TENSOR NETWORKS & NEURAL NETWORK STATES: from chiral topological order to image classification

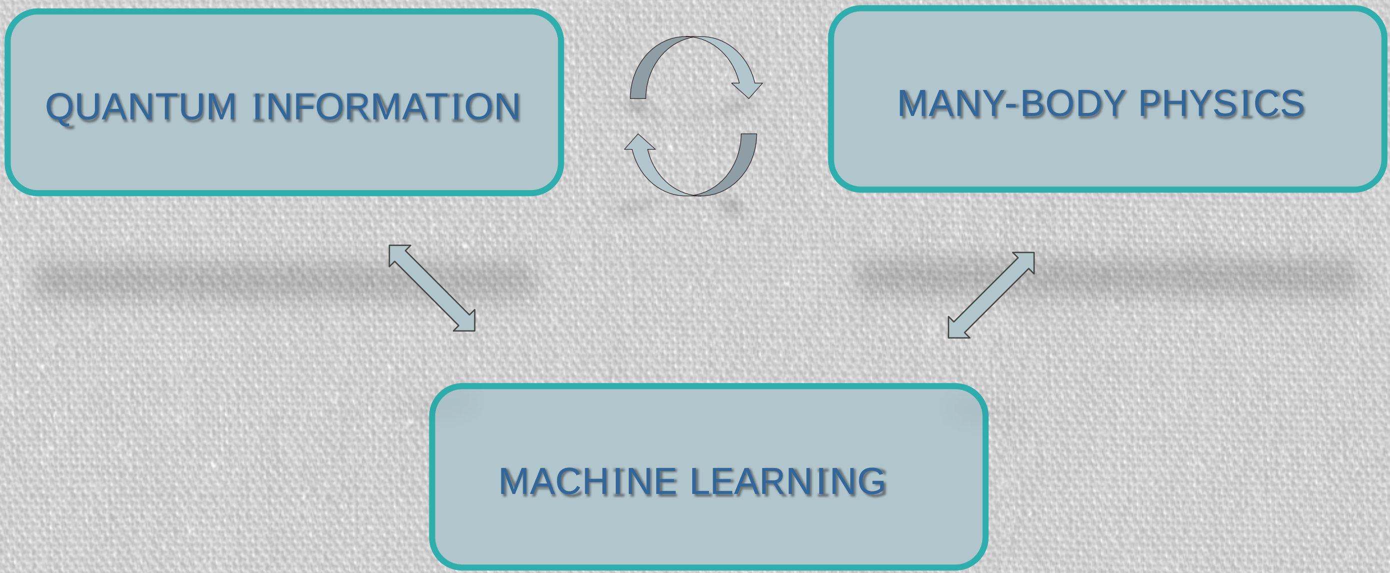
APS Workshop PHYSICS: WHAT IS NEXT?

Riverhead, NY, 8 October, 2018

Yvan Glasser
Nicola Pancotti
Mauriz August
Ivan Rodríguez

MPQ
Max-Planck-Institut für Quantenoptik

erc



Tensor networks and neural network states

- Applications in quantum many-body problems

  Glasser, Pancotti, Aubust, Rodríguez, JIC, Phys. Rev. X 8, 011006 (2018)

- Applications in machine learning

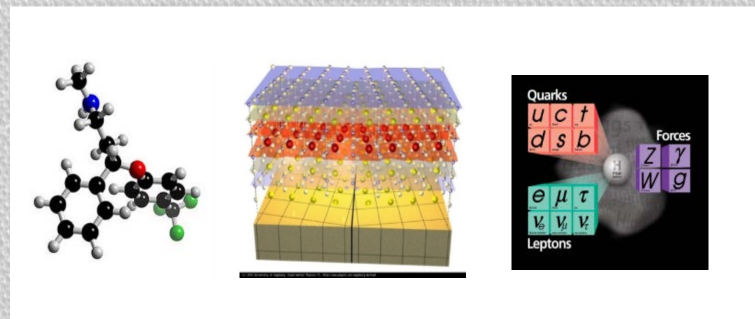
  Glasser, Pancotti, Cirac, arxiv:1806.05964

Applications in Quantum Many-Body Physics

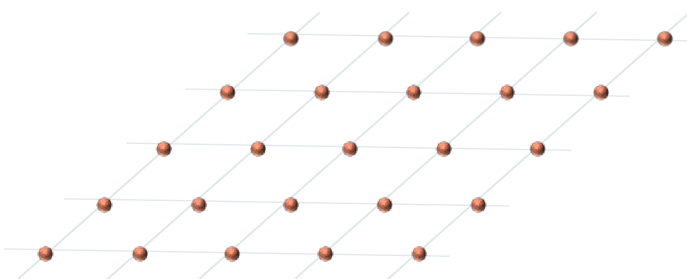
Glasser, Pancotti, August, Rodríguez, JIC, Phys. Rev. X 8, 011006 (2018)

see also Clark, arXiv:1710.03545

QUANTUM MANY-BODY SYSTEMS



Hamiltonian H



Physics at low temperatures

HILBERT SPACE

d^N

$$|\Psi\rangle = \sum \psi_{s_1,s_2,\ldots,s_N} \, | s_1, s_2, \ldots, s_N \rangle$$

Exponential problem

QUANTUM MANY-BODY SYSTEMS

Connection to machine learning:

$$|\Psi\rangle = \sum \psi(s)\,|\,s_1, s_2, \ldots, s_N\rangle$$

We can approximate

$$\psi(s) \Rightarrow \psi_w(s)$$

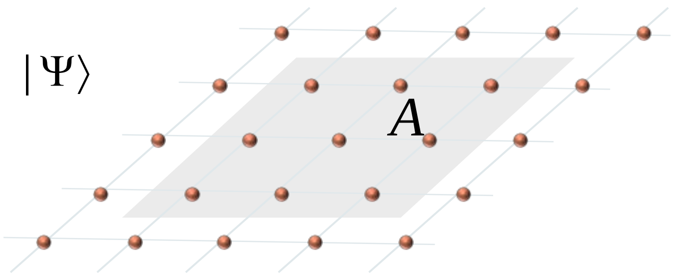
Family of states
w: parameters

Variational method:

$$\min_w \frac{\langle \Psi_w \,|\, H \,|\, \Psi_w \rangle}{\langle \Psi_w \,|\, \Psi_w \rangle}$$

How do we choose $\psi(s)$?

ENTANGLEMENT AND AREA LAW



How much entangled?

Lattice in any physical dimension and geometry

Local Hamiltonians: $H = \sum_n h_n$

Thermal equilibrium: $T = 0$

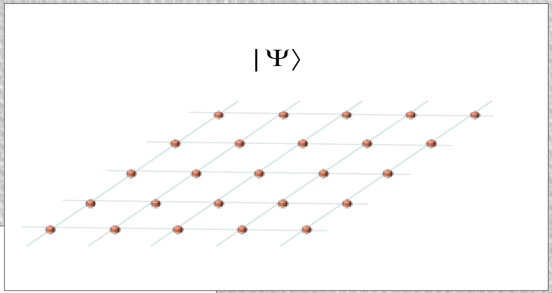
Very little!

$E(\Psi) \sim |\partial A|$

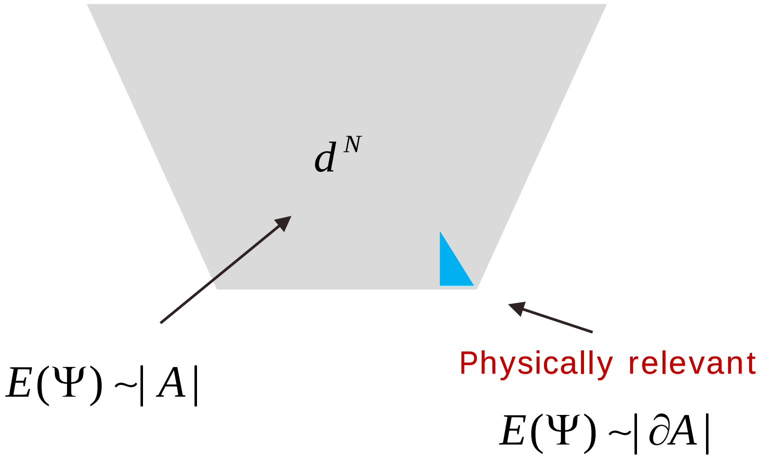
Area law

ENTANGLEMENT AND AREA LAW

What do we learn?



EXPONENTIAL HILBERT SPACE



TENSOR NETWORK STATES

$$|\Psi\rangle = \sum \psi_{s_1,s_2,\ldots,s_N} \,|\, s_1, s_2, \ldots, s_N \rangle$$

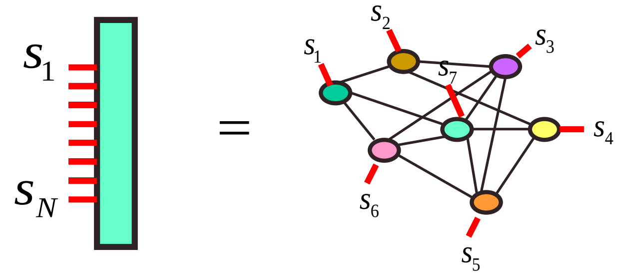
Local
Hamiltonian

Very little entanglement

$$\psi_{s_1,s_2,s_N} \approx \sum_{\alpha,\beta,\ldots} A^{s_1}_{\alpha,\beta,\gamma,\delta} B^{s_2}_{\gamma,\varepsilon,\eta,\mu} C^{s_3}_{\eta,\sigma,\tau,\zeta} \cdots$$

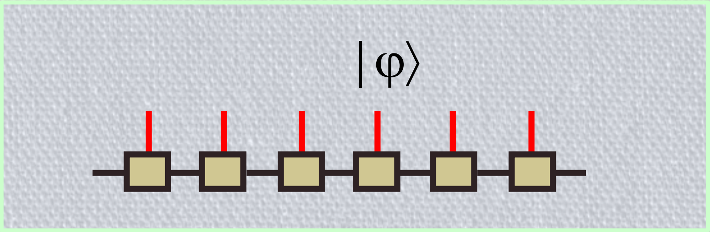
Tensor Networks

$\psi_{s_1,s_2,\ldots,s_N}$



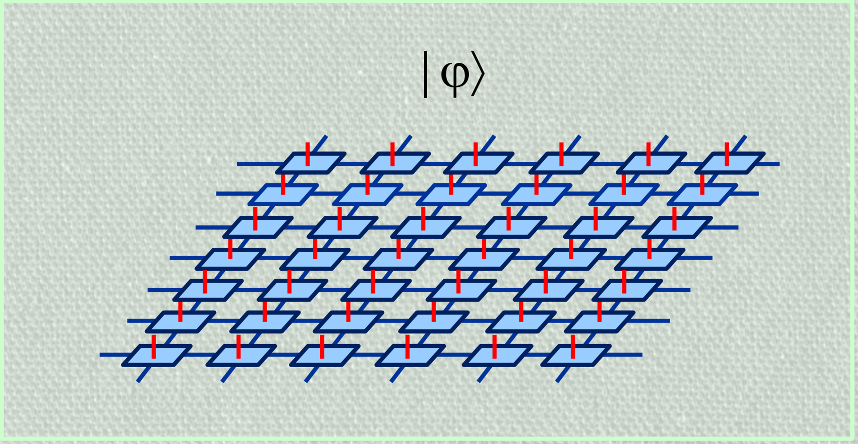
TENSOR NETWORK STATES:

MPS



Fannes, Narchtergaele, Werner, CMP **144**, 443 (1992)

PEPS



Verstraete, JIC, arxiv:0407066

Parameters w are the tensor themselves

TENSOR NETWORK STATES:

Variational method:

$$\min_{w} \frac{\langle \Psi_w | H | \Psi_w \rangle}{\langle \Psi_w | \Psi_w \rangle}$$

Requires contracting tensors

1 Dimension D^3

2 Dimensions D^{10}

3 Dimensions D^{18}

Monte-Carlo methods:

$$E_w = \frac{\langle \psi_w | H | \psi_w \rangle}{\langle \psi_w | \psi_w \rangle} = \sum_{\mathrm{s}} p(\mathrm{s}) E(\mathrm{s})$$

$$p(\mathrm{s}) = \frac{|\psi(\mathrm{s})|^2}{\sum_{\mathrm{s}} |\psi(\mathrm{s})|^2}$$

TENSOR NETWORKS

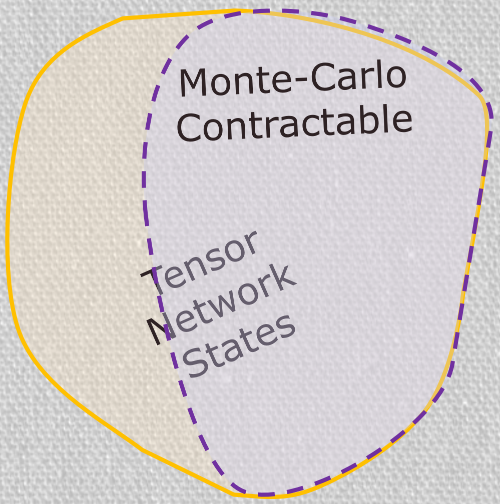
Schuch, Wolf, Verstraete, JIC, PRL **100**, 040501 (2008)

$$E_w = \sum_{s} p(s) E(s)$$

$$p(s) = \frac{|\psi(s)|^2}{\sum_s |\psi(s)|^2}$$

- Consider TN which can be contracted using MC methods

  $\Psi(s)$ and $E(s)$ can be efficiently computed



TENSOR NETWORKS

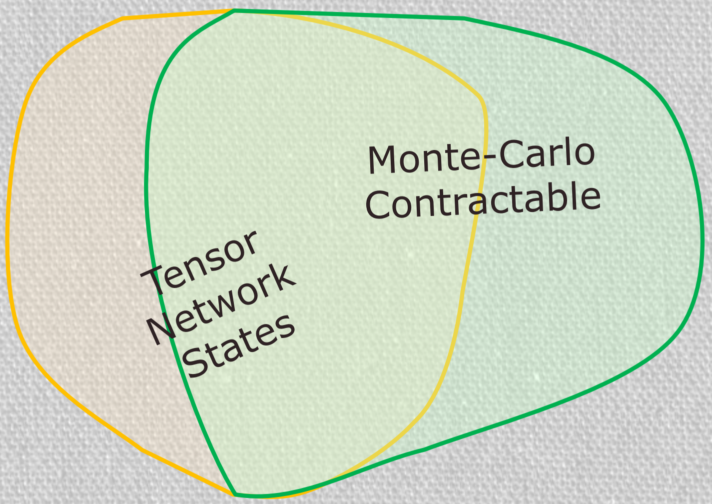
Schuch, Wolf, Verstraete, JIC, PRL **100**, 040501 (2008)

$$E_w = \sum_s p(s) E(s)$$

$$p(s) = \frac{|\psi(s)|^2}{\sum_s |\psi(s)|^2}$$

- Consider TN which can be contracted using MC methods

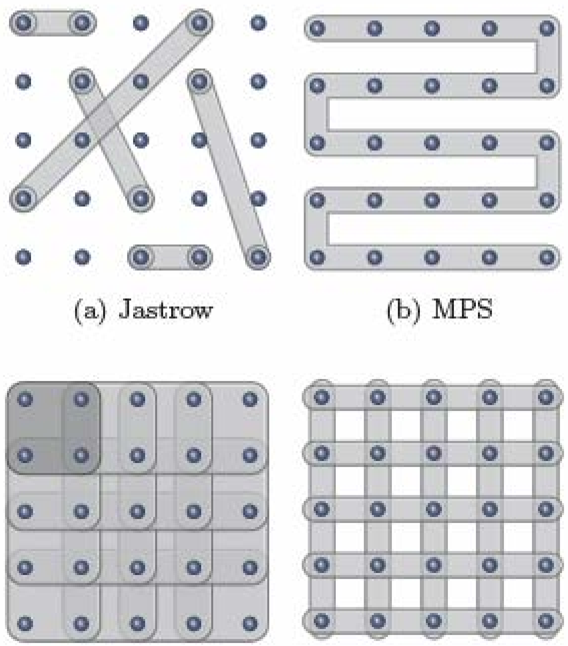
  $\Psi(s)$ and $E(s)$ can be efficiently computed



- Extend the family of TN
- Minimize the energy using Variational Monte-Carlo

STRING-NET STATES

$$|\psi\rangle = \sum_{s_1,\dots,s_N} \psi_w(\mathrm{s})\,|s_1,\dots,s_N\rangle$$



(a) Jastrow (b) MPS

(c) EPS (d) SBS

Jastrow

$$\psi_w(\mathrm{s}) = \prod_{i<j} f_{i,j}(s_i, s_j)$$

MPS

$$\psi_w(\mathrm{s}) = \mathrm{Tr}\left[\prod_j A_j^{s_j}\right]$$

Sandvik, Vidal, PRL **99**, 220602 (2007)
Schuch, Wolf, Verstraete, JIC, PRL **100**, 040501 (2008)
Silora, Chang, Chou, Pollmann, Kao, PRB **91**, 165113 (2015)

String-Bond States

$$\psi_w(\mathrm{s}) = \prod_I \mathrm{Tr}\left[\prod_{j\in I} A_{I,j}^{s_j}\right]$$

Schuch, Wolf, Verstraete, JIC, PRL **100**, 040501 (2008)

Entangled-plaquette states

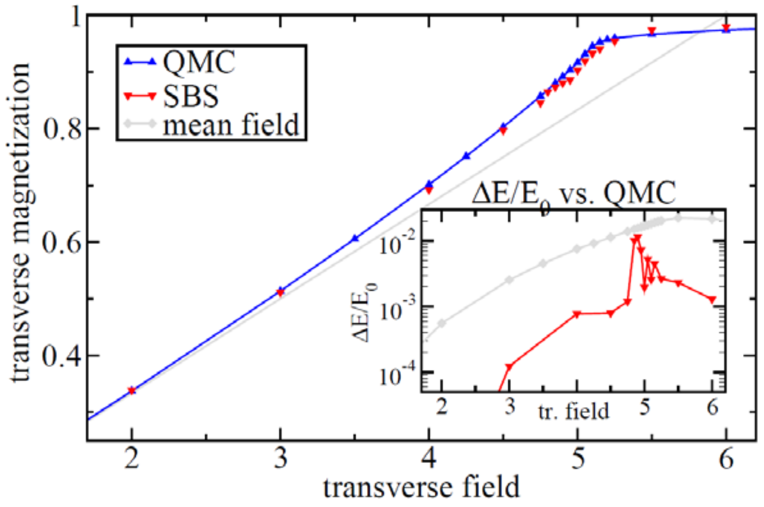
$$\psi_w(\mathrm{s}) = \left[\prod_p C_p^{s_p}\right]$$

Mezzacapo, Schuch, Boninsegni, JIC, NJP **11**, 083026 (2009)

STRING-BOND STATES

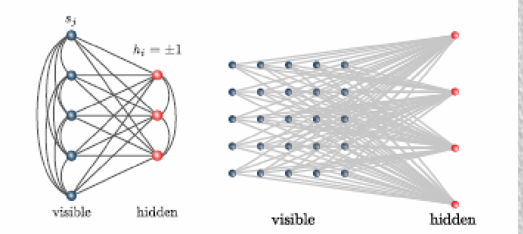
Ising Model in 3D with a transverse magnetic field

8x8x8 lattice



Sfondrini, Cerrillo, Schuch, JIC, Phys. Rev. B 81, 214426 (2010)

BOLTZMANN MACHINES



(a) Deep Boltzmann Machine

(b) Restricted Boltzmann Machine in 2D

RESTRICTED BOLTMANN MACHINE (RBM)

$$\psi_w(\mathrm{s}) \prec \sum_{\mathrm{h}} e^{H(\mathrm{s},\mathrm{h})}$$

$$H = \sum_j a_j s_j + \sum_i b_i h_i + \sum_{i,j} c_{i,j} s_i s_j + \sum_{i,j} w_{i,j} s_i h_j$$

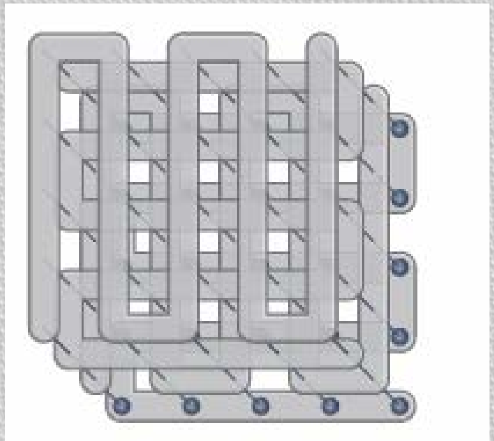
$$\psi_w(s) = e^{\sum a_j s_j} \prod_i \cosh\left[b_i + \sum_j w_{i,j} s_i \right]$$

Ackley, Hinton, Sejnowski, Cognitive Science **9**, 147 (1985)
Carleo, Troyer, Science 355, 602 (2017)

RESTRICTED BOLTZMANN MACHINES

$$\psi_w(s) = e^{\sum a_j s_j} \prod_i \cosh\left[b_i + \sum_j w_{i,j} s_i \right]$$
$$= \prod_i \mathrm{Tr}\left[A_{i,j}^{s_j} \right]$$

$$A_{i,j}^{s_j} = \begin{pmatrix} e^{b_i + w_{i,j} s_j} & 0 \\ 0 & e^{-b_i - w_{i,j} s_j} \end{pmatrix}$$



RBM is contained in the class of SBS

EXAMPLES

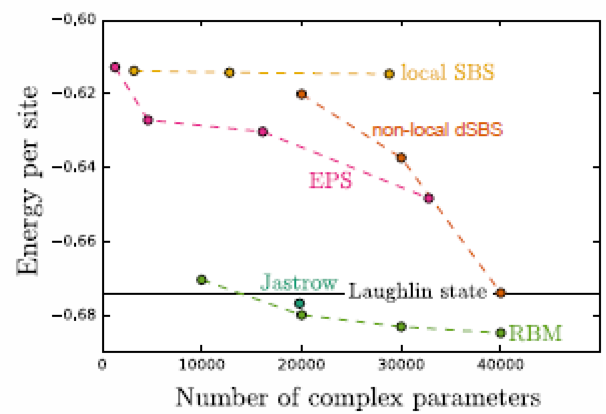
SPIN CHAINS

Majumdar-Gosh

AKLT

CHIRAL TOPLOGICAL STATES

$$H = J\sum_{\langle i,j\rangle}^{N-1} \mathrm{S}_i\,\mathrm{S}_j + J_\chi \sum_{\langle i,j,k\rangle}^{N} \mathrm{S}_i\left(\mathrm{S}_j \times \mathrm{S}_k\right)$$



(b) 10x10 lattice

See also Nomura, Barmawan, Yamji, Imada, arXiv:1709.06475

See also Kaubruegger, Pastori, Budich, arXiv:1710.04713

Applications in Machine learing

Glasser, Pancotti, Cirac, arxiv:1806.05964

SUPERVISED MACHINE LEARNING

IMAGE CLASSIFICATION

SUPERVISED MACHINE LEARNING

PROBABILITY

$$p(x, y) = p_w(x, y)$$

Adjustable parameters

CONDITIONAL PROBABILITY

$$p(y / x) = \frac{p(x, y)}{\sum_z p(x, z)}$$

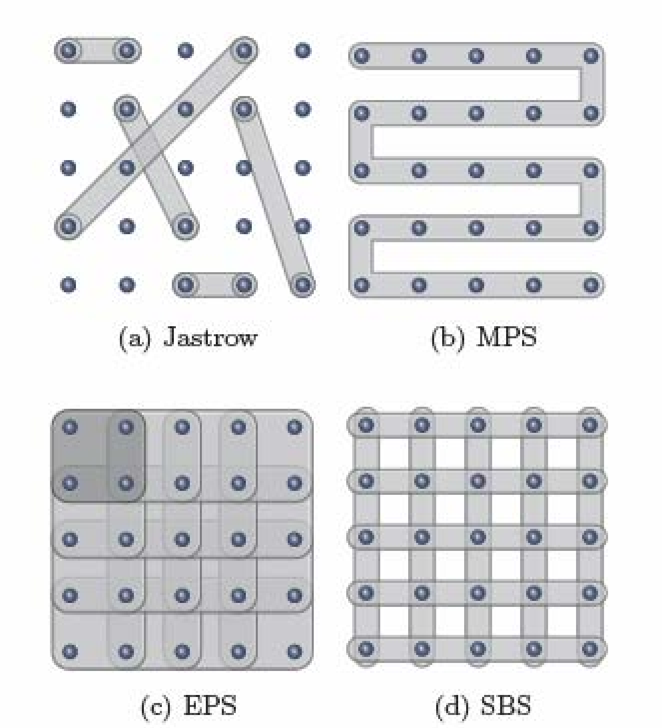
Use maximal likelyhood
No need to compute normalization
No need to use Monte-Carlo

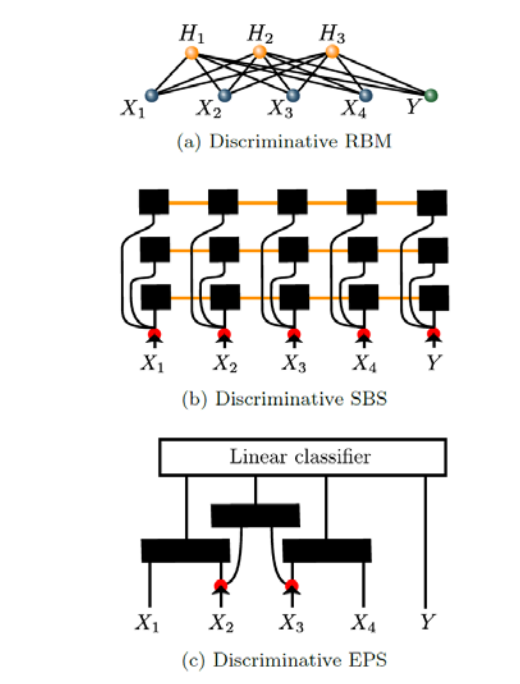
TRAINING: COST FUNCTION

$$F = -\sum_{i=1}^{D} \log[p(y_i / x_i)]$$

GENERALIZED TENSOR NETWOKS

$$p(x, y) = e^{GTN(x,y)}$$



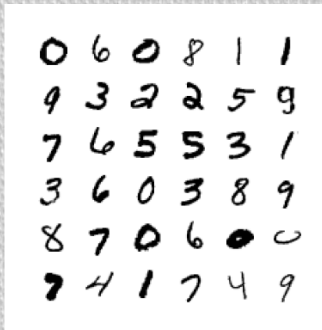


See also:

Critch and Morton, Symm. Integrability Geom. Methods Appl. (2014)
Robeva and Seigal, arXiv:1710.01437 (2017)
Cohen, Sharir, Shashua, 33rd int. Conf. Machine Learning (2016)
Levine, Sharir, Cohen, Shashua, arXiv:1803.09780 (2018)
Stoudernmire, Schwab, Adv. Neural Info. Pro. Sys. 29, 4799 (2016)
Liu, Ran, Wittek, Peng, arcia, Su, Lewenstein, arXiv:1710.04833
Liu, Zhang, Lewenstein, Ran, arXiv:1803.0911

SUPERVISED LEARNING WITH GTN

MNIST dataset



FASHION MNIST

28x28 greyscale images

Validation set: 55000
Test set: 10000

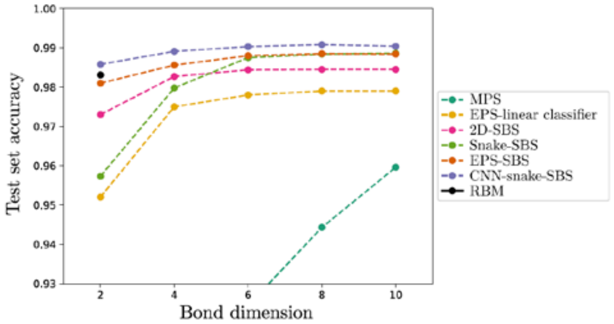


FIG. 14. Test set accuracy of different generalized tensor networks on the MNIST dataset.

Method	Accuracy
Support Vector Machine	84.1%
EPS + linear classifier	86.3%
Multilayer perceptron	87.7%
EPS-SBS	88.6%
Snake-SBS	89.2%
AlexNet	89.9%
CNN-snake-SBS	92.3%
GoogLeNet	93.7%

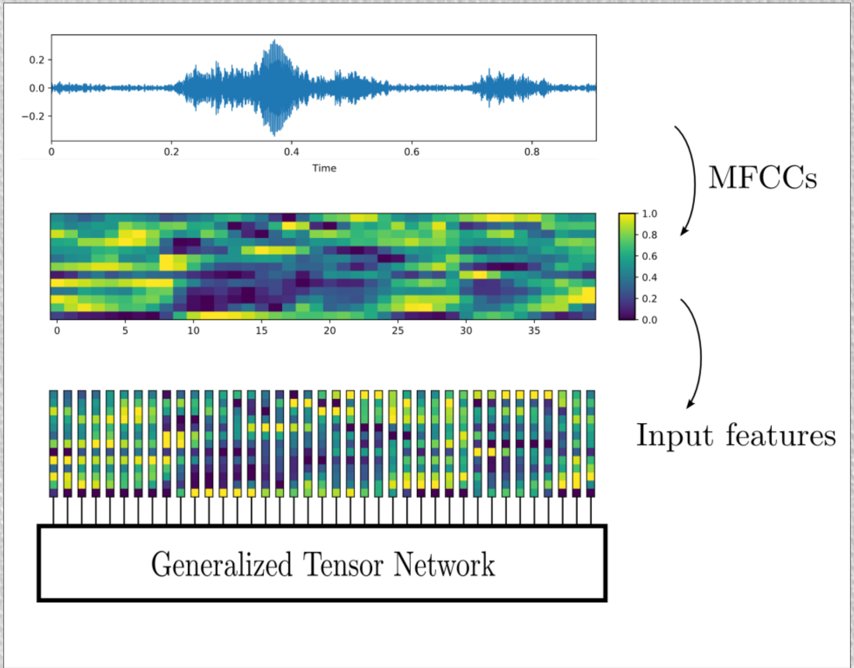
LeCun, Cortes, Burges, http://yann.lecun.com/exdb/minst (1998)

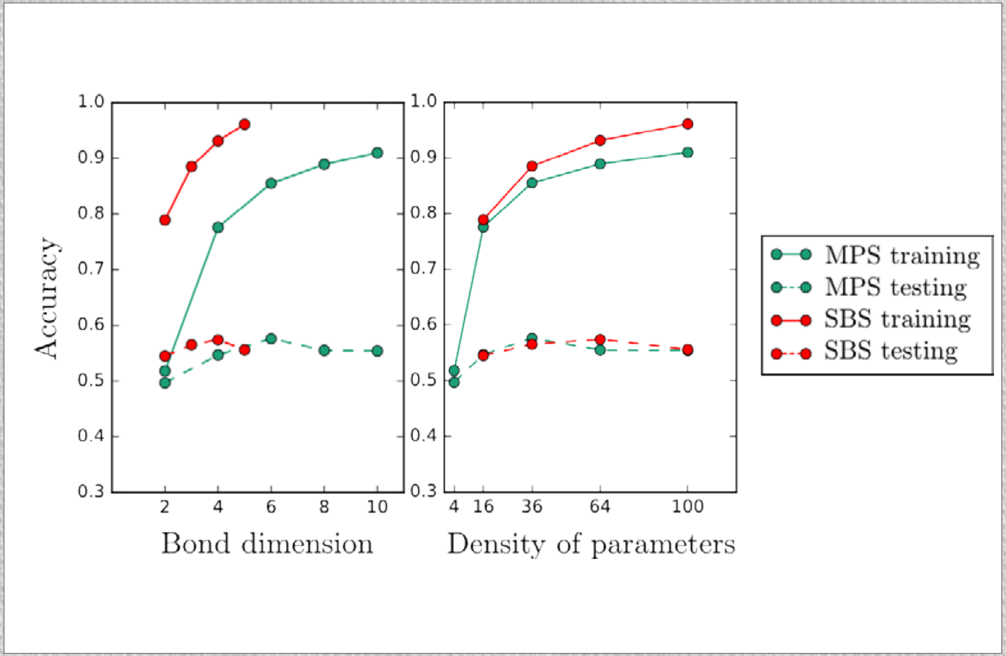
Xiao, Rasul, Vollgraf, arXiv:1708.07747 (2017)

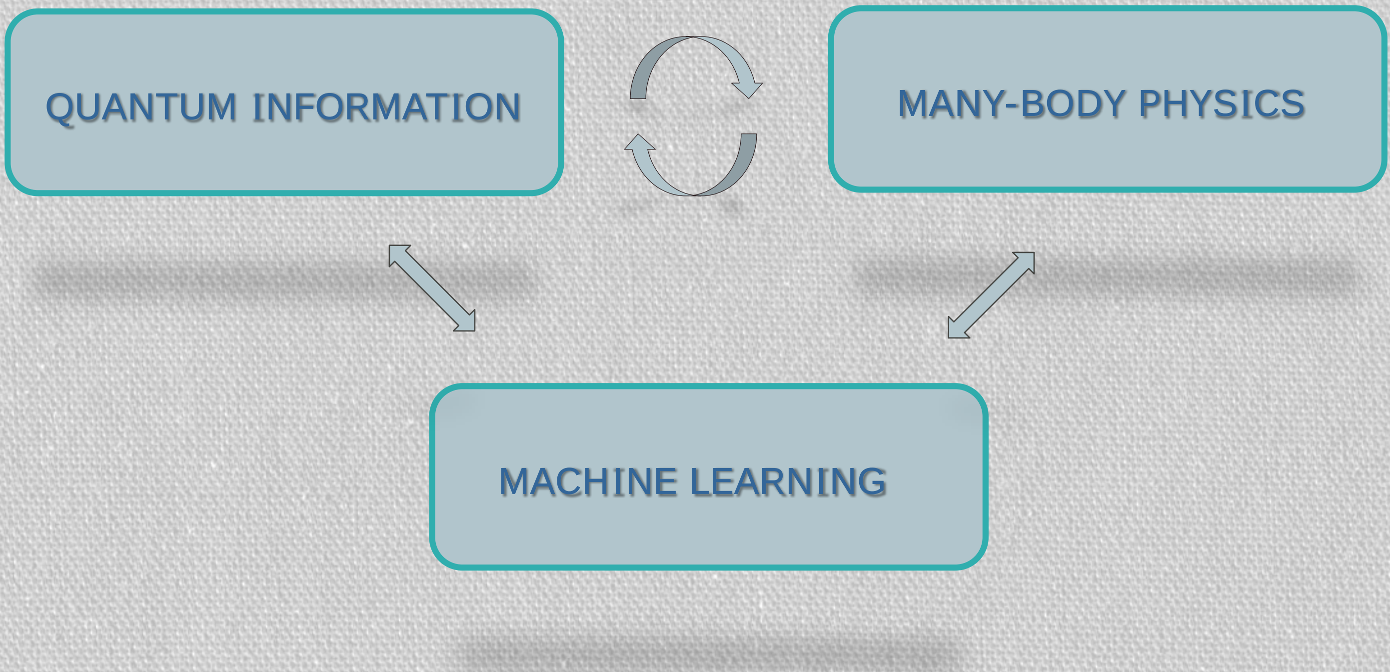
SUPERVISED LEARNING WITH GTN

URBANSOUND8K

10-class classification of urban sounds : air conditioner, car horn, children playing, dog bark, drilling, engine idling, gun shot, jackhammer, siren, street music







OUTLOOK

- Analysis and benchmark in other models
- Application to Lattice Gauge Theories
- Better understanding of the family of states

Acknowledgments