Defectivness of Tensor Network Varieties

[A. Bernardi, C. De Lazzari, F. Gesmundo]

Alessandra Bernardi

University of Trento, Italy

Feb 8, 2024

Overview

- Motivations
- Definitions
- Dimension and defectiveness
- Further Questions

Motivations - Why TNS

From Quantum Physics

- Tensor space has high dimension: $\dim(V^{\otimes d}) = \dim(V_i)^d$. Quickly intractable. Requires too large memory to reprensent a tensor.
- Given a quantum many-body wave function, specifying its coefficients in a given local basis does not give any intuition about the structure of the *entanglement* between its constituents:

$$e_0 \otimes e_0 \otimes e_0 + e_1 \otimes e_1 \otimes e_1$$

$$T = \sum_{i,j,k=1}^{d} t_{i,j,k} e_i \otimes e_j \otimes e_k$$

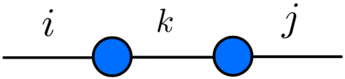
with $\{e_l\}$ orthonormal and $t_{i,j,k} \in \mathbb{R}_{>0}$

Motivations - Why TNS

A Tensor Network has this information directly available in its description in terms of a network of quantum correlations (physicists "mantra": *The interesting states live in a corner of the Hilbert space, the same where TNS are*).

Matrix product $AB = C$: $\sum_{k=1}^{m} a_{i,k} b_{k,j} = (c_{i,j})_{i=1,\ldots,n_1, j=1,\ldots,n_2}$.

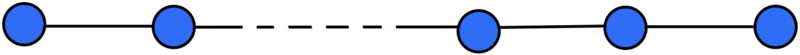
The network of correlations makes explicit the effective lattice geometry in which the state actually lives



A TN is a set of tensors where some, or all, indices are contracted according to some pattern.

Motivations - Why TNS

Matrix product states



Reduced number of parameters

$$dm^2 \dim(V) << \dim(V)^d$$

MPS are accurate representations of physical states with limited bond length m.

Highlight entangled structure of state. The corresponding spaces of tensors are only locally entangled because interactions (entanglement) in the physical world appear to just happen locally.

Motivations - Why TNS

So if physicists know in advance that the object of interest has to be a TNS (or very close to), than one looks only inside the region where TNS are.

For example: Computing the ground state of a Hamiltonian

$$\rho : \mathcal{H} \to \mathbb{R},\ v \mapsto \rho(v) = \frac{v^\dagger H v}{v^\dagger v}$$

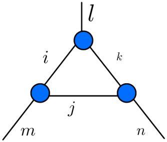
$$\lambda_0 = min\{\rho(v) : v \in \mathcal{H}\}$$

one can look only at

$$\lambda_0 \sim \min\{\rho(v) : v \in TNS\}.$$

Algorithms for approximating tensors on TNS, algorithms to evolve Hamiltonians on TN...

Motivations - Why dimension of TNS



$$T = \sum_{i,j,k} T_{mij} T_{lik} T_{njk} \in V_m \otimes V_l \otimes V_n. \qquad (1)$$

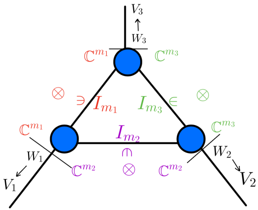
$$TNS_{\Gamma(v,e)} = \overline{\{T \in V_m \otimes V_l \otimes V_n \mid \exists T_{mij}, T_{lik}, T_{njk} s.t. T = (1)\}}$$

Motivations - Why dimension of TNS

If we are interested in study a paremetrized object we need to know if the number of parameters is essential, and, if not, to reduce them as much as possible.
The right number of parameters is the Dimension:

ϕ : Parameter space $\rightarrow$ TNS variety
$\dim(\phi^{-1}(T)) = ?$

Definition - Graph Tensor



- Fix a graph $\Gamma(\mathbf{v}(\Gamma), \mathbf{e}(\Gamma))$, $d := \sharp v(\Gamma)$
- Fix the weights $\mathbf{m} = (m_e, e \in \mathbf{e}(\Gamma))$ = *bond dimensions*
- Consider $I_{m_e} \in \mathbb{C}^{m_e} \otimes \mathbb{C}^{m_e}$ at e and Tensor them: $\bigotimes_{e \in e(\Gamma)} I_{m_e}$
- It naturally lives in $\bigotimes_{e \in e(\Gamma)} \mathbb{C}^{m_e} \otimes \mathbb{C}^{m_e}$ but we think it as an element of $\bigotimes_{v \in v(\Gamma)} (\bigotimes_{e \ni v} \mathbb{C}^{m_e}) := \bigotimes_{v \in v(\Gamma)} W_v$ obtained by grouping together the spaces incident at the same vertex:

$$T(\Gamma, \mathbf{m}) := \bigotimes_{e \in e(\Gamma)} I_{m_e} \in \bigotimes_{v \in v(\Gamma)} W_e$$

Definition - TNS

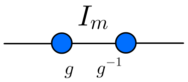
$TNS^{\Gamma}_{\mathbf{m},\mathbf{n}} \subset V_1 \otimes \cdots \otimes V_d$ associated to the tensor network $(\Gamma, \mathbf{m}, \mathbf{n})$

$$\begin{array}{rcl} \Phi : Hom(W_1, V_1) \times \cdots \times Hom(W_d, V_d) & \to & V_1 \otimes \cdots \otimes V_d \\ (X_1, \ldots, X_d) & \mapsto & (X_1 \otimes \cdots \otimes X_d)(T(\Gamma, \mathbf{m})) \end{array}$$

$$Im(\Phi) = TNS^{\Gamma,0}_{\mathbf{m},\mathbf{n}}$$

$$TNS^{\Gamma}_{\mathbf{m},\mathbf{n}} = \overline{Im(\Phi)} \subset V_1 \otimes \cdots \otimes V_d$$

Example Matrix multiplication



$$\sum_{k=1}^{m} a_{ik} b_{kj}$$

$T(\Gamma, m) = I_m \in \mathbb{C}^m \otimes \mathbb{C}^m = W_1 \otimes W_2$ Fix V_1, V_2

$$\Phi : Hom(W_1, V_1) \times Hom(W_2, V_2) \to V_1 \otimes V_2$$

$$\Phi(X_1, X_2) = (X_1, X_2) \cdot I_m = (X_1, X_2) \cdot \sum_{i=1}^{m} e_i \otimes e_i = \sum_{i=1}^{m} X_1 e_i \otimes X_2 e_i =$$

$$= \sum_{i=1}^{m} X_1 e_i (X_2 e_i)^T = X_1 I_m X_2^T = X_1 X_2^T$$

In this case $TNS^{\Gamma}_{\mathbf{m},\mathbf{n}} = \{M \in V_1 \otimes V_2 : rank(M) \leq m\} = TNS^{\Gamma,0}_{\mathbf{m},\mathbf{n}}$

Why graph tensor is better

The multilinear multiplication is nothing but evaluation. Evaluating the graph tensor $T(\Gamma, m)$ is easier than evaluating other tensors.

- Given $T \in V_1 \otimes \cdots \otimes V_d$ and a graph Γ
- start with small m and evaluate $T(\Gamma, m)$: hope to find linear maps $X_1, \ldots, X_d$ s.t.

$$(X_1 \otimes \cdots \otimes X_d)(T(\Gamma, m)) = T$$

([Christandl-Gesmundo-Stilck Franca- Werner '20] find a class of tensors for which the evaluation is easy)

Properties

- One can assume that all $m_e > 1$, otherwise remove the edge from the graph.
- Monotonicity:

$$\text{If } m' \leq m \text{ (entry-wise) then } TNS^{\Gamma}_{m',n} \subseteq TNS^{\Gamma}_{m,n}$$

- Universality: If Γ is connected then

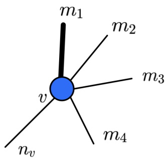
$$TNS^{\Gamma}_{m,n} = V_1 \otimes \cdots \otimes V_d$$

if m_e is large enough for every $e \in e(\Gamma)$.

Reductions

- We may assume all bond dimensions associated to the edges incident a fixed vertex are *balanced*: Fix a vertex v and $e_1, \ldots, e_k \in v$; If

$$m_{e_k} > n_v \cdot m_{e_1} \cdots m_{e_{k-1}}, \qquad m_{e_k} \text{ is } \textit{overabundant}$$



then

$$TNS_{\mathbf{m},\mathbf{n}} = TNS_{\overline{\mathbf{m}},\mathbf{n}}$$

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

Reductions

Definition (Landsberg-Qi-Ye '12)

A vertex $v \in \mathbf{v}$ is called

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

Theorem (BDG'23)

If the vertex v is supercritical let $N = \dim W_d = \prod_{e \ni d} m_e$ and $\mathbf{n}' = (n'_v : v \in \mathbf{v}(\Gamma))$ be the d-tuple of local dimensions s.t. $n'_v = n_v$ if $v \neq d$ and $n'_d = N$. Then

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

Studying the orbit of $T(\Gamma, m)$ does not say anything about tensors in $TNS_\Gamma(m,n) \setminus TNS^0_\Gamma(m,n)$.

Theorem (Landsberg-Qi-Ye '12)

- *If* Γ *doesn't have cycles, then* $TNS^0_\Gamma(m,n) = TNS_\Gamma(m,n)$
- *otherwise* $TNS_\Gamma(m,n) \setminus TNS^0_\Gamma(m,n) \neq \emptyset$

[Christandl-Lucia-Vrana-Werner '20] There exist tensors of physical interest on the boundary.

What's known

- [Haegeman, Marien, Osborne, Verstraete, 2014]: MPS with open boundary conditions.
- [Buczynska, Buczynski, Michalek, 2015]: Perfect binary trees, Train train.

Dimension

If $f : X \to Y$ map between varieties, then

$$\dim(\overline{Im(f)}) = \dim X - \dim f^{-1}(y)$$

for y generic in $Im(f)$.

We study the fibers of

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

$$(X_1, \ldots, X_d) \mapsto (X_1 \otimes \cdots \otimes X_d)(T(\Gamma, \mathbf{m}))$$

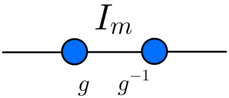
Obviously in the fiber

Ex: Matrix case

$$\Phi : Hom(\mathbb{C}^m, V_1) \times Hom(\mathbb{C}^m, V_2) \to V_1 \otimes V_2$$

with $\Phi(X_1, X_2) = X_1 \cdot I_m \cdot X_2^t$.

$\Phi(X_1, X_2) = \Phi(X_1 g, X_2 (g^{-1})^t)$ for every $g \in GL_m$.
The fiber containing (X_1, X_2) contains the entire GL_m-orbit.



$$\sum_{k=1}^{m} a_{ik} b_{kj}$$

The fiber containing $(X_v : v \in \mathbf{v}(\Gamma))$ contains its entire $\mathcal{G}_{\Gamma,\mathbf{m}}$-orbit, where $\mathcal{G}_{\Gamma,\mathbf{m}} = \times_{e \in \mathbf{e}(\Gamma)} GL_{m_e}$ *gauge subgroup* of Γ.

The role of this group in the theory of tensor network was known and it is expected that it entirely controls the value of dim TNS. In fact, it is **expected** that in "most" cases the exact value of the dimension is

$$\min\{\underbrace{\sum_v (n_v \times \prod_{e \ni v} m_e) - d + 1}_{\dim \times_v \mathbb{P}(Hom(W_v, V_v))} - \underbrace{\sum_e (m_e^2 - 1)}_{\dim \mathcal{G}_{\Gamma,m}}, \prod_v n_v\}$$

This computation does not take care of two facts:

- the possible existence of the stabilizer under the action of the gauge subgroup of a generic d-tuple of linear maps,
- there may be something else in the fiber.

Main theorem

Theorem (BDG'23)

$$\dim(TNS^{\Gamma}_{m,n}) \leq$$

$$\min\{\underbrace{\sum_{v}(n_v \times \prod_{e\ni v} m_e) - d + 1}_{\dim \times_v \mathbb{P}(Hom(W_v, V_v))} - (\underbrace{\sum_{e}(m_e^2 - 1)}_{\dim \mathcal{G}_{\Gamma,m}} - \underbrace{\dim Stab_{\mathcal{G}_{\Gamma,m}}(X)}_{??}), \prod_{v} n_v\}$$

Luckily...

Theorem (Derksen-Makam-Walter'20)

$\dim(Stab_{\mathcal{G}_{\Gamma,\mathbf{m}}}(X)) = 0$ *in "most" cases*

(the action of $\mathcal{G}_{\Gamma,m}$ on $\times_v Hom(W_v, V_v)$ is generically stable, i.e. there exists an element v in the parameter space s.t. $Stab_G(v)$ is a finite group).

Two important ones:

- Γ *is a cycle, called matrix product states;*
- Γ *is a grid, called projected entangled pair states.*

Theorem (Haegeman-Mariën-Osborne-Verstraete '14)

Matrix product states with open boundary conditions $(m_0 = m_d = 1)$

$$\dim TNS^{\Gamma}_{\mathbf{m},\mathbf{n}} = \min\left\{ \sum_{i=1}^{d} n_i m_{i-1} m_i - \sum_{j=1}^{d-1} m_i^2, \prod_{1}^{d} n_i \right\}$$

Theorem (Buczynska, Buczynski, Michalek '15)

Perfect binary, TT have expected dimension.

Main theorem

Theorem (BDG'23)

If $(\Gamma, \mathbf{m}, \mathbf{n})$ *is a subcritical tensor network with no overabundant bond dimension, then*

$$\dim(TNS^{\Gamma}_{m,n}) \leq$$

$$\min\{\underbrace{\sum_v (n_v \times \prod_{e \ni v} m_e) - d + 1}_{\dim \times_v \mathbb{P}(Hom(W_v, V_v))} - (\underbrace{\sum_e (m_e^2 - 1)}_{\dim \mathcal{G}_{\Gamma, m}} - \underbrace{\dim Stab_{\mathcal{G}_{\Gamma,m}}(X)}_{??}), \prod_v n_v\}$$

If $(\Gamma, \mathbf{m}, \mathbf{n})$ *is a supercritical case the the bound is sharp and* $\dim Stab_{\mathcal{G}_{\Gamma,m}}(X)) = 0$

$$\dim(TNS^{\Gamma}_{m,n}) = \min\{\sum_v (n_v \times \prod_{e \ni v} m_e) - d + 1 - \sum_e (m_e^2 - 1), \prod_v n_v\}$$

Subcritical defective examples

$\Delta: \quad m = (2,2,2)$

	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

$\square: \quad m = (2,2,2,2)$

	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

Δ: $m = (2,2,2)$, $n = (2,3,4)$

- $T(\Gamma, \mathbf{m}) \in \mathbb{C}^{2\times 2} \otimes \mathbb{C}^{2\times 2} \otimes \mathbb{C}^{2\times 2}$
- $TNS_\Gamma(\mathbf{m}, \mathbf{n}) \subseteq \mathbb{P}(\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^4)$.

Let $T \in \mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^4$. Consider the flattening

$$T_1 : \mathbb{C}^2 \to \mathbb{C}^3 \otimes \mathbb{C}^4.$$

Then $L_T = \mathbb{P}(Im(T_1))$ is a line in $\mathbb{P}(\mathbb{C}^3 \otimes \mathbb{C}^4)$ (or a single point).

Theorem (BDG'23)

$T \in TNS_\Gamma(\mathbf{m}, \mathbf{n})$ if and only if

- *either $rank(L_T) = 1$*
- *or L_T intersects $\{A : rank(A) \leq 2\}$ in at least two points (counted with multiplicity).*

$\dim TNS_\Gamma(\mathbf{m}, \mathbf{n}) = 2m(n_1 + n_2 - m) + 1 = 21 < 23$ (proj dim).

In progress

Theorem (BG)

Fix $\mathbf{m} = (2,2,m)$ *and* $\mathbf{n} = (2, n_1, n_2)$, *on the triangle graph* Δ. *Write* $V_i = \mathbb{C}^{n_i}$ *and* $V_0 = \mathbb{C}^2$.

$$\mathcal{TNS}^{\Delta}_{\mathbf{m},\mathbf{n}} = \overline{\left\{ T \in \mathbb{P}(V_0 \otimes V_1 \otimes V_2) : \begin{array}{l} Im(T : V_0^* \to V_1 \otimes V_2) \\ \text{is a line intersecting} \\ \sigma_m^{n_1 \times n_2} \text{ in at least two points} \end{array} \right\}}.$$

In particular, its projective dimension is

$$\dim \mathcal{TNS}^{\Delta}_{\mathbf{m},\mathbf{n}} = 2m(n_1 + n_2 - m) + 1.$$

defectiveness $\delta = m^2 - 1$

Further Questions

- Classify all sub-critical cases where the upper bound is not reached: they have some interesting peculiar geometric properties.
- Which is "the best" $TNS^{\Gamma}_{m,n}$ a given T belongs to?
  - Γ can be reasonably chosen from the context. One may work on decreasing m. How to choose m s.t. a given $T \in TNS^{\Gamma,0}_{m,n}$?
  - Very well established procedures to find a "good enough" approximation of T on a given $TNS^{\Gamma}_{m,n}$.
- How much the reduction of parameters can improve the efficiency of algorithms?

Further Work

Preliminary [DL]:

- She applied the reduction of parameters in a global search (variational NLCG) to possibly improve the search of ground state of the AKLT model on MPS with open boundary conditions (only the Gauge in the fiber).
- The variation of the NLCG she 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: we reduce the number of coordinates of the gradient.
- We notice a gain in runtime to convergence, compared to the standard NLCG. (The global method preserves the symmetries of the tensor network, differently from the majority of sequential methods.)

Further Work

Results on $MPS(\mathbf{m} = 2, \mathbf{n} = 3, \sharp sites = d)$

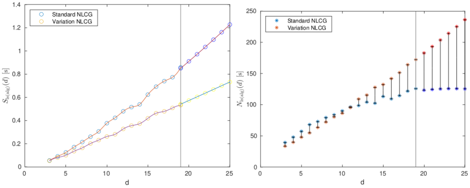


Figure 6.20: Left: Comparison of time of the single line search $S_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 25$: Algorithm 6 is faster in the line search. Right: Comparison of the number of line searches $N_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 25$: for $d > 11$, Algorithm 5 needs less iterations to reach convergence.

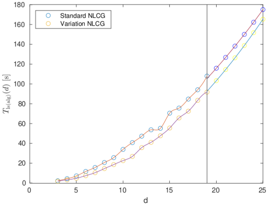


Figure 6.21: Comparison of the overall time of line search $T_{\mathrm{ls(alg)}}(d) = S_{\mathrm{ls(alg)}}(d)N_{\mathrm{ls(alg)}}(d)$, for $d = 3, \ldots, 25$.

Advertisement - hiring PhD

TENORS
Tensor modEliNg, geOmetRy and optimiSation
Marie Skłodowska-Curie Doctoral Network
2024-2027

Tensors are nowadays ubiquitous in many domains of applied mathematics, computer science, signal processing, data processing, machine learning and in the emerging area of quantum computing. TENORS aims at fostering cutting-edge research in tensor sciences, stimulating interdisciplinary and intersectoriality knowledge developments between algebraists, geometers, computer scientists, numerical analysts, data analysts, physicists, quantum scientists, and industrial actors facing real-life tensor-based problems.

Partners:

1. Inria, Sophia Antipolis, France (B. Mourrain, A. Mantzaflaris)
2. CNRS, LAAS, Toulouse, France (D. Henrion, V. Magron, M. Skomra)
3. NWO-I/CWI, Amsterdam, the Netherlands (M. Laurent)
4. Univ. Konstanz, Germany (M. Schweighofer, S. Kuhlmann, M. Michałek)
5. MPI, Leipzig, Germany (B. Sturmfels, S. Telen)
6. Univ. Tromsoe, Norway (C. Riener, C. Bordin, H. Munthe-Kaas)
7. Univ. degli Studi di Firenze, Italy (G. Ottaviani)
8. Univ. degli Studi di Trento, Italy (A. Bernardi, A. Oneto, I. Carusotto)
9. CTU, Prague, Czech Republic (J. Marecek)
10. ICFO, Barcelona, Spain (A. Acin)
11. Artelys SA, Paris, France (M. Gabay)

Associate partners:

1. Quandela, France
2. Cambridge Quantum Computing, UK.
3. Bluetensor, Italy.
4. Arva AS, Norway.
5. HSBC Lab., London, UK.

15 PhD positions (2024-2027)

(recruitment expected around Oct. 2024)

Scientific coord: B. Mourrain
Adm. manager: Linh Nguyen

1

THANKS!