz classes for the simulation of

quantum spin chains. From the mathematical point of view, uniform matrix product states are parametrized by traces of products of matrices. This fact naturally links them to the classical trace algebra theory and consequently to the use of representation theory techniques. Even in this very relevant and first nontrivial class of tensor network variety, several problems related to the dimension are not completely solved. In the thesis, we approach the problem of studying the linear span of uniform matrix product states. The problem proves to be challenging even in the first nontrivial case of bond and local dimensions equal to two. In this case, we provide an upper bound on the dimension of the linear span. Then, we conjecture a value of the dimension which is in agreement with the given upper bound and which is strongly supported by numerical computations. In the general setting instead, we prove a theorem that provides nontrivial trace relations for the variety of uniform matrix product states. In particular, the result implies that the linear span of the variety is strictly included in the ambient space as long as the number of sites is at least quadratic in the bond dimension, improving the state of the art.

Finally, we address the ground state approximation problem. We set up the theoretic framework to implement the nonlinear conjugate gradient method on the variational class of matrix product states with open boundary conditions. We propose a variation of the method which shows that the knowledge on the dimension of the variety can enhance the performances of the global algorithm. The analysis that we provide in this framework is preliminary and requires further work. The most interesting attempt we want to highlight is the application of the nonlinear conjugate gradient method and our proposition on the homogeneous matrix product states. Since the dimension of this variety is site-independent, the runtime of the algorithm is acceptable taking into account that the method globally updates the whole parameter space. Further, the method preserves the symmetry properties of the network also in the finite chain configuration, contrary to the majority of sequential methods.

Tensor network varieties are attractive geometric objects. We venture into the problem of their dimension and we contribute by providing our original theoretical results. We could investigate the topic from three different perspectives. We prove the general upper bound on the dimension of the tensor network variety using algebraic geometric techniques. We study the linear span of uniform matrix product states through representation theory and the Cayley-Hamilton theorem, which lies at the basis of classical trace algebra theory. We approach the problem of determining ground states, arising from quantum physics, which has led us to the study of a global numerical method and to move our first steps towards the development of a variation of the algorithm, based again on dimensional considerations.

Bibliography

[AH95] J. Alexander and A. Hirschowitz, *Polynomial interpolation in several variables*, J. Alg. Geom. **4** (1995), no. 2, 201–222.

[AKLT88] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, *Valence bond ground states in isotropic quantum antiferromagnets*, Condensed matter physics and exactly soluble models, Springer, 1988, pp. 253–304.

[AMS09] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*, Optimization Algorithms on Matrix Manifolds, Princeton University Press, 2009.

[AR08] E. S. Allman and J. A. Rhodes, *Phylogenetic ideals and varieties for the general Markov model*, Advances in Applied Mathematics **40** (2008), no. 2, 127–148.

[BBM15] W. Buczyńska, J. Buczyński, and M. Michałek, *The Hackbusch conjecture on tensor formats*, J. Math. Pures Appl. **104** (2015), no. 4, 749–761.

[BC92] M. Ben-Or and R. Cleve, *Computing algebraic formulas using a constant number of registers*, SIAM J. Comput. **21** (1992), no. 21, 54–58.

[BDLG21] A. Bernardi, C. De Lazzari, and F. Gesmundo, *Dimension of tensor network varieties*, GitHub page (2021), https://fulges.github.io/code/BDG-DimensionTNS.html.

[BDLG22] ———, *Dimension of tensor network varieties*, Communication in Contemporary Mathematics **0** (2022), no. 0, 2250059.

[Ben09] Y. Bengio, *Learning deep architectures for AI*, Now Publishers Inc, 2009.

[Blä13] M. Bläser, *Fast Matrix Multiplication*, Theory of Computing, Graduate Surveys **5** (2013), 1–60.

[BLF22] T. Barthel, J. Lu, and G. Friesecke, *On the closedness and geometry of tensor network state sets*, Letters in Mathematical Physics **112** (2022), no. 4, 1–33.

[BM05] N. Bray and J. R. Morton, *Equations defining hidden markov models*, Algebraic statistics for computational biology, Cambridge University Press, 2005, pp. 237–249.

[Bri05] M. Brion, *Lectures on the geometry of flag varieties*, pp. 33–85, Birkhäuser Basel, 2005.

[BRVR18] J. Bryan, Z. Reichstein, and M. Van Raamsdonk, *Existence of locally maximally entangled quantum states via geometric invariant theory*, Annales Henri Poincaré, vol. 19, Springer, 2018, pp. 2491–2511.

[BSU16] M. Bachmayr, R. Schneider, and A. Uschmajew, *Tensor networks and hierarchical tensors for the solution of high-dimensional partial differential equations*, Foundations of Computational Mathematics **16** (2016), no. 6, 1423–1472.

[CCX+18] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, *Equivalence of restricted boltzmann machines and tensor network states*, Physical Review B **97** (2018), no. 8, 085104.

[CGFmcW21] M. Christandl, F. Gesmundo, D. S. França, and A. H. Werner, *Optimization at the boundary of the tensor network variety*, Phys. Rev. B **103** (2021), 195139.

[CGL+20] A. Conner, F. Gesmundo, J. M. Landsberg, E. Ventura, and Y. Wang, *Towards a geometric approach to Strassen's asymptotic rank conjecture*, Collectanea Mathematica (2020), 1–24.

[CGMZ21] M. Christandl, F. Gesmundo, M. Michałek, and J. Zuiddam, *Border rank nonadditivity for higher order tensors*, SIAM Journal on Matrix Analysis and Applications **42** (2021), no. 2, 503–527.

[CLO+16] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic, *Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions*, Foundations and Trends in Machine Learning **9** (2016), no. 4-5, 249–429.

[CLVW20] M. Christandl, A. Lucia, P. Vrana, and A. H. Werner, *Tensor network representations from the geometry of entangled states*, SciPost Phys. **9** (2020), no. 3, 42.

[CM14] A. Critch and J. R. Morton, *Algebraic geometry of matrix product states*, Symmetry Integrability and Geometry-methods and Applications **10** (2014), 095.

[CMS19] A. Czapliński, M. Michałek, and T. Seynnaeve, *Uniform matrix product states from an algebraic geometer's point of view*, arXiv: Mathematical Physics (2019).

[CPGSV21] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, *Matrix product states and projected entangled pair states: Concepts, symmetries, theorems*, Rev. Mod. Phys. **93** (2021), 045003.

[CTDL+77] Cohen-Tannoudji, B. Diu, F. Laloë, S. R. Hemley, and N. Ostrowsky, *Quantum machanics, v. 1*, John Wiley, 1977.

[CVZ19] M. Christandl, P. Vrana, and J. Zuiddam, *Asymptotic tensor rank of graph tensors: beyond matrix multiplication*, Computational Complexity **28** (2019), no. 1, 57–111.

[dG78] H. F. de Groote, *On varieties of optimal algorithms for the computation of bilinear mappings I. The isotropy group of a bilinear mapping*, Th. Comp. Science **7** (1978), no. 1, 1–24.

[DL22] C. De Lazzari, *Variational nonlinear conjugate gradient on matrix product states with open boundary conditions*, GitHub repository (2022), https://github.com/claudia-dela/NLCG_MPS_open-boundaries/.

[DLMS22a] C. De Lazzari, H. J. Motwani, and T. Seynnaeve, *The linear span of uniform matrix product states*, arXiv preprint arXiv:2204.10363 (2022).

[DLMS22b] ———, *The linear span of uniform matrix product states*, GitHub repository (2022), https://github.com/harshitmotwani2015/uMPS/.

[DLMS22c] ———, *The linear span of uniform matrix product states*, GitHub repository (2022), https://github.com/claudia-dela/uMPS_highest-weight-vectors/.

[DM21] H. Derksen and V. Makam, *Maximum likelihood estimation for matrix normal models via quiver representations*, SIAM Journal on Applied Algebra and Geometry **5** (2021), no. 2, 338–365.

[DMPY12] Z. Dvir, G. Malod, S. Perifel, and A. Yehudayof, *Separating multilinear branching program formulas*, Proc. of the 44th Annual ACM Symp. on Th. of Comp. (New York), STOC '12, 2012, pp. 615–624.

[DMW20] H. Derksen, V. Makam, and M. Walter, *Maximum likelihood estimation for tensor normal models via castling transforms*, arXiv:2011.03849 (2020).

[ECP10] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, Reviews of modern physics **82** (2010), no. 1, 277.

[EH16] D. Eisenbud and J. Harris, *3264 and All That - A Second Course in Algebraic Geometry*, Cambridge University Press, Cambridge, 2016.

[ERSS05] N. Eriksson, K. Ranestad, B. Sturmfels, and S. Sullivant, *Phylogenetic algebraic geometry*, Projective varieties with unexpected properties (2005), 237–255.

[EV11] G. Evenbly and G. Vidal, *Tensor network states and geometry*, Journal of Statistical Physics **145** (2011), no. 4, 891–918.

[EV14] ———, *Algorithms for entanglement renormalization: boundaries, impurities and interfaces*, Journal of Statistical Physics **157** (2014), no. 4, 931–978.

[FH13] W. Fulton and J. Harris, *Representation theory: a first course*, vol. 129, Springer Science & Business Media, 2013.

[FNW92] M. Fannes, B. Nachtergaele, and R. F. Werner, *Finitely correlated states on quantum spin chains*, Communications in mathematical physics **144** (1992), no. 3, 443–490.

[FR64] R. Fletcher and C. M. Reeves, *Function minimization by conjugate gradients*, The Computer Journal **7** (1964), no. 2, 149–154.

[Ges16] F. Gesmundo, *Geometric Aspects of Iterated Matrix Multiplication*, J. Algebra **461** (2016), 42–64.

[GLW18] F. Gesmundo, J. M. Landsberg, and M. Walter, *Matrix product states and the quantum max-flow/min-cut conjectures*, J. Math. Phys. **55** (2018), no. 10, 102205.

[Gra10] L. Grasedyck, *Hierarchical singular value decomposition of tensors*, SIAM journal on matrix analysis and applications **31** (2010), no. 4, 2029–2054.

[Gre14] J. Greene, *Traces of matrix products*, The Electronic Journal of Linear Algebra **27** (2014), 716–734.

[GS20] D. R. Grayson and M. E. Stillman, *Macaulay 2, a software system for research in algebraic geometry*, http://www.math.uiuc.edu/Macaulay2/, July 2020, version 1.16.

[Hac12] W. Hackbusch, *Tensor spaces and numerical tensor calculus*, vol. 42, Springer, 2012.

[Hac19] ———, *Tensor spaces and numerical tensor calculus*, Springer Series in Computational Mathematics, vol. 56, Springer Cham, 2019.

[Hal15] B. Hall, *Lie groups, lie algebras, and representations, volume 222 of graduate texts in mathematics*, Springer International Publishing, Cham, Switzerland, **21** (2015), 22–25.

[Har13] R. Hartshorne, *Algebraic geometry*, vol. 52, Springer Science & Business Media, 2013.

[Has07] M. B. Hastings, *An area law for one-dimensional quantum systems*, Journal of statistical mechanics: theory and experiment **2007** (2007), no. 08, P08024.

[HCO+11] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, *Time-dependent variational principle for quantum lattices*, Phys. Rev. Lett. **107** (2011), 070601.

[HGS+20] L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, and J. I. Cirac, *Geometry of variational methods: dynamics of closed quantum systems*, SciPost Phys. **9** (2020), 48.

[HLO+16] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, *Unifying time evolution and optimization with matrix product states*, Physical Review B **94** (2016), no. 16, 165116.

[HLT12] F. Holweck, J.-G. Luque, and J.-Y. Thibon, *Geometric descriptions of entangled states by auxiliary varieties*, J. Math. Phys. **53** (2012), no. 10, 102203.

[HMOV14] J. Haegeman, M. Mariën, T. J. Osborne, and F. Verstraete, *Geometry of matrix product states: Metric, parallel transport, and curvature*, J. Math. Phys. **55** (2014), no. 2, 021902.

[HOV13] J. Haegeman, T. J. Osborne, and F. Verstraete, *Post-matrix product state methods: To tangent space and beyond*, Physical Review B **88** (2013), no. 7, 075133.

[HP18] J. Hauschild and F. Pollmann, *Efficient numerical simulations with tensor networks: Tensor network python (tenpy)*, SciPost Physics Lecture Notes (2018), 005.

[HRS12] S. Holtz, T. Rohwedder, and R. Schneider, *On manifolds of tensors of fixed tt-rank*, Numerische Mathematik **120** (2012), no. 4, 701–731.

[HVDH21] M. Hauru, M. Van Damme, and J. Haegeman, *Riemannian optimization of isometric tensor networks*, SciPost Physics **10** (2021), no. 2, 040.

[KN79] G. Kempf and L. Ness, *The length of vectors in representation spaces*, Lecture Notes in Math. **732** (1979), 233–243.

[Lan12] J. M. Landsberg, *Tensors: geometry and applications*, Representation theory **381** (2012), no. 402, 3.

[Lan17] ———, *Geometry and complexity theory*, Cambridge Studies in Advanced Mathematics, vol. 169, Cambridge University Press, Cambridge, 2017.

[Lau96] S. L. Lauritzen, *Graphical models*, Oxford Statistical Science Series, vol. 17, The Clarendon Press Oxford University Press, New York, 1996.

[Lee13] J. M Lee, *Smooth manifolds*, Introduction to smooth manifolds, Springer, 2013, pp. 1–31.

[LOV15] C. Lubich, I. V. Oseledets, and B. Vandereycken, *Time integration of tensor trains*, SIAM Journal on Numerical Analysis **53** (2015), no. 2, 917–941.

[LQY12] J. M. Landsberg, Y. Qi, and K. Ye, *On the geometry of Tensor Network States*, Quantum Inf. Comput. **12** (2012), no. 3-4, 346–254.

[LT03] J.-G. Luque and J.-Y. Thibon, *Polynomial invariants of four qubits*, Phys. Rev. A **67** (2003), no. 4, 042303, 5.

[Mat20] The Mathworks, Inc., Natick, Massachusetts, *MATLAB version 9.7.0.1319299 (R2019b)*, 2020.

[MFK94] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, Springer Science & Business Media, vol. 34, Springer-Verlag, Berlin, 1994.

[MHO13] A. Milsted, J. Haegeman, and T. J. Osborne, *Matrix product states and variational methods applied to critical quantum field theory*, Phys. Rev. D **88** (2013), 085030.

[Mov17] R. Movassagh, *Generic local hamiltonians are gapless*, Physical review letters **119** (2017), no. 22, 220504.

[MSV19] M. Michałek, T. Seynnaeve, and F. Verstraete, *A tensor version of the quantum Wielandt theorem*, SIAM Journal on Matrix Analysis and Applications **40** (2019), no. 3, 1125–1130.

[NV18] Miguel Navascues and Tamas Vertesi, *Bond dimension witnesses and the structure of homogeneous matrix product states*, Quantum **2** (2018), 50.

[NW06] J. Nocedal and S. Wright, *Numerical optimization*, Springer Science & Business Media, 2006.

[ÖR95] S. Östlund and S. Rommer, *Thermodynamic limit of density matrix renormalization*, Phys. Rev. Lett. **75** (1995), 3537–3540.

[Orú14] R. Orús, *A practical introduction to tensor networks: Matrix product states and projected entangled pair states*, Annals of Physics **349** (2014), 117–158.

[Ose11] I. V. Oseledets, *Tensor-train decomposition*, SIAM Journal on Scientific Computing **33** (2011), no. 5, 2295–2317.

[OT09] I. V. Oseledets and E. E. Tyrtyshnikov, *Breaking the curse of dimensionality, or how to use svd in many dimensions*, SIAM Journal on Scientific Computing **31** (2009), no. 5, 3744–3759.

[PESV14] R. NC Pfeifer, G. Evenbly, S. Singh, and G. Vidal, *Ncon: A tensor network contractor for matlab*, arXiv:1402.0939 (2014).

[PGVWC07] D. Pérez-García, F. Verstraete, M. M. Wolf, and J. I. Cirac, *Matrix product state representations*, Quantum Information & Computation **7** (2007), no. 5, 401–430.

[PP+08] K. B. Petersen, M. S. Pedersen, et al., *The matrix cookbook*, Technical University of Denmark **7** (2008), no. 15, 510.

[Pro76] C. Procesi, *The invariant theory of n× n matrices*, Advances in mathematics **19** (1976), no. 3, 306–381.

[Pro07] ———, *Lie groups: An approach through invariants and representations*, Universitext, Springer, New York, 2007.

[PVV11] B. Pirvu, F. Verstraete, and G. Vidal, *Exploiting translational invariance in matrix product state simulations of spin chains with periodic boundary conditions*, Phys. Rev. B **83** (2011), 125104.

[RS78] M. Reed and B. Simon, *Iv: Analysis of operators*, vol. 4, Elsevier, 1978.

[RS19] E. Robeva and A. Seigal, *Duality of graphical models and tensor networks*, Inf. Inference **8** (2019), no. 2, 273–288.

[Sch11] U. Schollwöck, *The density-matrix renormalization group in the age of matrix product states*, Annals of physics **326** (2011), no. 1, 96–192.

[Sha94] I. R. Shafarevich, *Basic algebraic geometry. 1 - Varieties in projective space*, second ed., Springer-Verlag, Berlin, 1994.

[Sib68] K.S. Sibirskii, *Algebraic invariants for a set of matrices*, Siberian Mathematical Journal **9** (1968), no. 1, 115–124.

[Sta13] R. P. Stanley, *Algebraic combinatorics*, Undergraduate Texts in Mathematics, Springer, New York, 2013, Walks, trees, tableaux, and more.

[STG+19] P. Silvi, F. Tschirsich, M. Gerster, J. Jünemann, D. Jaschke, M. Rizzi, and S. Montangero, *The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems*, SciPost Physics Lecture Notes (2019), 008.

[The20] The Sage Developers, *Sagemath, the Sage Mathematics Software System (Version 9.0)*, 2020, `https://www.sagemath.org`.

[VC04] F. Verstraete and J. I. Cirac, *Renormalization algorithms for quantum-many body systems in two and higher dimensions*, arXiv preprint cond-mat/0407066 (2004).

[VC17] P. Vrana and M. Christandl, *Entanglement distillation from Greenberger-Horne-Zeilinger shares*, Comm. in Math. Ph. **352** (2017), no. 2, 621–627.

[VHCV16] L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, *Gradient methods for variational optimization of projected entangled-pair states*, Physical Review B **94** (2016), no. 15, 155123.

[VHV19] L. Vanderstraeten, J. Haegeman, and F. Verstraete, *Tangent-space methods for uniform matrix product states*, SciPost Phys. Lect. Notes (2019), 7.

[Vid04] G. Vidal, *Efficient simulation of one-dimensional quantum many-body systems*, Physical review letters **93** (2004), no. 4, 040502.

[Vid07] ———, *Entanglement renormalization*, Physical review letters **99** (2007), no. 22, 220405.

[Vid08] ———, *Class of quantum many-body states that can be efficiently simulated*, Physical review letters **101** (2008), no. 11, 110501.

[VMC08] F. Verstraete, V. Murg, and J. I. Cirac, *Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems*, Advances in Physics **57** (2008), no. 2, 143–224.

[VWPGC06] F. Verstraete, M. M. Wolf, D. Pérez-García, and J. I. Cirac, *Criticality, the area law, and the computational power of projected entangled pair states*, Physical review letters **96** (2006), no. 22, 220601.

[Whi93] S. R. White, *Density-matrix algorithms for quantum renormalization groups*, Physical review b **48** (1993), no. 14, 10345.

[YL18] K. Ye and L.-H. Lim, *Tensor network ranks*, arXiv:1801.02662 (2018).

[Zak93] F. L. Zak, *Tangents and secants of algebraic varieties*, vol. 127, American Mathematical Soc., 1993.

[ZSVF+18] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman, *Variational optimization algorithms for uniform matrix product states*, Physical Review B **97** (2018), no. 4, 045145.