Quantum mechanics, tensor networks and machine learning

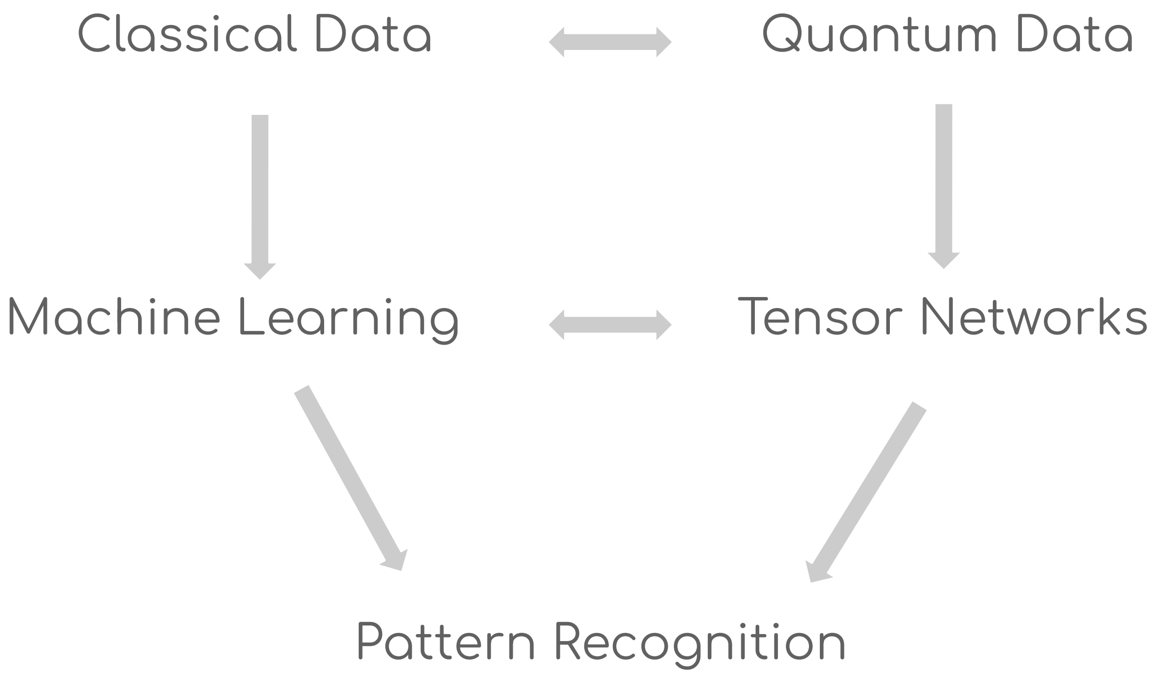
Nicola Pancotti,
quantum research scientist @ Amazon Web Services

in collaboration with

- **Max Planck Quantum Optics:**
  Ivan Glasser, J. Ignacio Cirac, Ivan Diego Rodriguez
- **Technical University Munich:**
  Moritz August
- **Free University Berlin:**
  Ryan Sweke, Jens Eisert

motivations

- Quantum mechanics and machine learning are intrinsically probabilistic theories
- Neural networks and tensor networks are two extremely successful paradigms in their respective fields
- Can we connect their mathematical formulation?
- Can we improve one by using the other?


Classical Data
Quantum Data
Machine Learning
Tensor Networks
Pattern Recognition

content of the talk

1. quantum mechanics and linear algebra
    a. wavefunctions are vectors
    b. observables are matrices
2. linear algebra and tensor networks
    a. Tensor networks as an efficient tool for certain problems in linear algebra
3. tensor networks and machine learning
    a. tensor networks for probabilistic modeling
    b. Examples for supervised and unsupervised learning

quantum mechanics

wavefunctions are vectors

$$\begin{array}{cccccccc} \downarrow & \uparrow & \downarrow & \downarrow & \uparrow & \uparrow & \downarrow & \downarrow \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{array}$$

Notation:

$$0 \to (1,0) = |0\rangle$$
$$1 \to (0,1) = |1\rangle$$

$$0100 \to (1,0) \otimes (0,1) \otimes (1,0) \otimes (1,0) = |0100\rangle$$

A generic vector in this Hilbert space can be expressed as:

$$|v\rangle = \vec{v} = \sum_{\{x\}} \omega_x |x\rangle \quad \text{with} \quad \sum_{\{x\}} |\omega_x|^2 = 1$$

$$|v\rangle = \vec{v} = \sum_{\{x\}} \omega_x |x\rangle$$

observables are matrices

$$\mathcal{O}\vec{v} = \vec{u}$$

two examples
$$\begin{cases} \text{energy} & E = \dfrac{\vec{v}^T H \vec{v}}{\vec{v}^T \vec{v}} \\ \text{magnetization} & M = \dfrac{\vec{v}^T Z \vec{v}}{\vec{v}^T \vec{v}} \end{cases}$$

Dynamical behavior: Schrödinger equation

$$\boxed{i\hbar \frac{d\vec{v}(t)}{dt} = H\vec{v}(t) \rightarrow \vec{v}(t) = e^{-iHt/\hbar}\vec{v}(0)}$$

tensor networks and linear algebra

matrix product representations

$$x : (x_1, x_2, \ldots, x_n)$$
$$x_i \in (0,1)$$

$$|v\rangle = \vec{v} = \sum_{\{x\}} \omega_x |x\rangle$$

$$|MPS\rangle = \sum_{\{x\}} \boxed{A_1^{x_1} A_2^{x_2} A_3^{x_3} \ldots A_N^{x_N}} |x_1 x_2 \ldots x_N\rangle$$

---

in physics, if

- physical degrees of freedom are arranged on a line: one dimension.
- the vector fulfills an area law

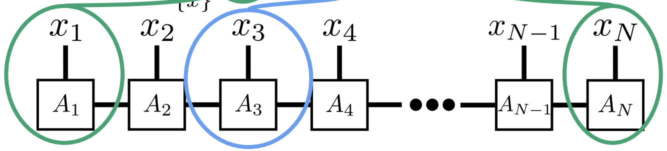
then matrix product states are a faithful representation

$$x : (x_1, x_2, \ldots, x_n)$$
$$x_i \in (0, 1)$$

graphical notation

for vectors

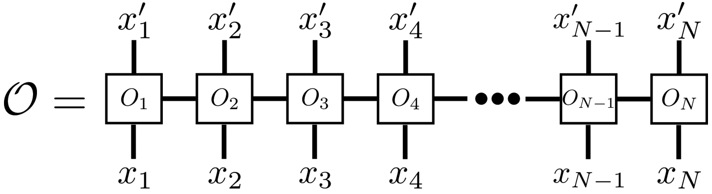
$$|MPS\rangle = \sum_{\{x\}} A_1^{x_1} A_2^{x_2} A_3^{x_3} \ldots A_N^{x_N} |x_1 x_2 \ldots x_N\rangle$$



Rank 2 tensors at the edges

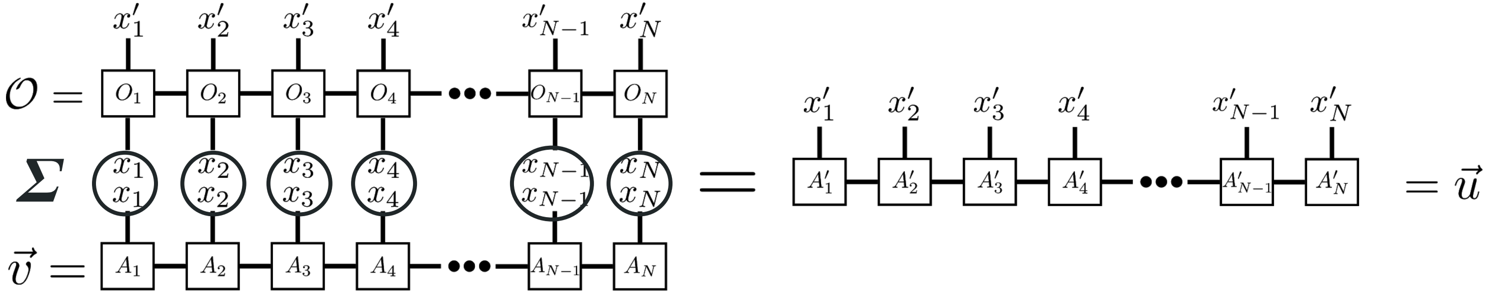
Rank 3 tensors in the bulk

for matrices



Efficient matrix-vector multiplication

$$\mathcal{O}\vec{v} = \vec{u}$$



tensor networks and machine learning

$$|v\rangle = \vec{v} = \sum_{\{x\}} \omega_x |x\rangle$$

Matrix Product States (MPS) :

$$|MPS\rangle = \sum_{\{x\}} A_1^{x_1} A_2^{x_2} A_3^{x_3} \dots A_N^{x_N} |x_1 x_2 \dots x_N\rangle$$

Boltzmann Machines (BM):

$$|BM\rangle = \sum_{\{x_1 x_2 \dots x_N\}} \sum_{\{h\}} \exp\left[\tilde{H}(\mathbf{x}, \mathbf{h})\right] |x_1 x_2 \dots x_N\rangle$$

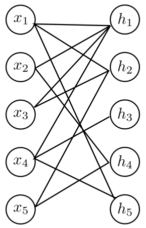
String Bond States (SBS):

$$|SBS\rangle = \sum_{\{x\}} \prod_j A_{1,j}^{x_1} A_{2,j}^{x_2} A_{3,j}^{x_3} \dots A_{N,j}^{x_N} |x_1 x_2 \dots x_N\rangle$$

Restricted Boltzmann Machines (RBM):

$$|RBM\rangle = \sum_{\{x_1 x_2 \dots x_N\}} \prod_j \cosh\left(\sum_i W_{ij} x_i\right) |x_1 x_2 \dots x_N\rangle$$

$$\tilde{H} = \sum_{ij} W_{ij} x_i h_j$$



Restricted Boltzmann Machines are a subclass of string bond states

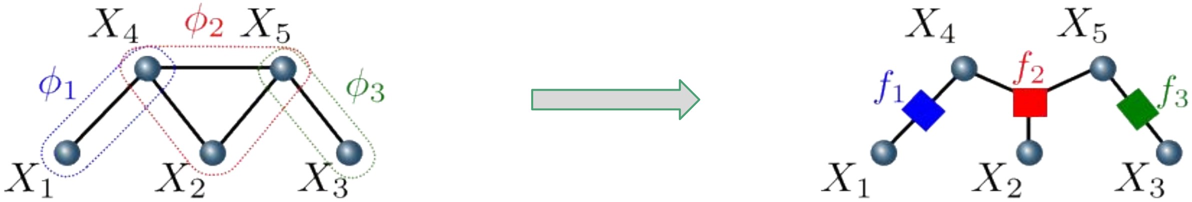
$$|RBM\rangle \rightarrow |SBS\rangle$$

PRX 8 (1), 011006

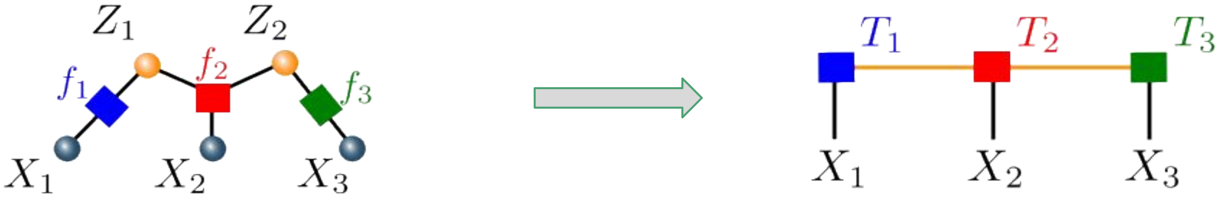
Relationship with graphical models

Graphical models are classical probabilistic models where one assumes a certain factorization of the probability density function

Without hidden units: $P(X_1, X_2, X_3, X_4, X_5) = \phi_1(X_1, X_4)\phi_2(X_4, X_2, X_5)\phi_3(X_5, X_3)$



With hidden units: $P(X_1, X_2, X_3) = \sum_{Z_1, Z_2} f_1(X_1, Z_1) f_2(X_2, Z_1, Z_2) f_3(X_3, Z_2)$

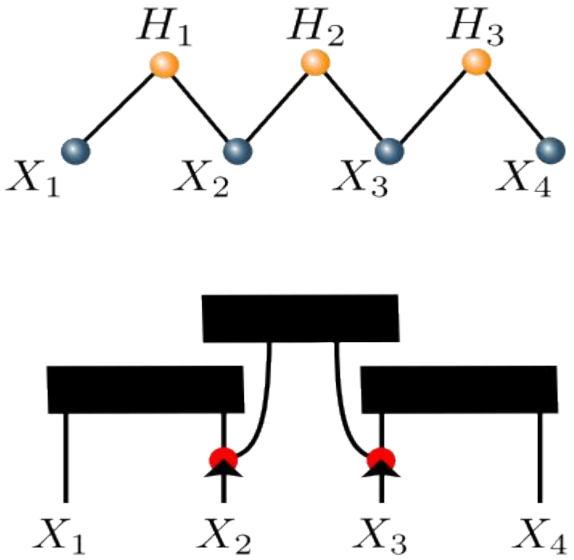


$$|RBM\rangle = \sum_{\{x_1 x_2 \ldots x_N\}} \prod_j \cosh\left(\sum_i W_{ij} x_i\right) |x_1 x_2 \ldots x_N\rangle$$

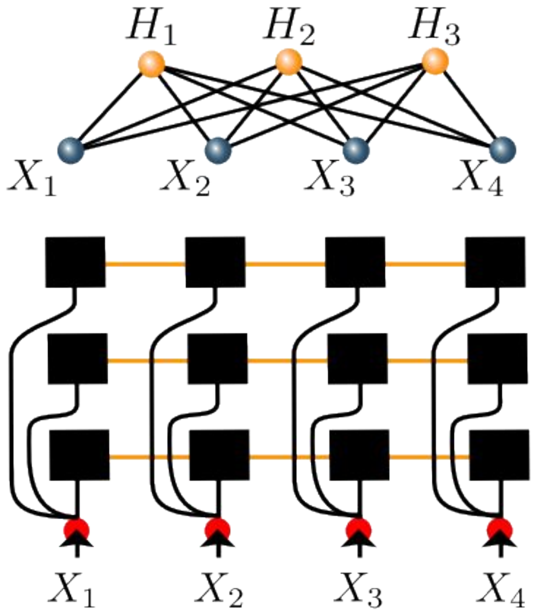
$$\tilde{H} = \sum_{ij} W_{ij} x_i h_j$$

Locality of the RBM

local connections

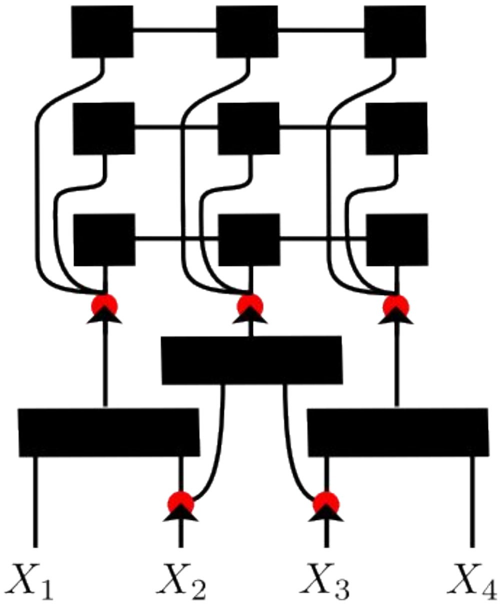


with local connections



IEEE Access 8, 68169-68182

Combining different models



classification

Image Classification

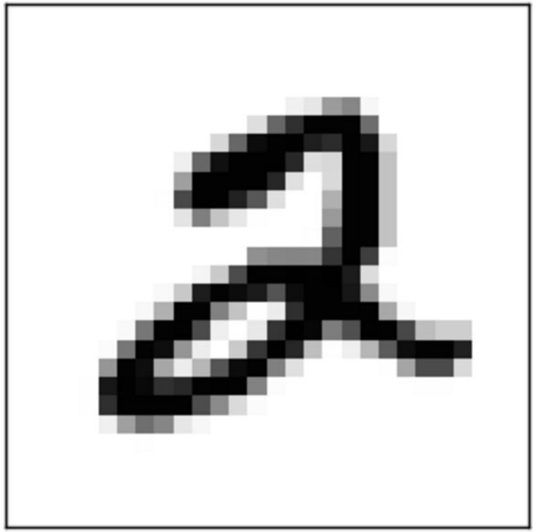
Goal: Given a dataset of images and corresponding labels, we want to predict the label of a new image

Image Classification

Goal: Given a dataset of images and corresponding labels, we want to predict the label of a new image

Choose a "model" : $p(x, y) = \mathsf{SBS}(x, y)$

Define a cost function : $-\sum_i \log(p(y_{x_i} | x_i))$

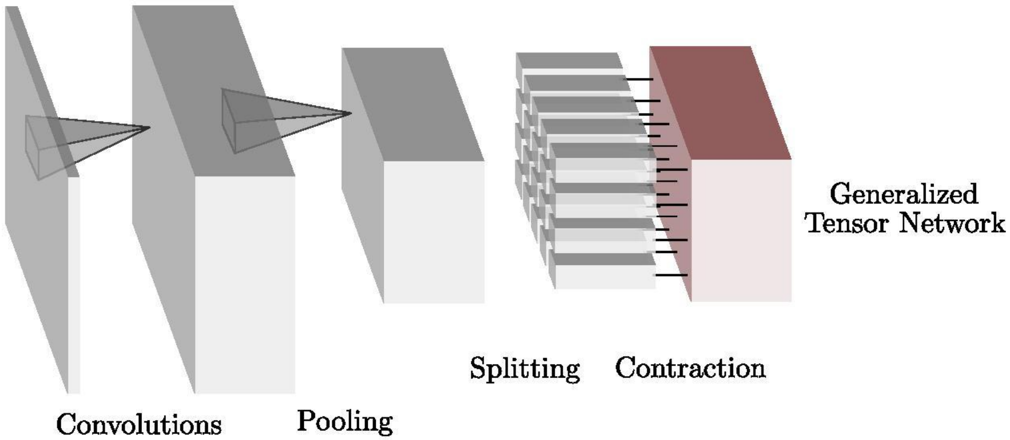


FashionMNIST

SVM	84.1%
Multilayer Perceptron	87.7%
SBS	89.0%
AlexNet	89.9%
1-layer CNN+SBS	92.3%
GoogLeNet	93.7%

IEEE Access 8, 68169-68182

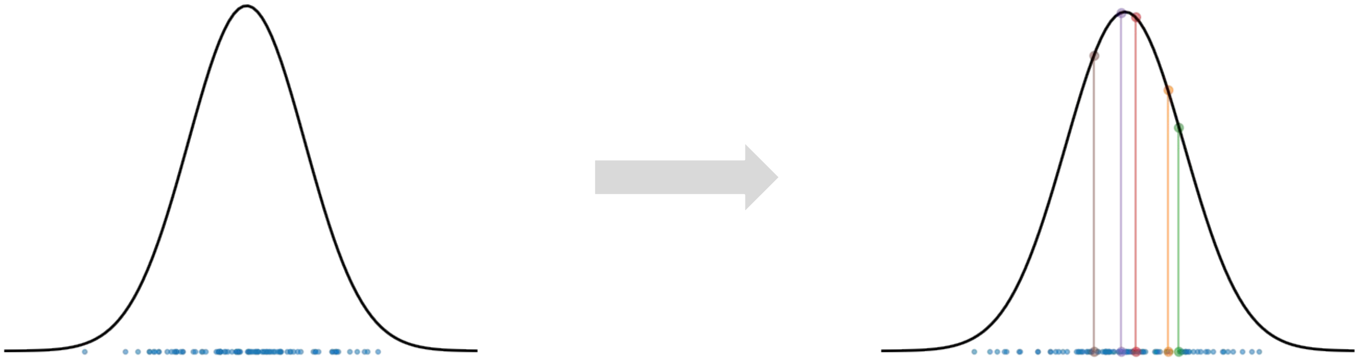
Architecture



IEEE Access 8, 68169-68182

maximum likelihood estimations

Maximum likelihood

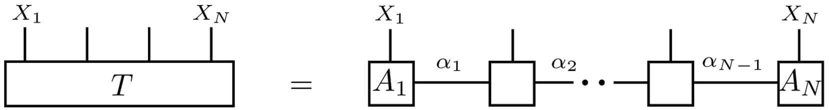


Learn from a database: $$\mathcal{L} = \sum_i \log P_i = \sum_i \log(T_i/Z)$$

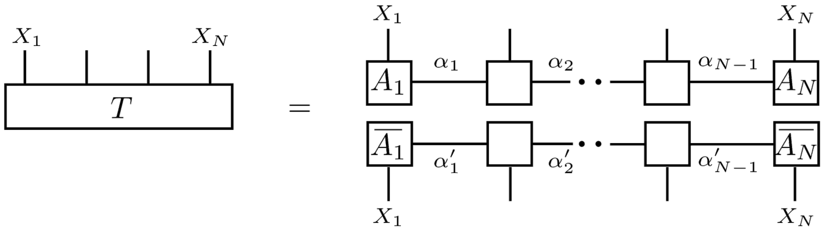
Learn from a distribution: $$\mathcal{D}(P \| T/Z) = \sum_x P_x \log\left(\frac{P_x}{T_x/Z}\right)$$

Some models for unsupervised learning

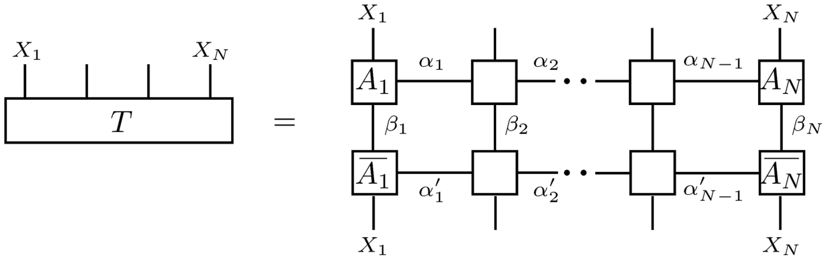
Tensor Train (MPS):



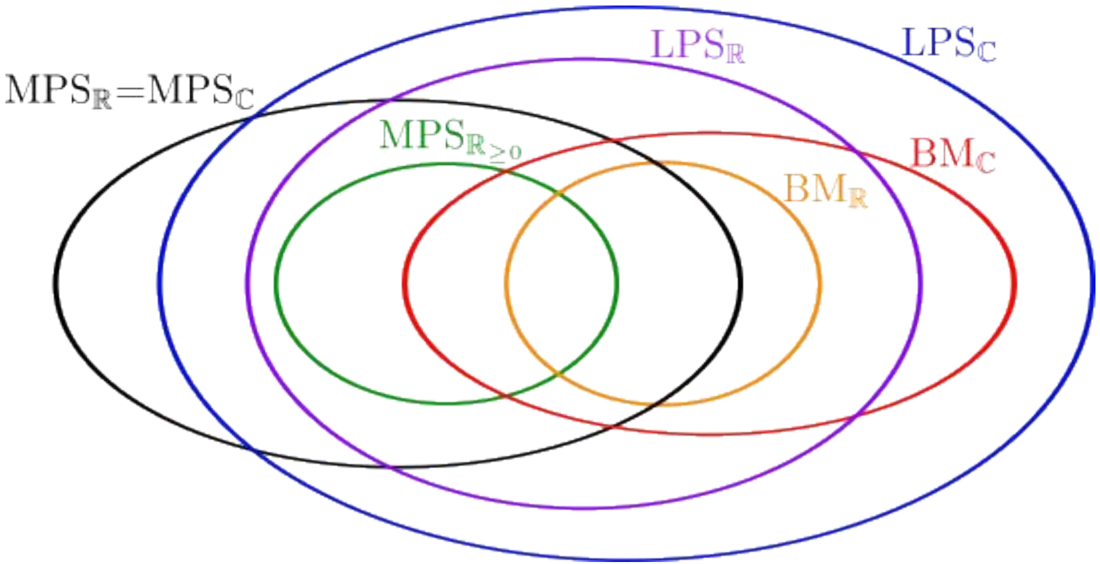
Born Machines:



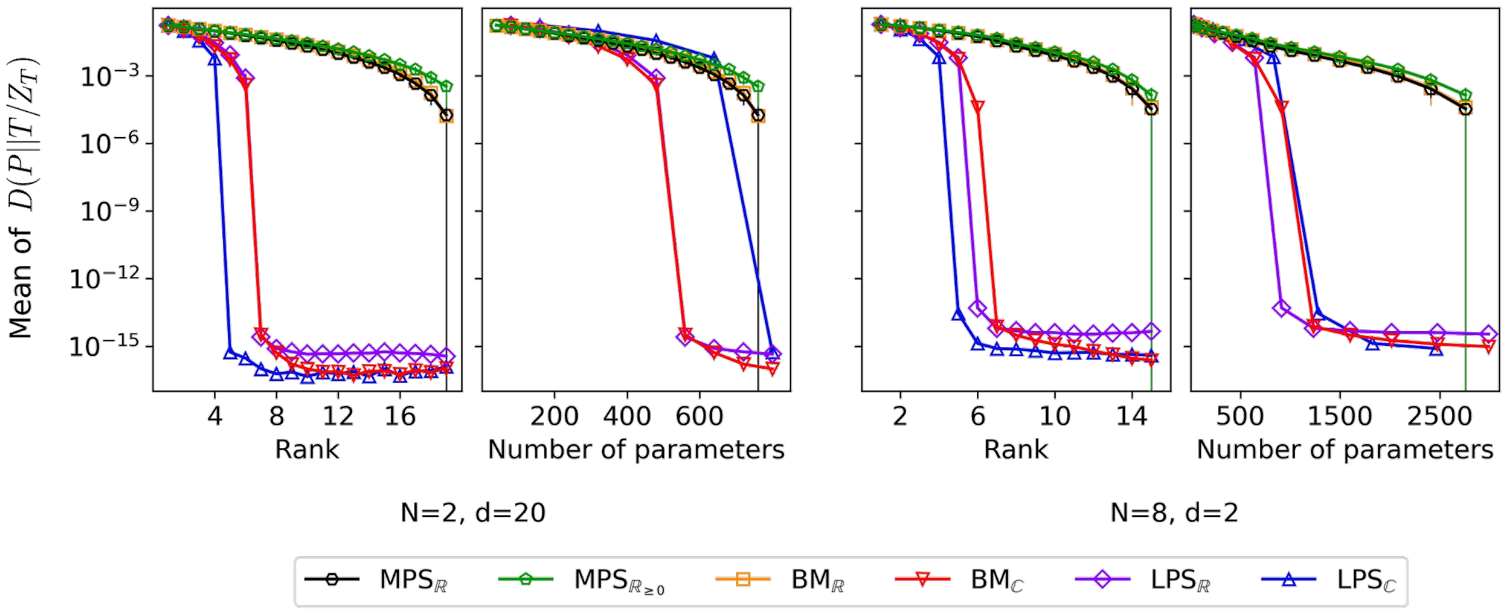
Locally Purified States:



Expressive power

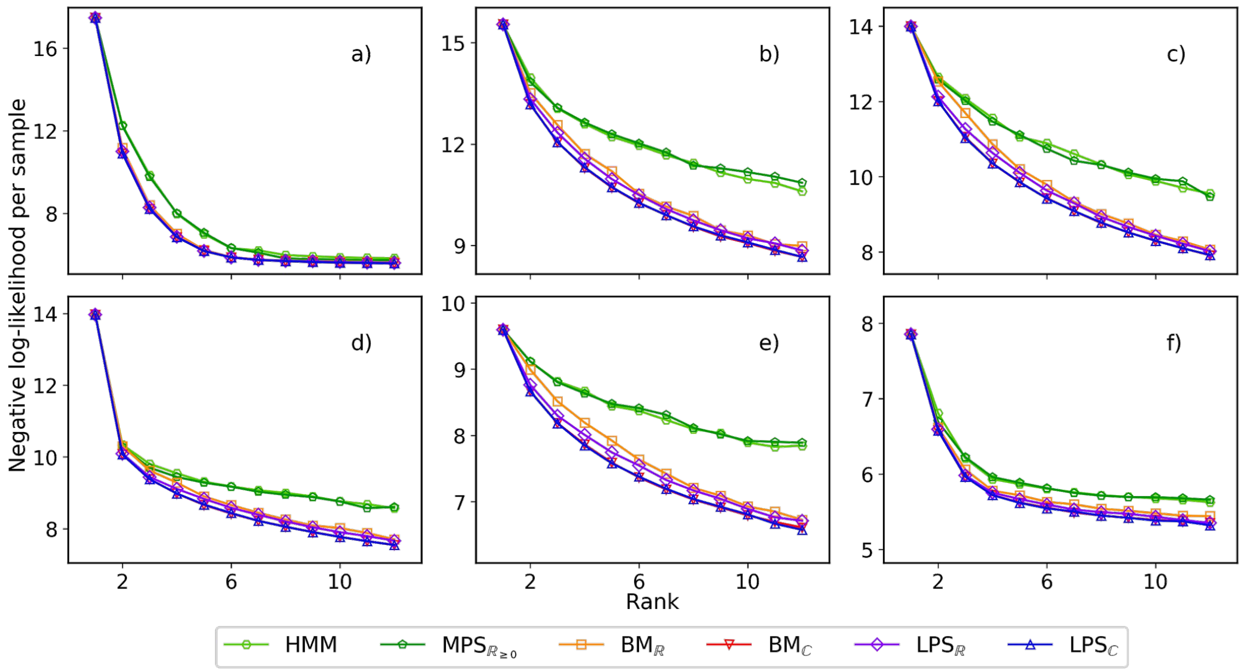


practical applications: random distributions



NeurIPS, 2019

practical applications: real data sets



NeurIPS, 2019

conclusions

- quantum mechanics and linear algebra
- linear algebra and tensor networks
- graphical models can be mapped to tensor networks
- tensor networks can be used for
    - classification problems
    - modeling probabilistic theories
- tensor networks can provide deeper mathematical insights

thanks for your attention