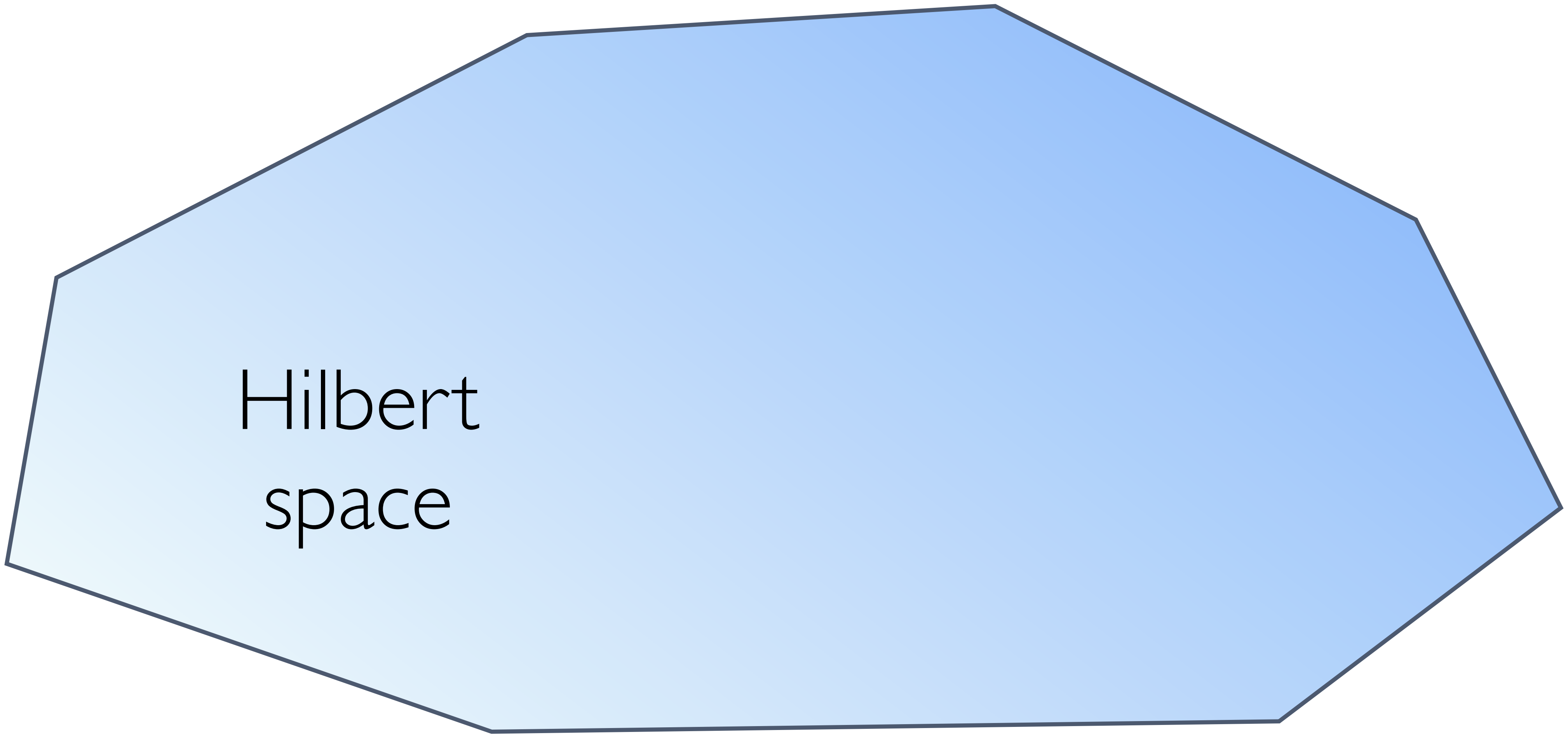
TENSOR NETWORK MACHINE LEARNING

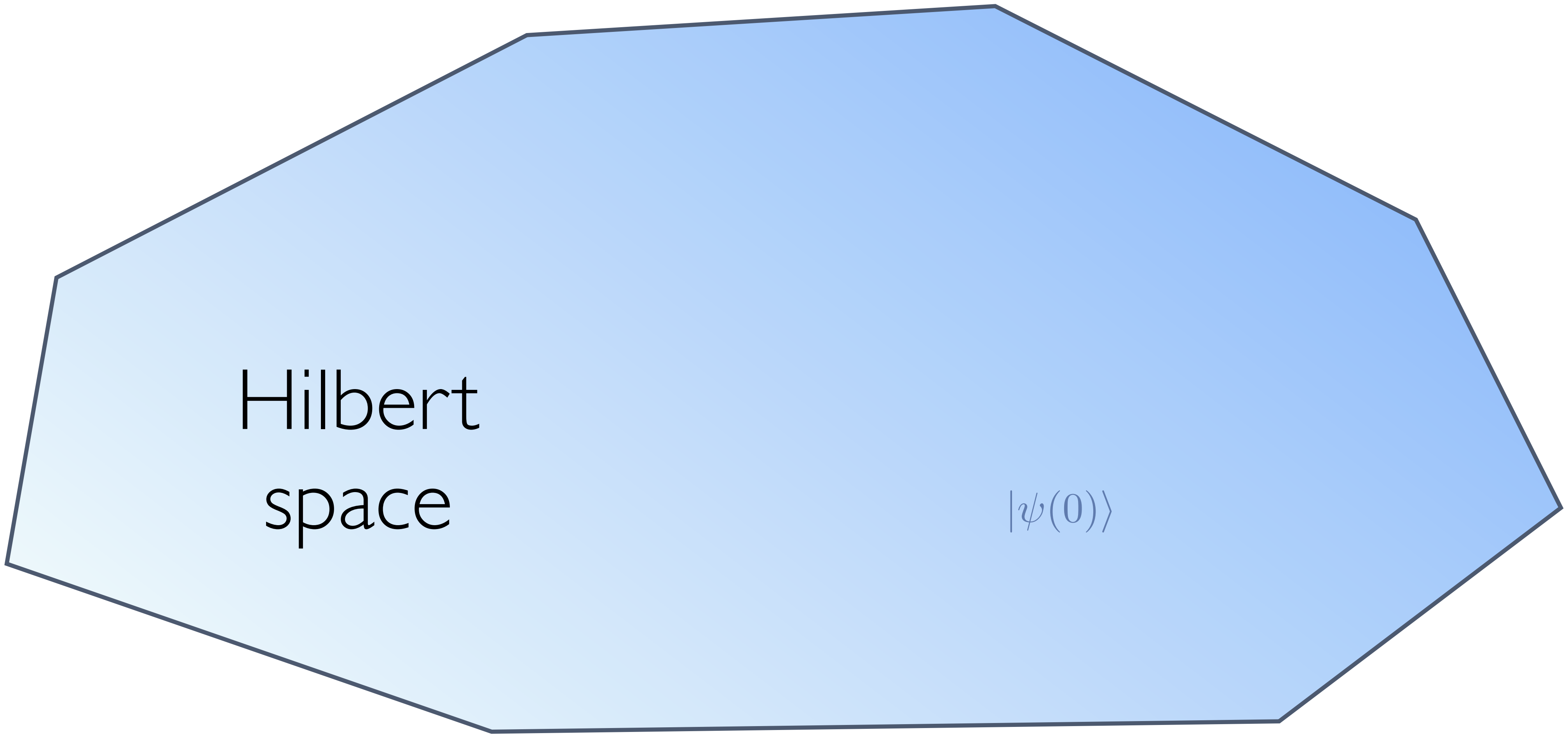
Simone Montangero

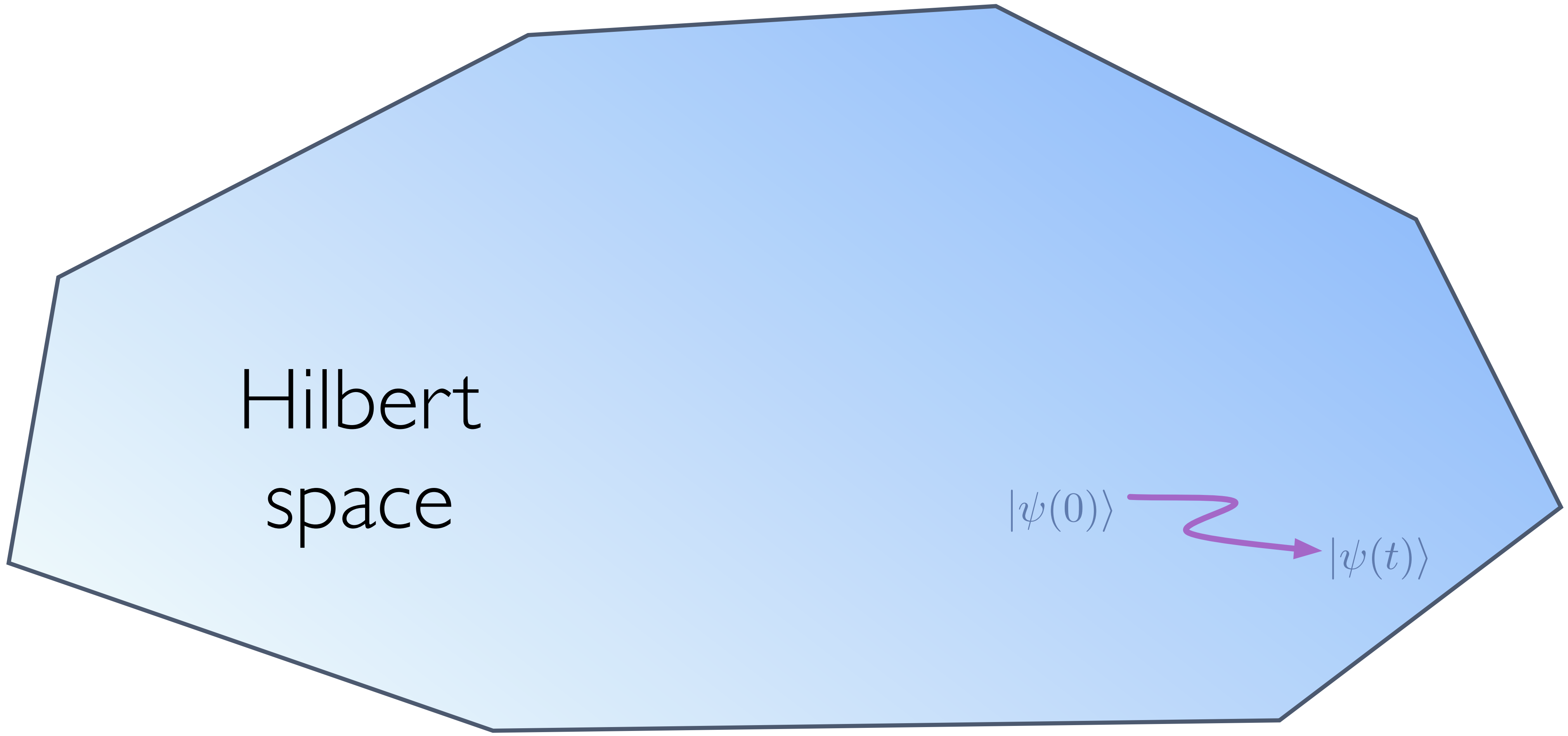
University of Padova

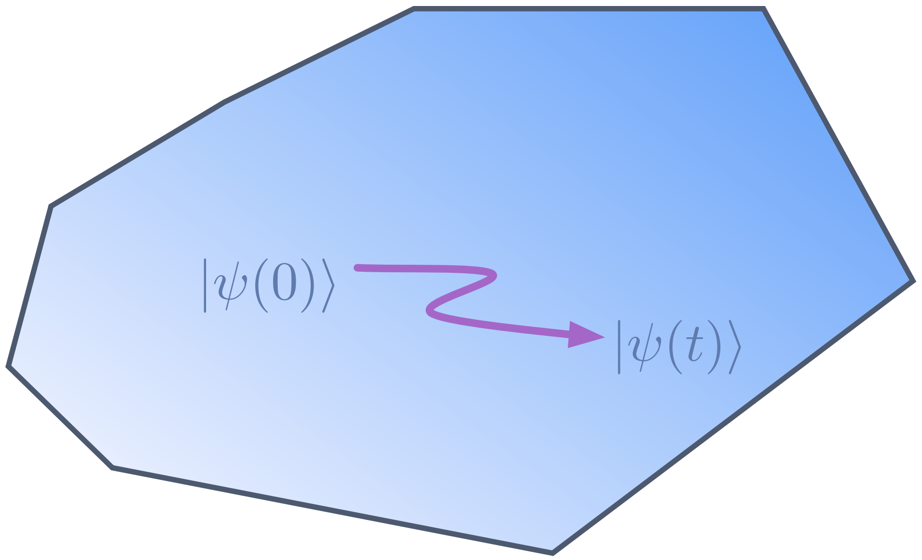
Dipartimento di Fisica e Astronomia Galileo Galilei

Università degli Studi di Padova


Hilbert
space


Hilbert space
$|\psi(0)\rangle$


Hilbert
space
$|\psi(0)\rangle$
$|\psi(t)\rangle$



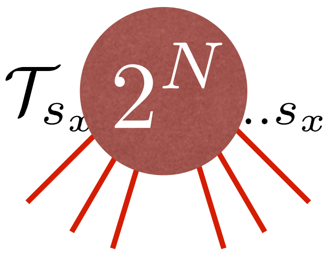
$|\psi(0)\rangle$
$|\psi(t)\rangle$

TENSOR NETWORK ANSATZ

$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1}, s_{x_2} \ldots s_{x_N}} \mathcal{T}_{s_{x_1}, s_{x_2} \ldots s_{x_N}} |s_{x_1}, s_{x_2} \ldots s_{x_N}\rangle$$

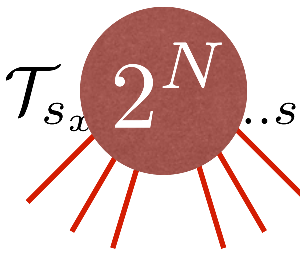
TENSOR NETWORK ANSATZ

$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1}, s_{x_2} \ldots s_{x_N}} \mathcal{T}_{s_{x_1}, s_{x_2} \ldots s_{x_N}} |s_{x_1}, s_{x_2} \ldots s_{x_N}\rangle$$



TENSOR NETWORK ANSATZ

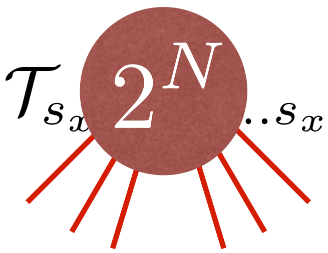
$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1}, s_{x_2} \dots s_{x_N}} \mathcal{T}_{s_{x_1}, s_{x_2} \dots s_{x_N}} |s_{x_1}, s_{x_2} \dots s_{x_N}\rangle$$

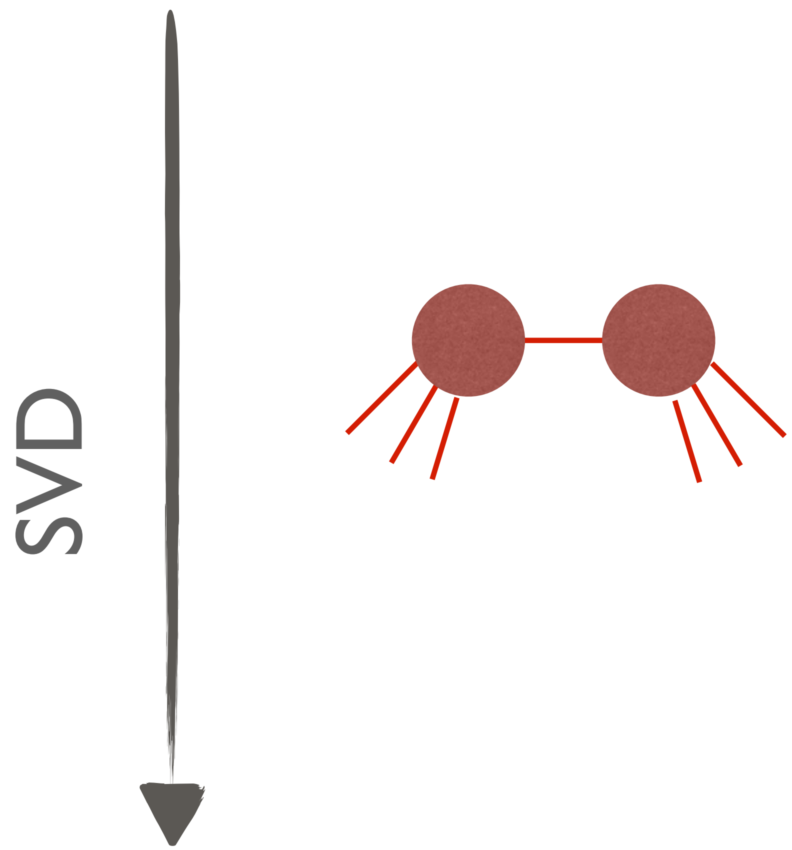


SVD

TENSOR NETWORK ANSATZ

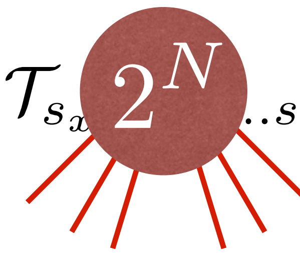
$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1},s_{x_2}\dots s_{x_N}} \mathcal{T}_{s_{x_1}\dots s_{x_N}} |s_{x_1}, s_{x_2} \dots s_{x_N}\rangle$$

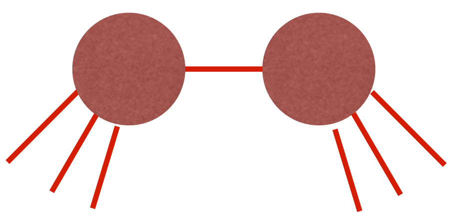


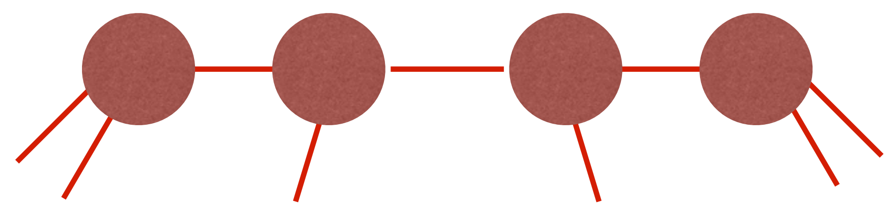


TENSOR NETWORK ANSATZ

$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1}, s_{x_2} \ldots s_{x_N}} \mathcal{T}_{s_{x_1} \ldots s_{x_N}} |s_{x_1}, s_{x_2} \ldots s_{x_N}\rangle$$

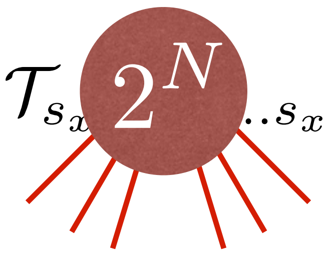


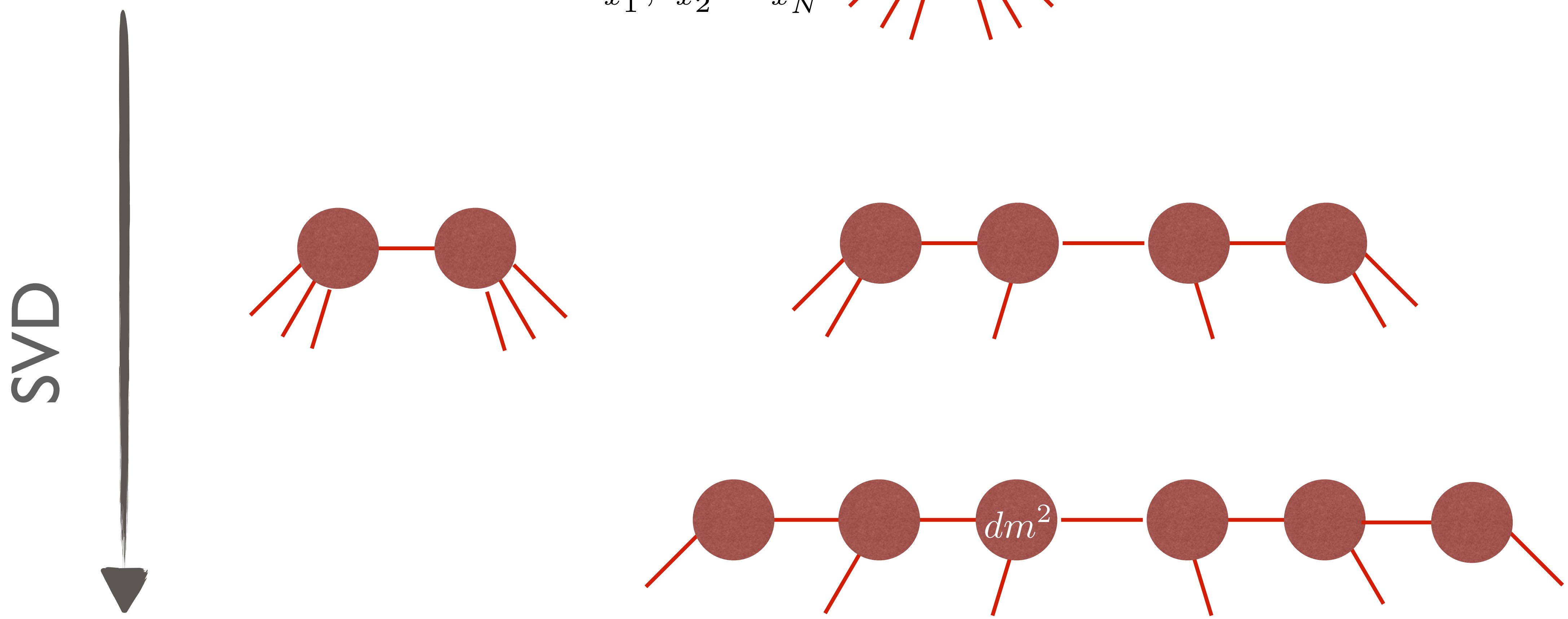




TENSOR NETWORK ANSATZ

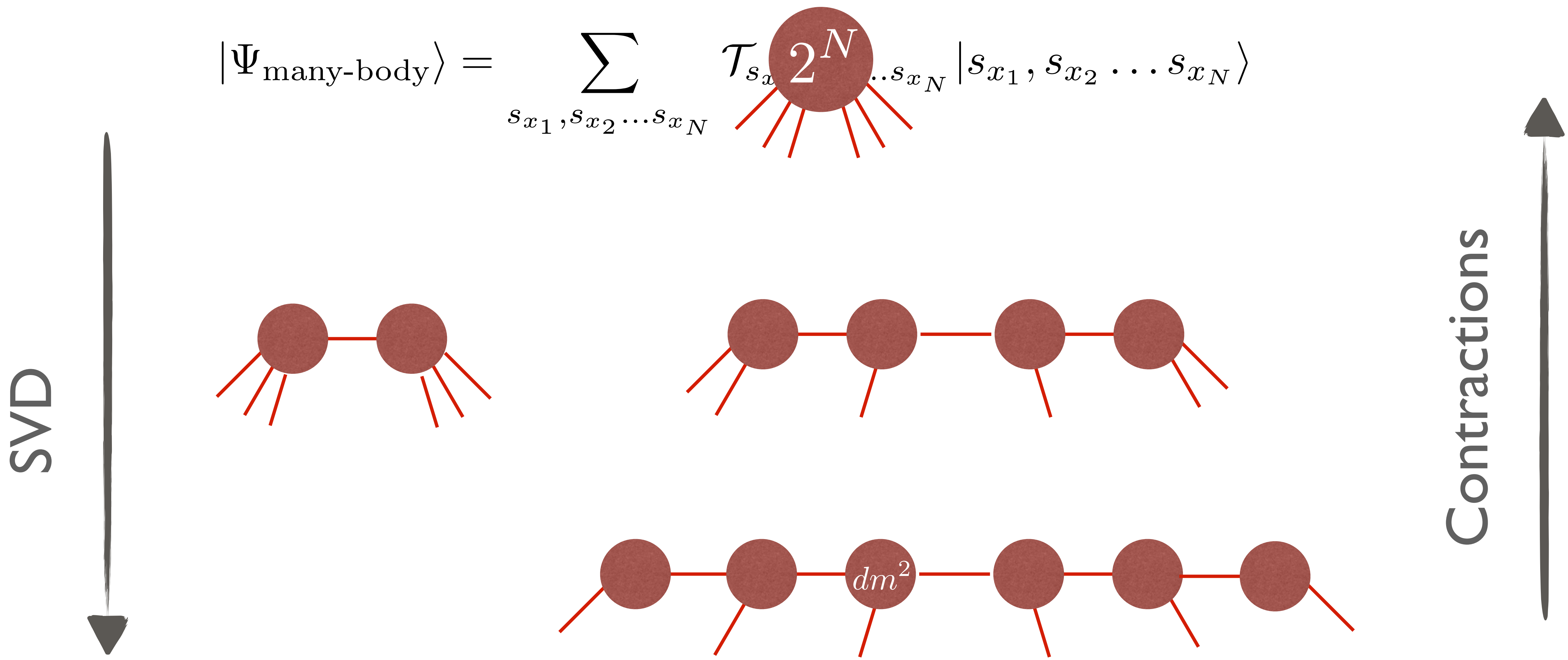
$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1},s_{x_2}\dots s_{x_N}} \mathcal{T}_{s_{x_1}\dots s_{x_N}} |s_{x_1}, s_{x_2}\dots s_{x_N}\rangle$$





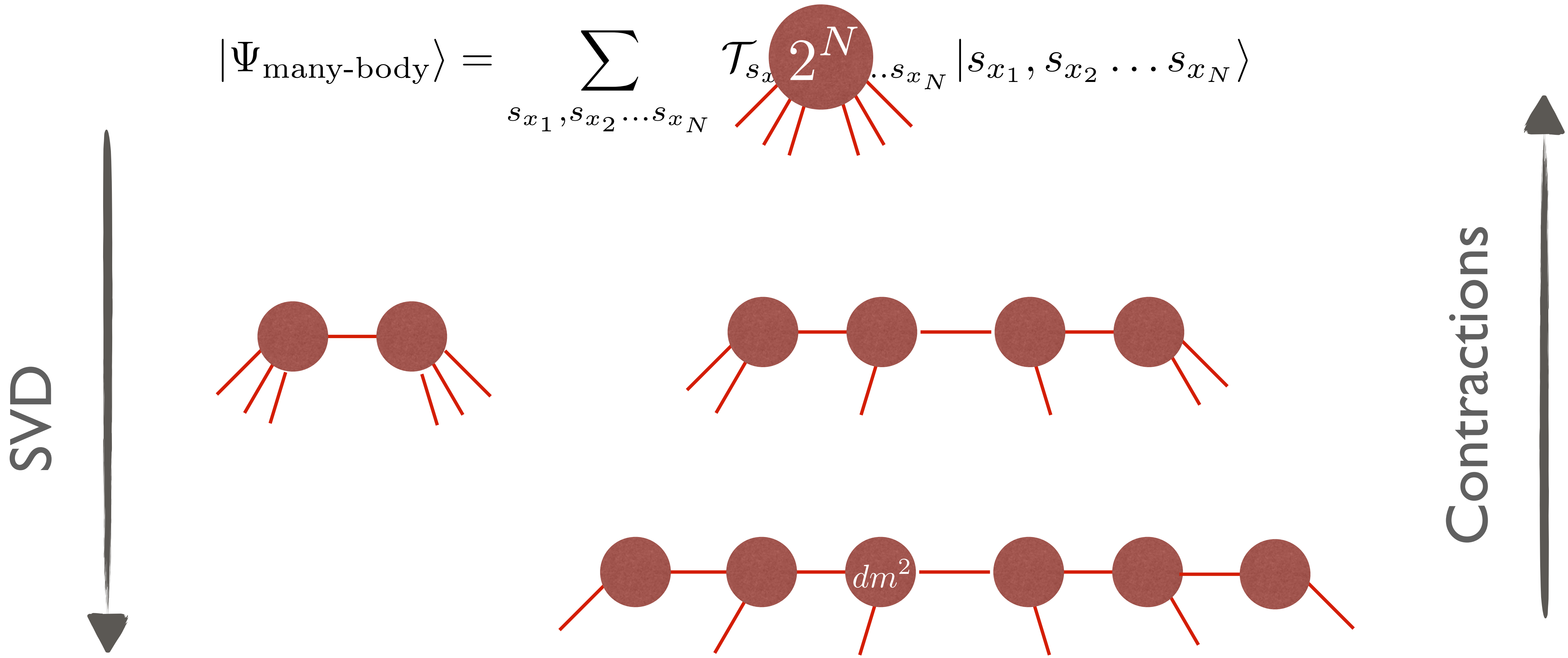
TENSOR NETWORK ANSATZ

$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1},s_{x_2}\dots s_{x_N}} \mathcal{T}_{s_{x_1}\dots s_{x_N}} |s_{x_1}, s_{x_2}\dots s_{x_N}\rangle$$



TENSOR NETWORK ANSATZ

$$|\Psi_{\text{many-body}}\rangle = \sum_{s_{x_1}, s_{x_2} \dots s_{x_N}} \mathcal{T}_{s_{x_1} \dots s_{x_N}} |s_{x_1}, s_{x_2} \dots s_{x_N}\rangle$$

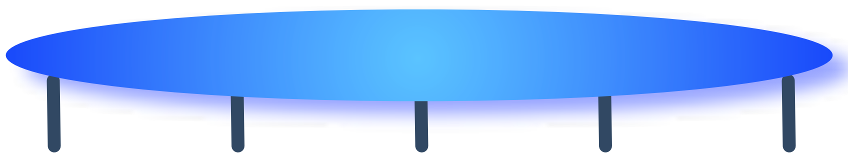


Assume:

$$|\Psi_{\text{MPS}}\rangle = \sum_{\{s_i\},\{\alpha_i\}} A^{(s_1)}_{\alpha_1} A^{(s_2)}_{\alpha_1,\alpha_2} \cdots A^{(s_N)}_{\alpha_{N-1}} |s_1, s_2, \cdots, s_N\rangle$$

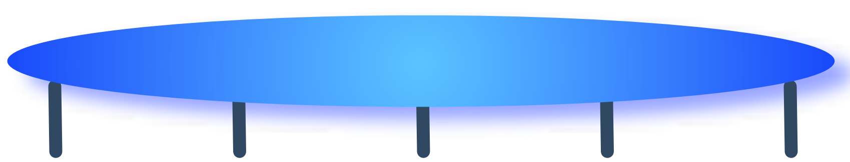
TENSOR NETWORKS STATES

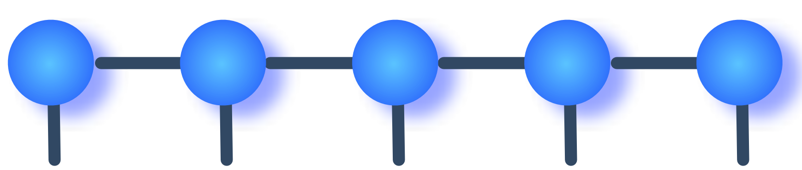
$\psi_{\alpha_1,\alpha_2,\ldots\alpha_N}$ $\mathcal{O}(d^N)$



TENSOR NETWORKS STATES

$\psi_{\alpha_1,\alpha_2,\ldots\alpha_N}$ $\mathcal{O}(d^N)$



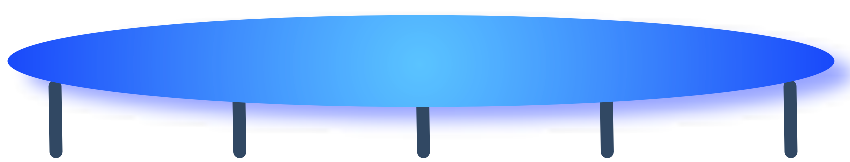


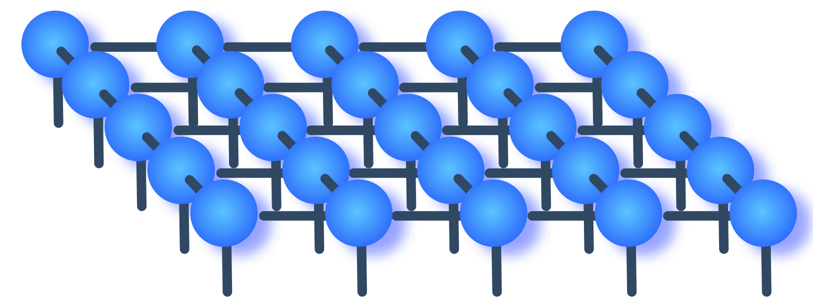
$\mathcal{O}(Ndm^2)$

$A_{\alpha_1}^{\beta_1} A_{\alpha_2}^{\beta_1\beta_2} \ldots A_{\alpha_N}^{\beta_{N-1}}$

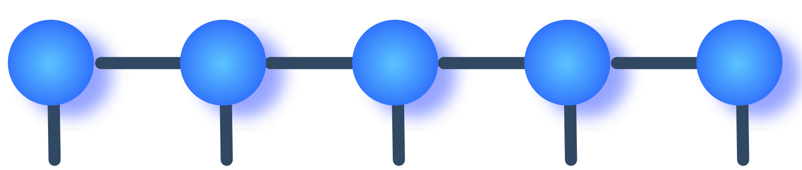
TENSOR NETWORKS STATES

$\psi_{\alpha_1,\alpha_2,\ldots\alpha_N}$ $\mathcal{O}(d^N)$





PEPS



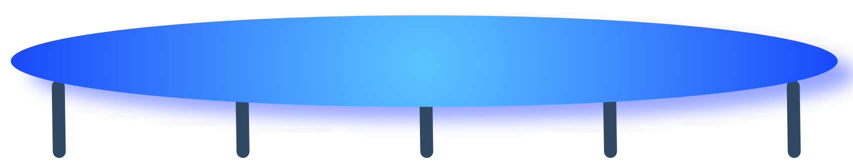
$\mathcal{O}(Ndm^2)$

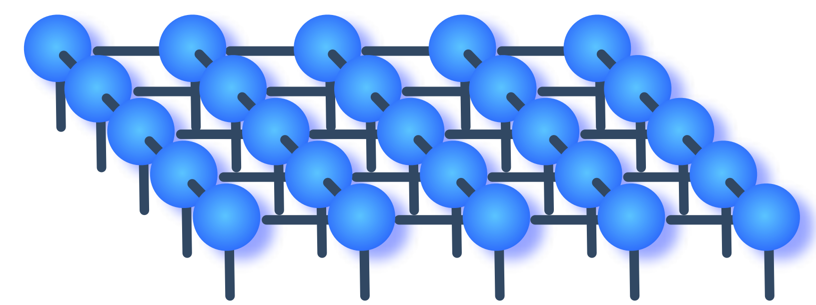
$A_{\alpha_1}^{\beta_1} A_{\alpha_2}^{\beta_1\beta_2} \ldots A_{\alpha_N}^{\beta_{N-1}}$

TENSOR NETWORKS STATES

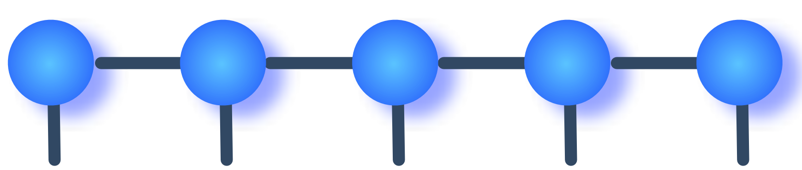
$$\psi_{\alpha_1,\alpha_2,\ldots\alpha_N}$$

$\mathcal{O}(d^N)$



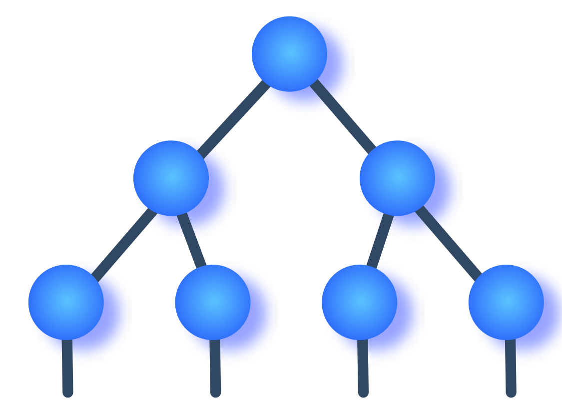


PEPS



$\mathcal{O}(Ndm^2)$

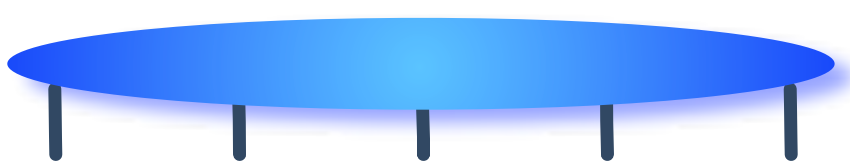
$$A_{\alpha_1}^{\beta_1} A_{\alpha_2}^{\beta_1\beta_2} \ldots A_{\alpha_N}^{\beta_{N-1}}$$

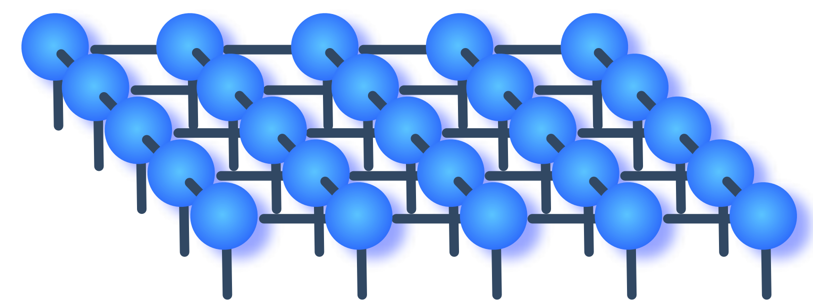


Tree Tensor Network

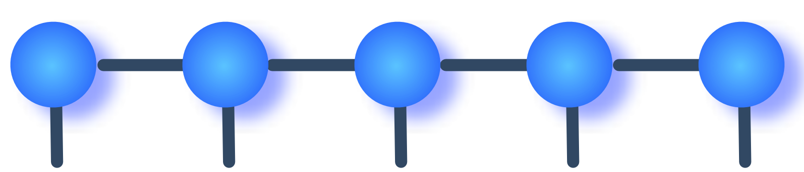
TENSOR NETWORKS STATES

$$\psi_{\alpha_1,\alpha_2,\ldots\alpha_N} \qquad \mathcal{O}(d^N)$$

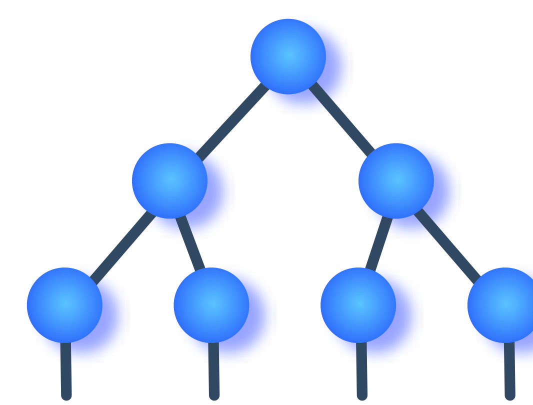




PEPS



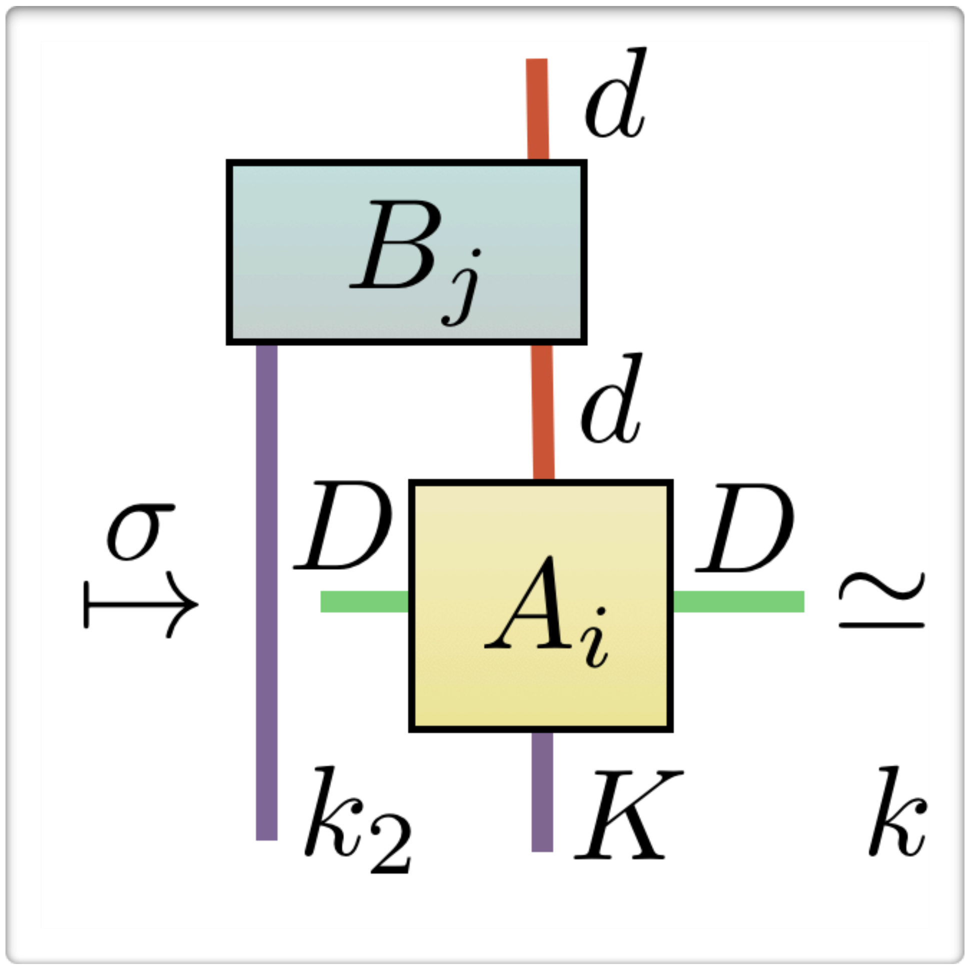
$$A^{\beta_1}_{\alpha_1} A^{\beta_1\beta_2}_{\alpha_2} \ldots A^{\beta_{N-1}}_{\alpha_N} \qquad \mathcal{O}(Ndm^2)$$



Tree Tensor Network

Tensor networks states are a compressed description of the system tunable between mean field and exact

TENSOR NETWORK ALGORITHMS

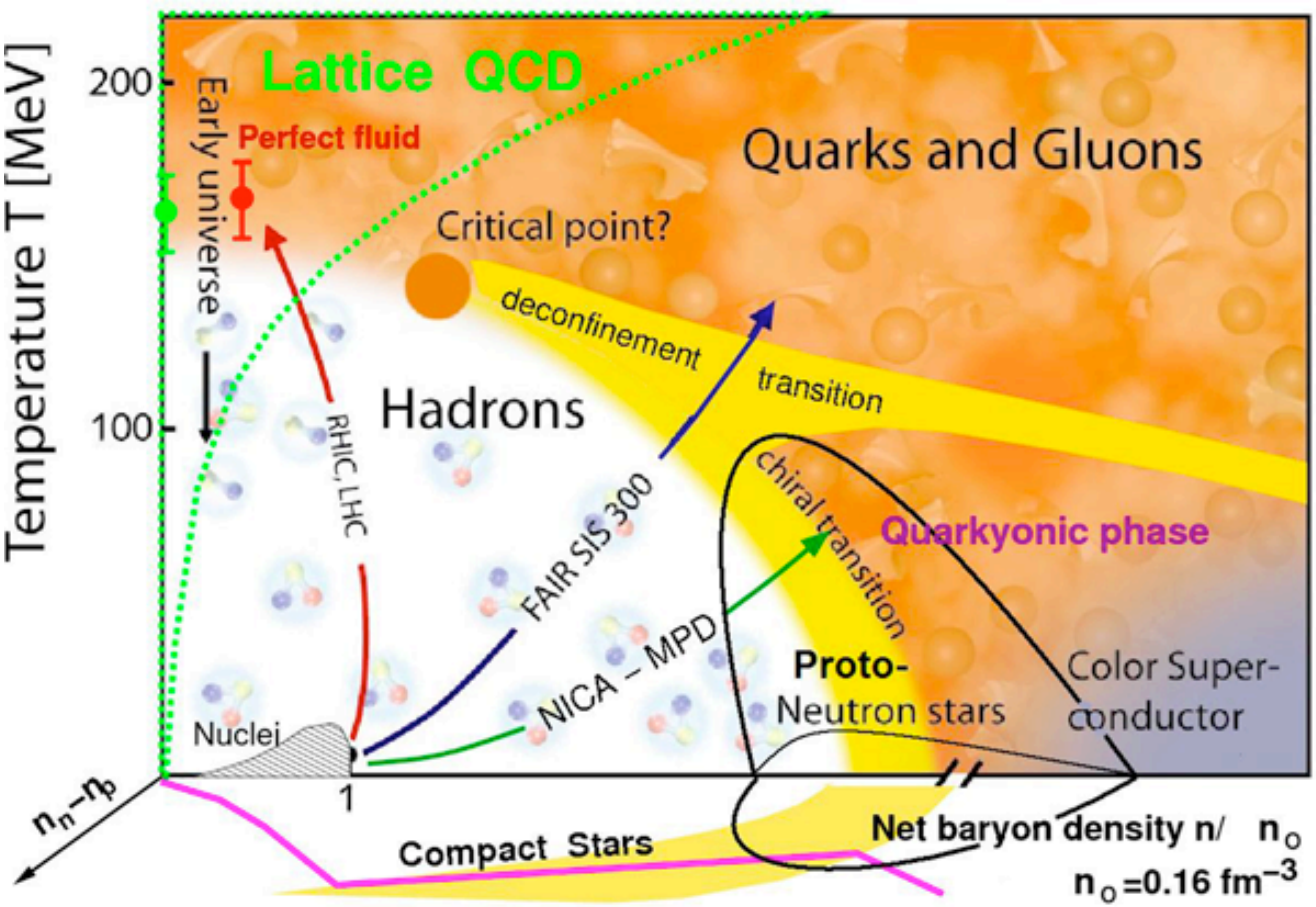


- *State of the art in 1D (poly effort)*
- *No sign problem*
- *Extended to open quantum systems*
- ***Machine learning***
- *Data compression (BIG DATA)*
- *Extended to lattice gauge theories*
- *Simulations of low-entangled systems of hundreds qubits*
- *Extended to quantum field theories*

S. Montangero "Introduction to Tensor Network Methods", Springer (2019)

U. Schollwock, RMP (2005) ***A. Cichocki, ECM (2013)*** ***I. Glasser, et al. PRX (2018)***

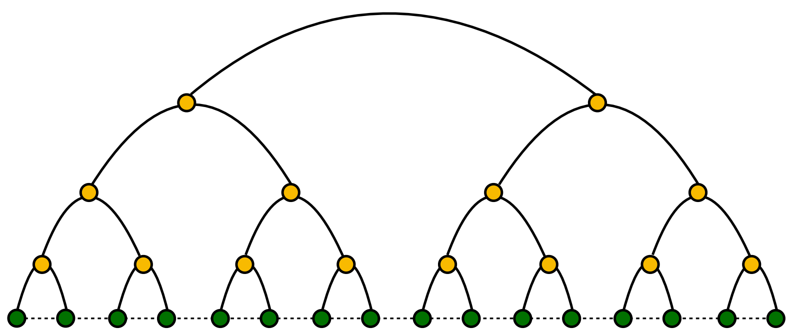
LATTICE GAUGE THEORIES

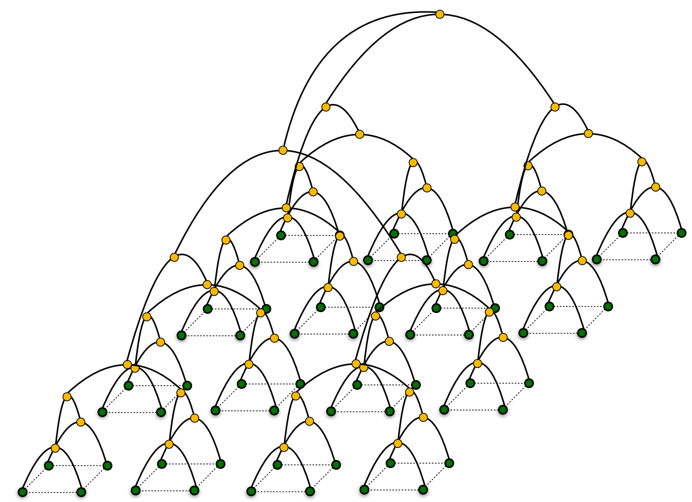


The current wisdom on the phase diagram of nuclear matter.

McLerran, L. Nucl.Phys.Proc.Suppl. 195 (2009) 275-280

3D TREE TENSOR NETWORK



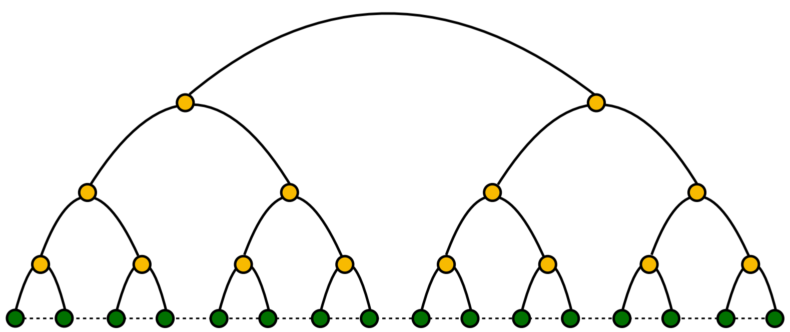


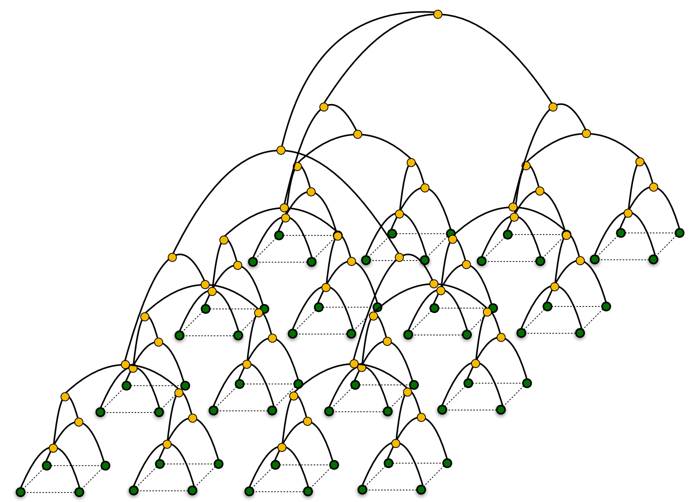
T. Felser, P. Silvi, M. Collura,
S. Montangero
PRX (2020)

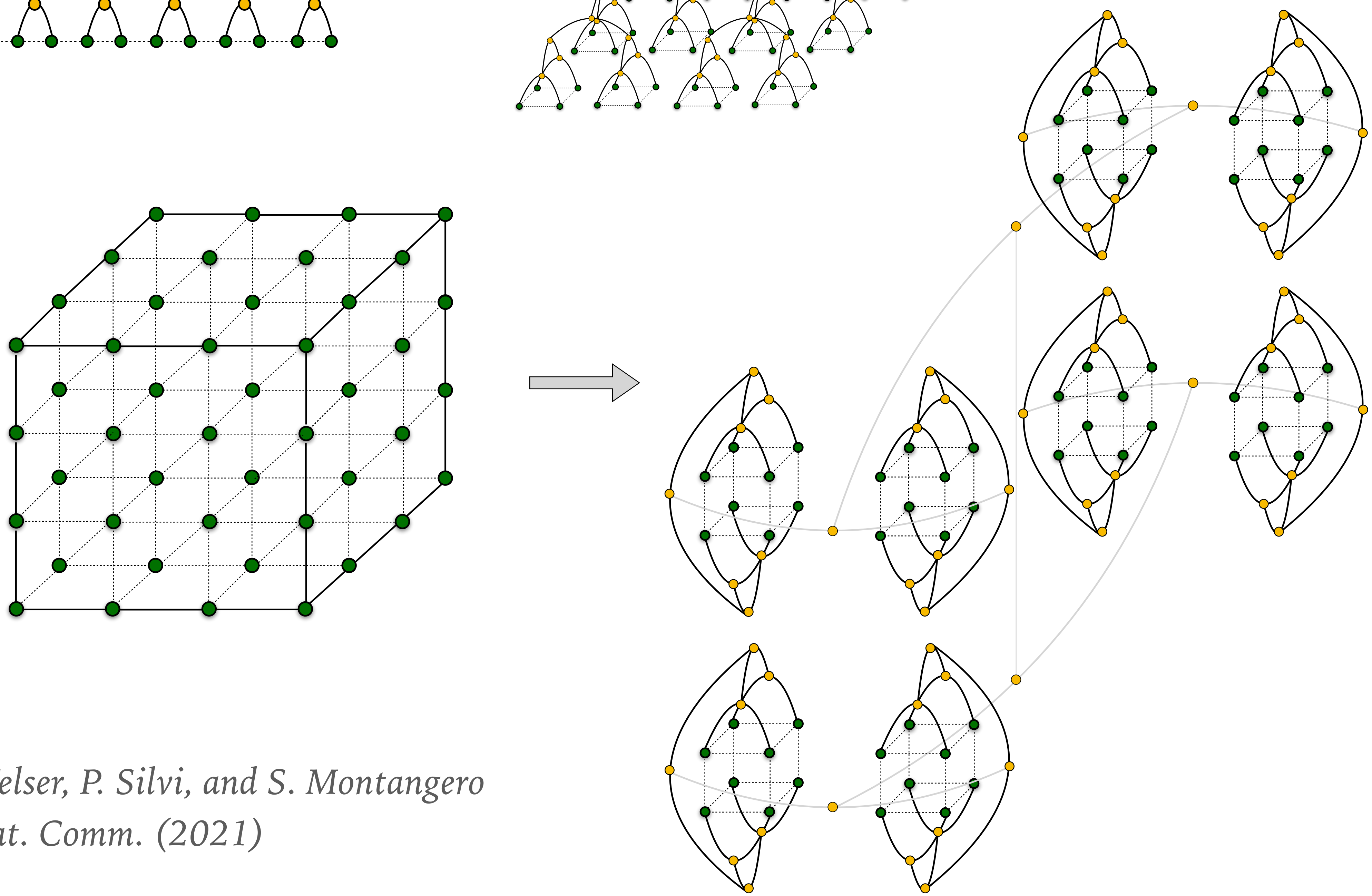
G. Magnifico, T. Felser, P. Silvi, and S. Montangero
Nat. Comm. (2021)

3D TREE TENSOR NETWORK

T. Felser, P. Silvi, M. Collura,
S. Montangero
PRX (2020)







G. Magnifico, T. Felser, P. Silvi, and S. Montangero
Nat. Comm. (2021)

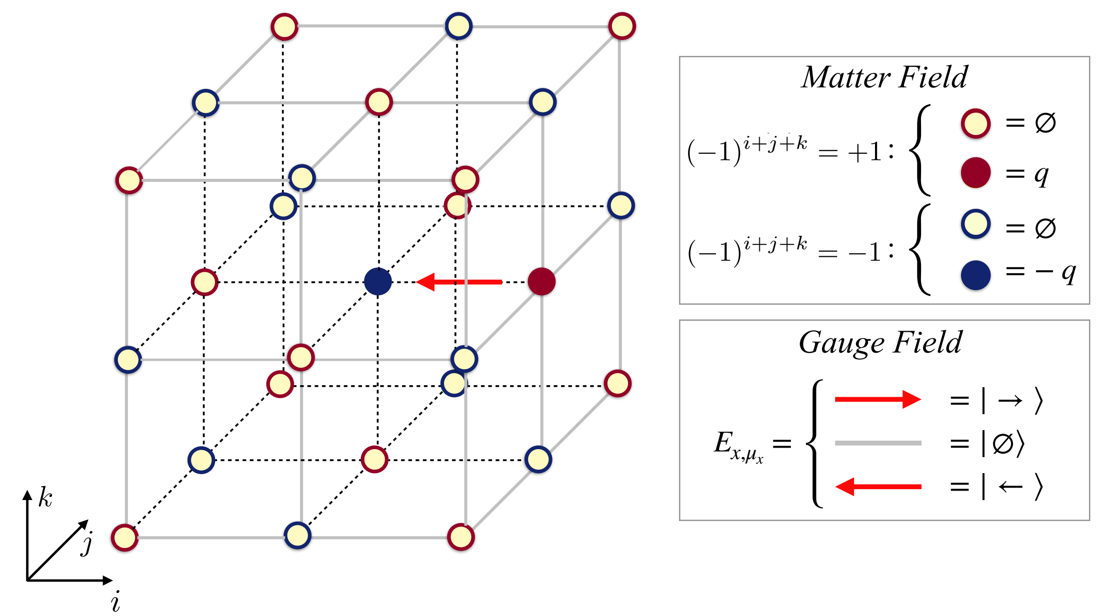
3D QUANTUM-LINK FORMULATION OF QED

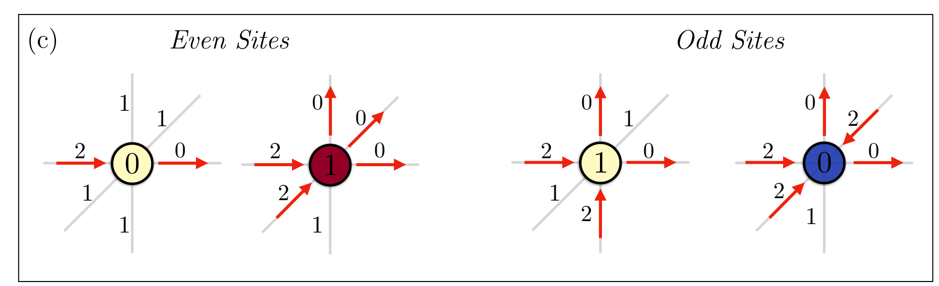
$$\hat{H} = -t \sum_{x,\mu} \left(\hat{\psi}_x^\dagger \, \hat{U}_{x,\mu} \, \hat{\psi}_{x+\mu} + \text{H.c.} \right)$$

$$+ \, m \sum_x (-1)^x \hat{\psi}_x^\dagger \hat{\psi}_x + \frac{g_e^2}{2} \sum_{x,\mu} \hat{E}_{x,\mu}^2$$

$$- \, \frac{g_m^2}{2} \sum_x \left(\square_{\mu_x,\mu_y} + \square_{\mu_x,\mu_z} + \square_{\mu_y,\mu_z} + \text{H.c.} \right)$$

$$\hat{G}_x = \hat{\psi}_x^\dagger \hat{\psi}_x - \frac{1-(-1)^x}{2} - \sum_\mu \hat{E}_{x,\mu}$$

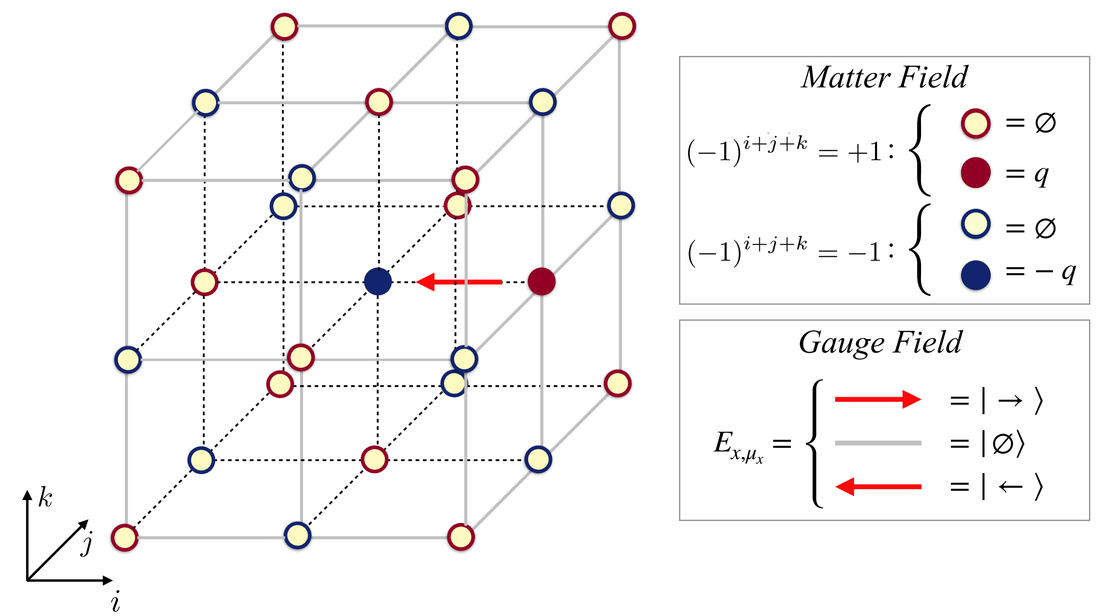




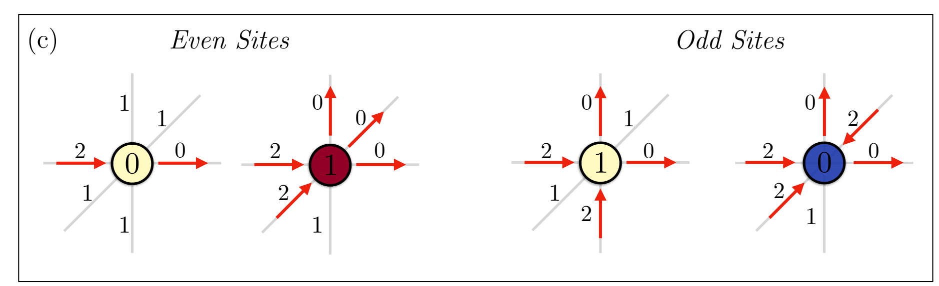
$$H_{pen} = \nu \sum_{x,\mu} \left(1 - \delta_{2,\hat{L}_{x,\mu}} \right)$$

3D QUANTUM-LINK FORMULATION OF QED

$$\hat{H} = -t \sum_{x,\mu} \left(\hat{\psi}_x^\dagger \, \hat{U}_{x,\mu} \, \hat{\psi}_{x+\mu} + \text{H.c.} \right)$$
$$+ \, m \sum_x (-1)^x \hat{\psi}_x^\dagger \hat{\psi}_x + \frac{g_e^2}{2} \sum_{x,\mu} \hat{E}_{x,\mu}^2$$
$$- \, \frac{g_m^2}{2} \sum_x \left(\Box_{\mu_x,\mu_y} + \Box_{\mu_x,\mu_z} + \Box_{\mu_y,\mu_z} + \text{H.c.} \right)$$



$$\hat{G}_x = \hat{\psi}_x^\dagger \hat{\psi}_x - \frac{1-(-1)^x}{2} - \sum_\mu \hat{E}_{x,\mu}$$



Local dimension 267, up to 12288 Hamiltonian operators

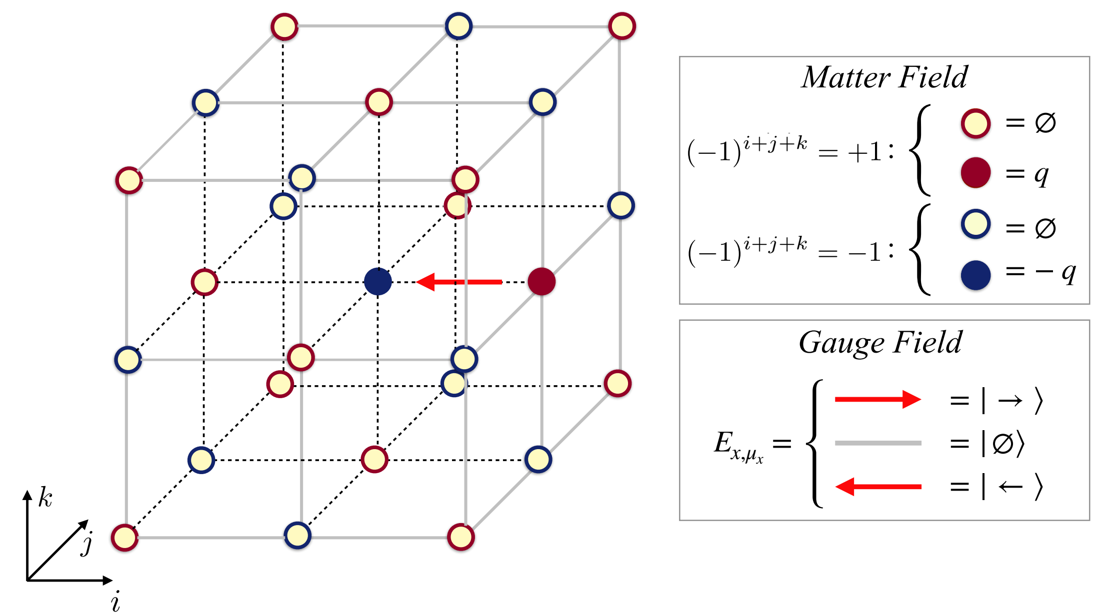
$$H_{pen} = \nu \sum_{x,\mu} \left(1 - \delta_{2,\hat{L}_{x,\mu}} \right)$$

3D QUANTUM-LINK FORMULATION OF QED

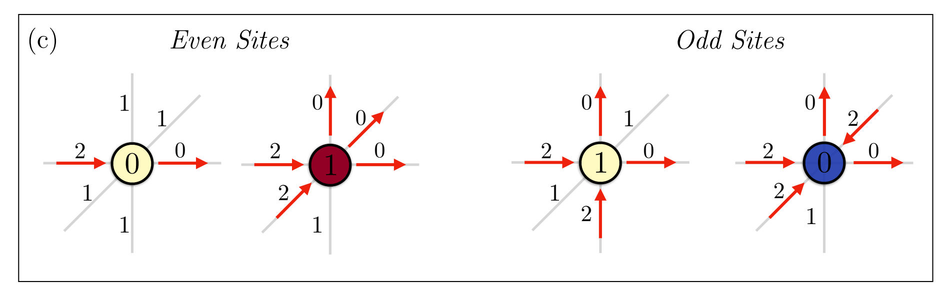
$$\hat{H} = -t \sum_{x,\mu} \left(\hat{\psi}_x^\dagger \, \hat{U}_{x,\mu} \, \hat{\psi}_{x+\mu} + \text{H.c.} \right)$$

$$+ \, m \sum_x (-1)^x \hat{\psi}_x^\dagger \hat{\psi}_x + \frac{g_e^2}{2} \sum_{x,\mu} \hat{E}_{x,\mu}^2$$

$$- \, \frac{g_m^2}{2} \sum_x \left(\Box_{\mu_x,\mu_y} + \Box_{\mu_x,\mu_z} + \Box_{\mu_y,\mu_z} + \text{H.c.} \right)$$



$$\hat{G}_x = \hat{\psi}_x^\dagger \hat{\psi}_x - \frac{1-(-1)^x}{2} - \sum_\mu \hat{E}_{x,\mu}$$



$$H_{pen} = \nu \sum_{x,\mu} \left(1 - \delta_{2,\hat{L}_{x,\mu}} \right)$$

Local dimension 267, up to 12288 Hamiltonian operators

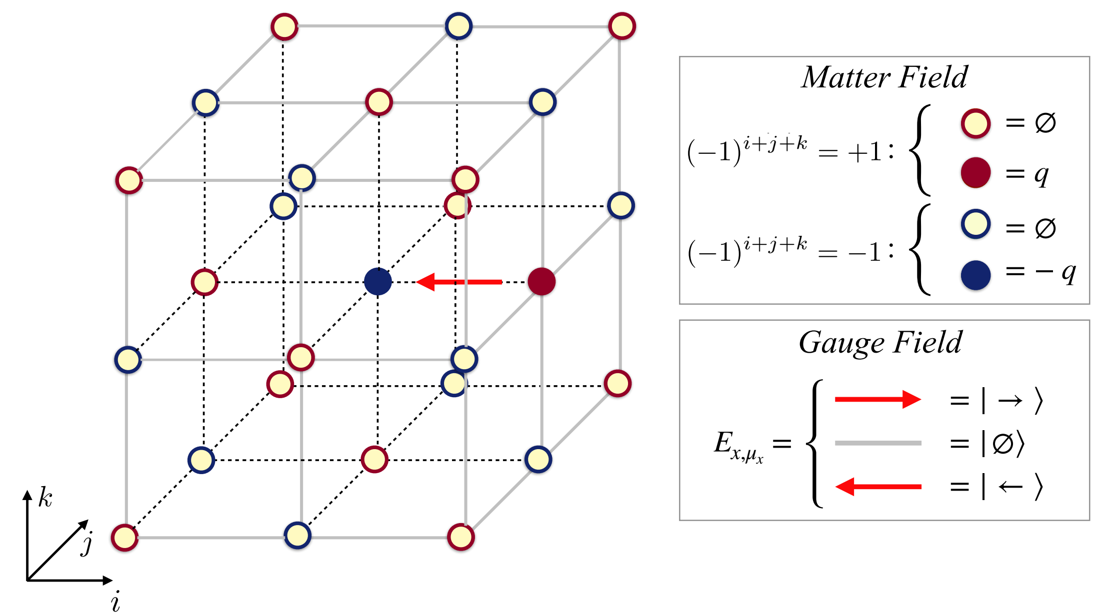
Up to 5 weeks x 64 cores of computational time

3D QUANTUM-LINK FORMULATION OF QED

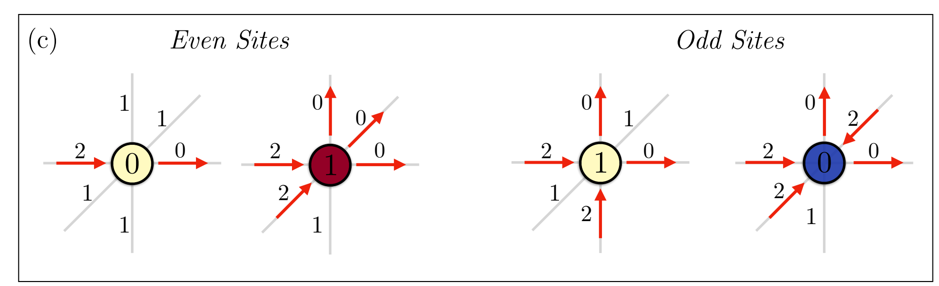
$$\hat{H} = -t \sum_{x,\mu} \left(\hat{\psi}_x^\dagger \, \hat{U}_{x,\mu} \, \hat{\psi}_{x+\mu} + \text{H.c.} \right)$$

$$+ \, m \sum_x (-1)^x \hat{\psi}_x^\dagger \hat{\psi}_x + \frac{g_e^2}{2} \sum_{x,\mu} \hat{E}_{x,\mu}^2$$

$$- \, \frac{g_m^2}{2} \sum_x \left(\Box_{\mu_x,\mu_y} + \Box_{\mu_x,\mu_z} + \Box_{\mu_y,\mu_z} + \text{H.c.} \right)$$



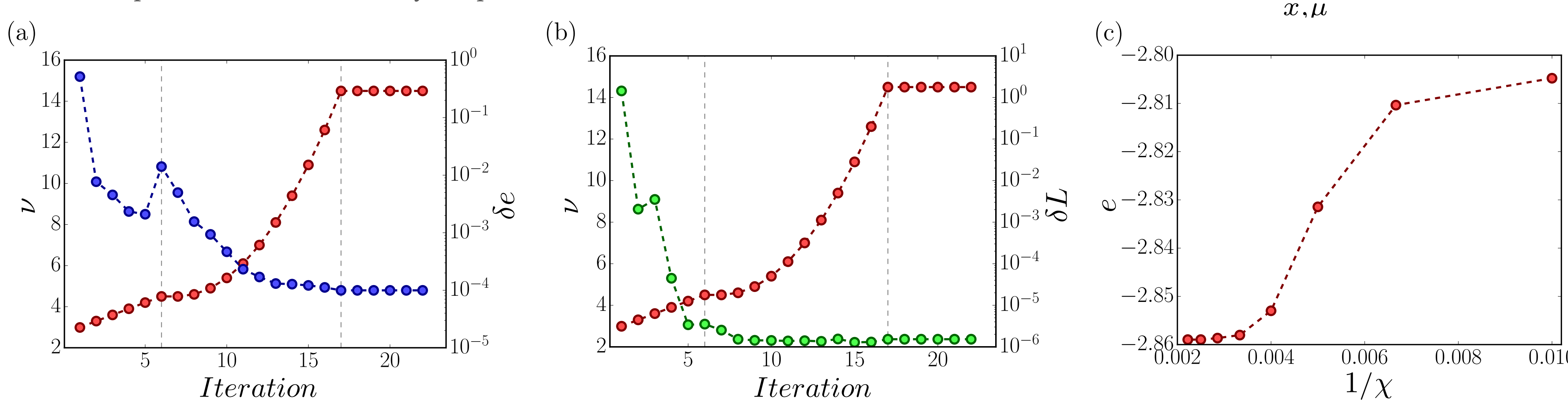
$$\hat{G}_x = \hat{\psi}_x^\dagger \hat{\psi}_x - \frac{1-(-1)^x}{2} - \sum_\mu \hat{E}_{x,\mu}$$



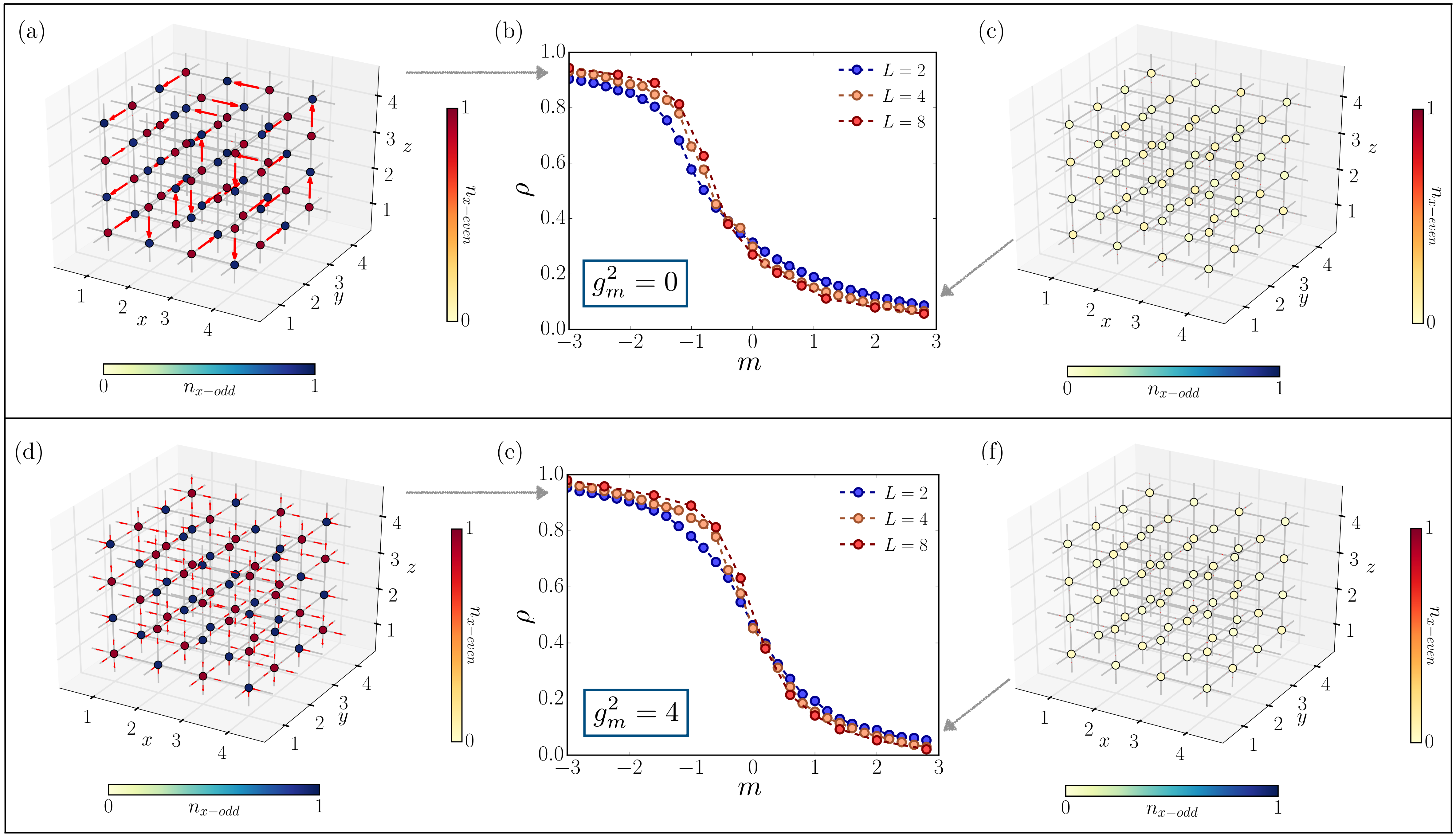
Local dimension 267, up to 12288 Hamiltonian operators

$$H_{pen} = \nu \sum_{x,\mu} \left(1 - \delta_{2,\hat{L}_{x,\mu}} \right)$$

Up to 5 weeks x 64 cores of computational time



QUANTUM PHASES



$m_c \approx +0.22$

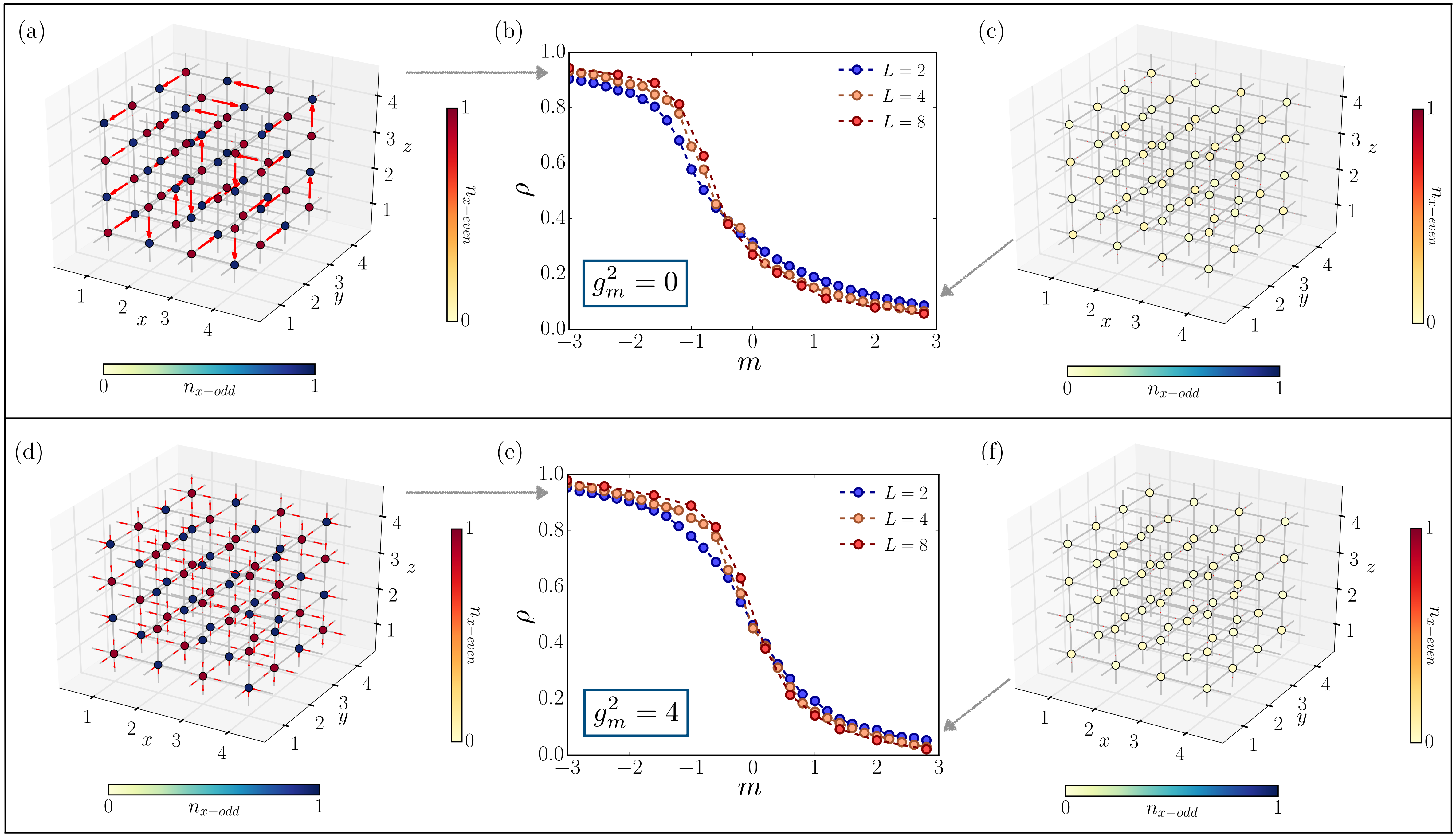
$g_m^2 = 8/g_e^2$

QUANTUM PHASES

Hilbert space of

~64x64x64 qubits!

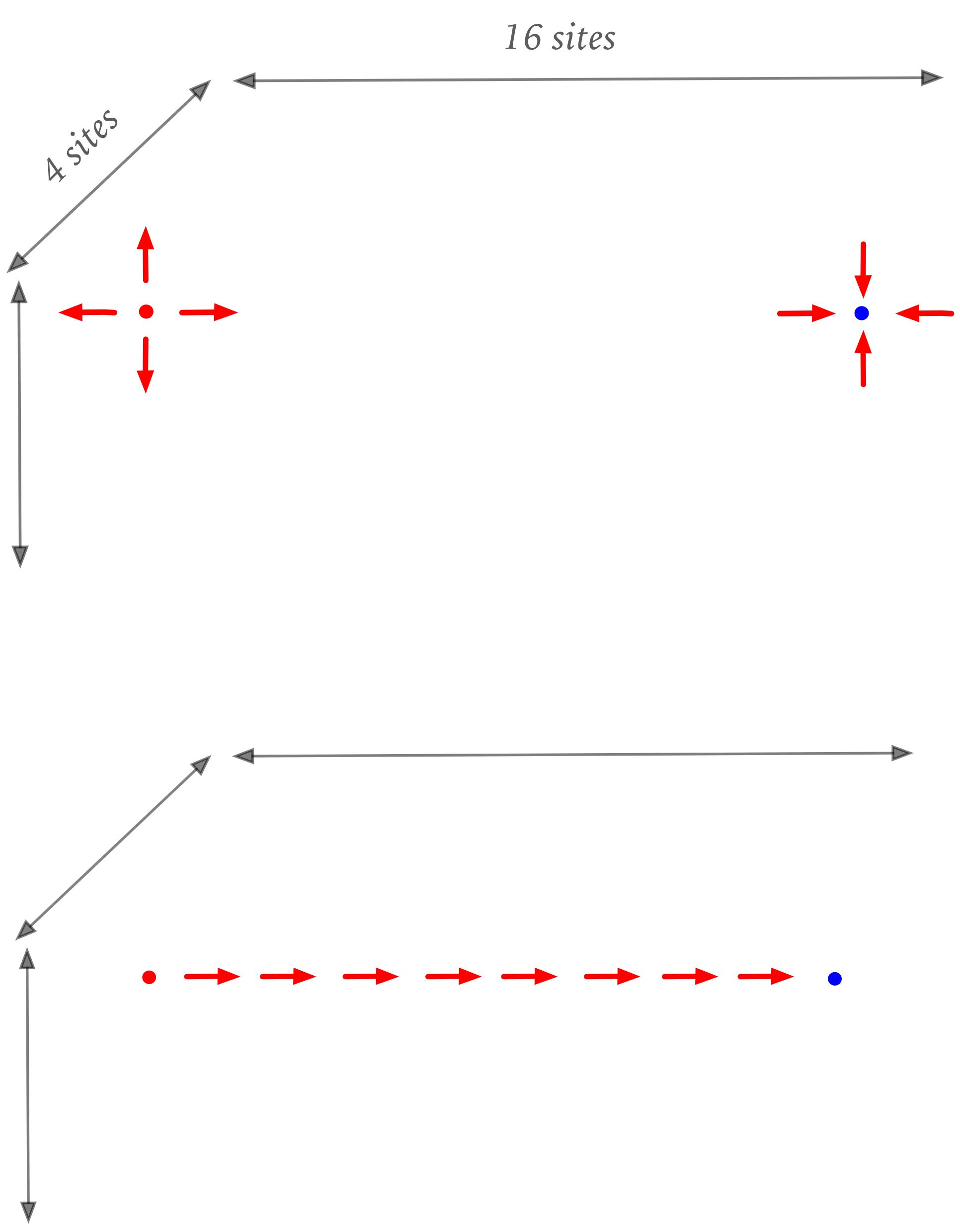
200Kb QRAM

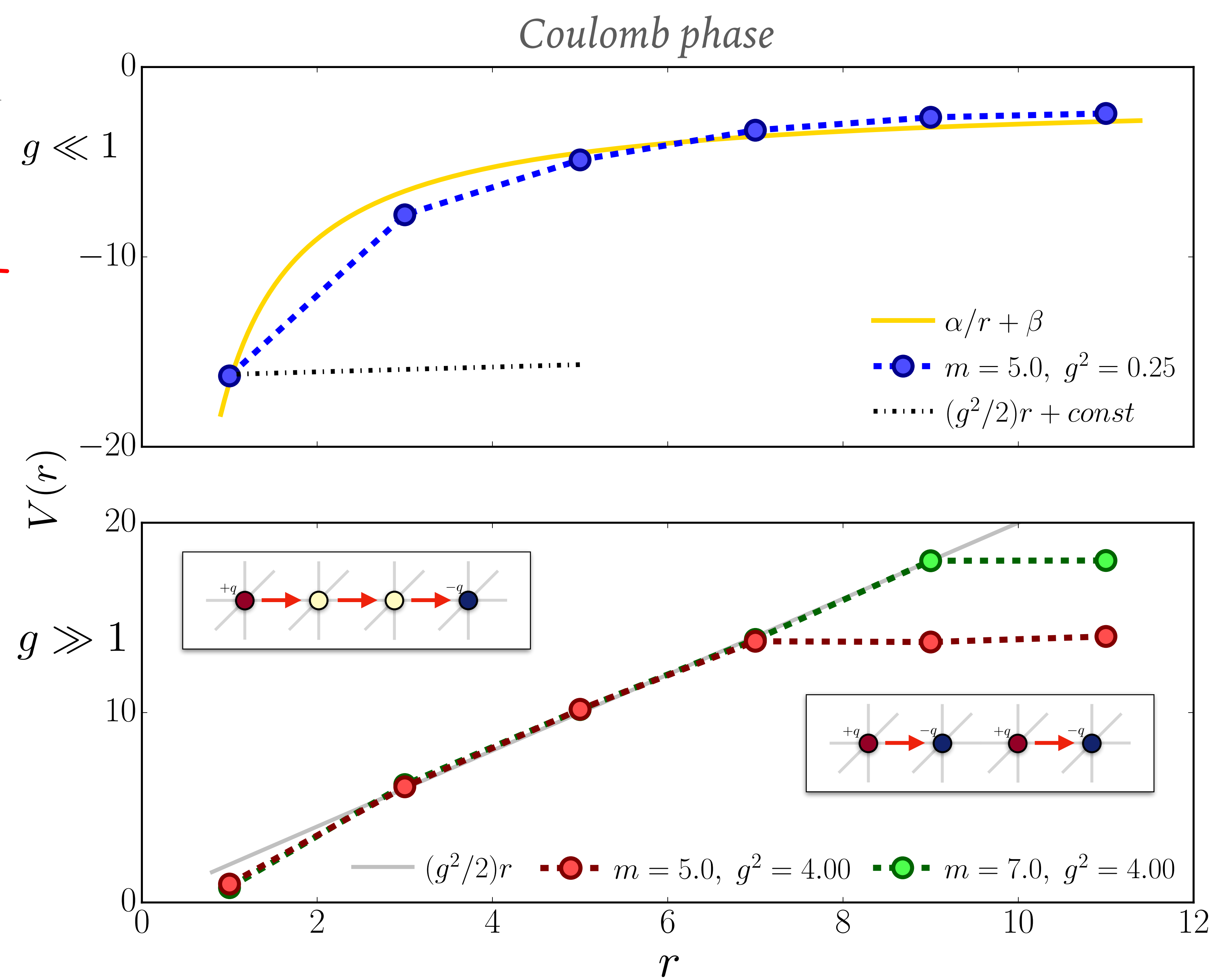


$m_c \approx +0.22$

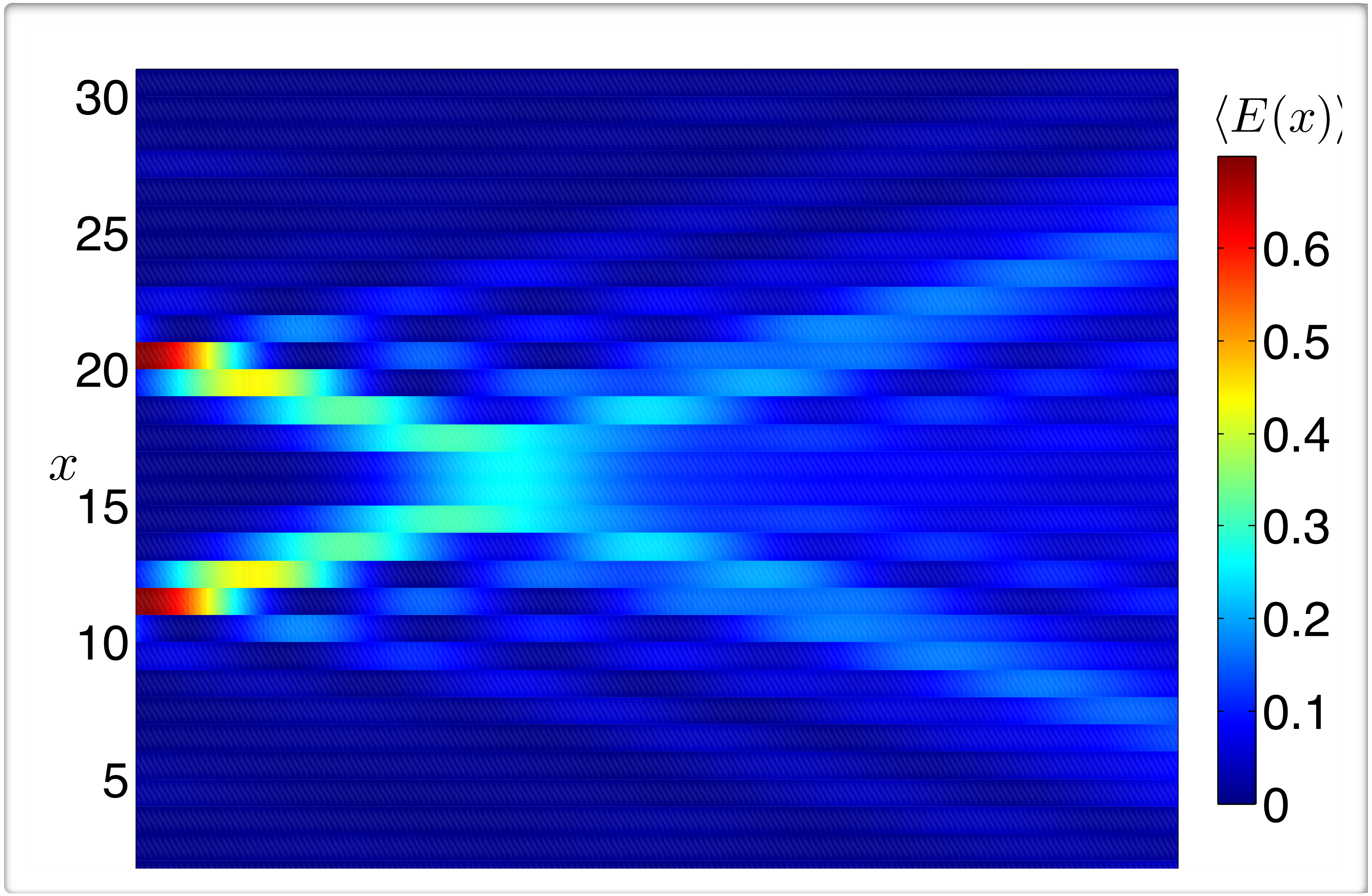
$g_m^2 = 8/g_e^2$

CONFINEMENT





$$g_e^2 = g^2/a,\ g_m^2 = 8/(g^2 a)$$

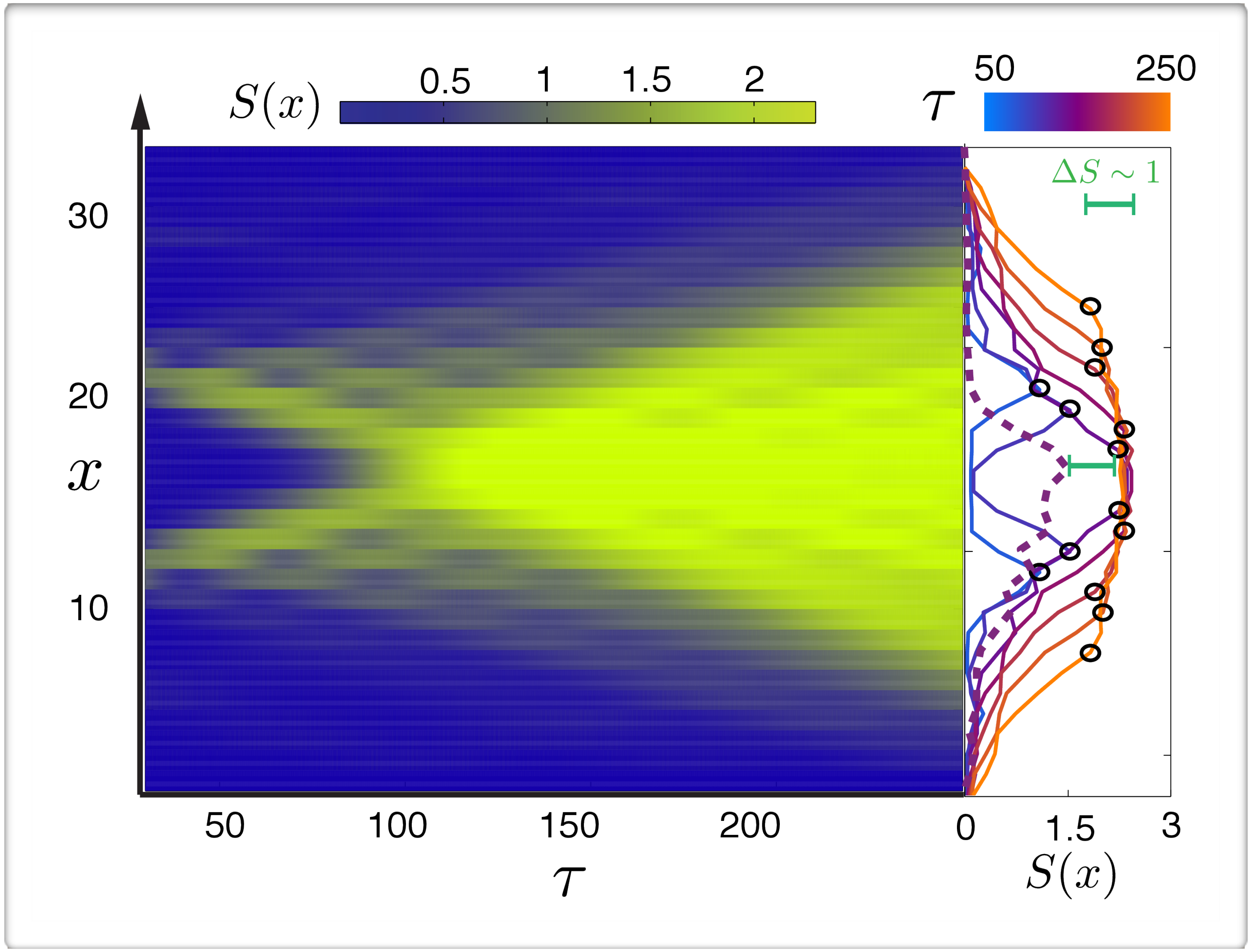


Space

Real time

MESONS SCATTERING

T. Pichler, E. Rico, M. Dalmonte, P. Zoller, and SM, PRX (2016)



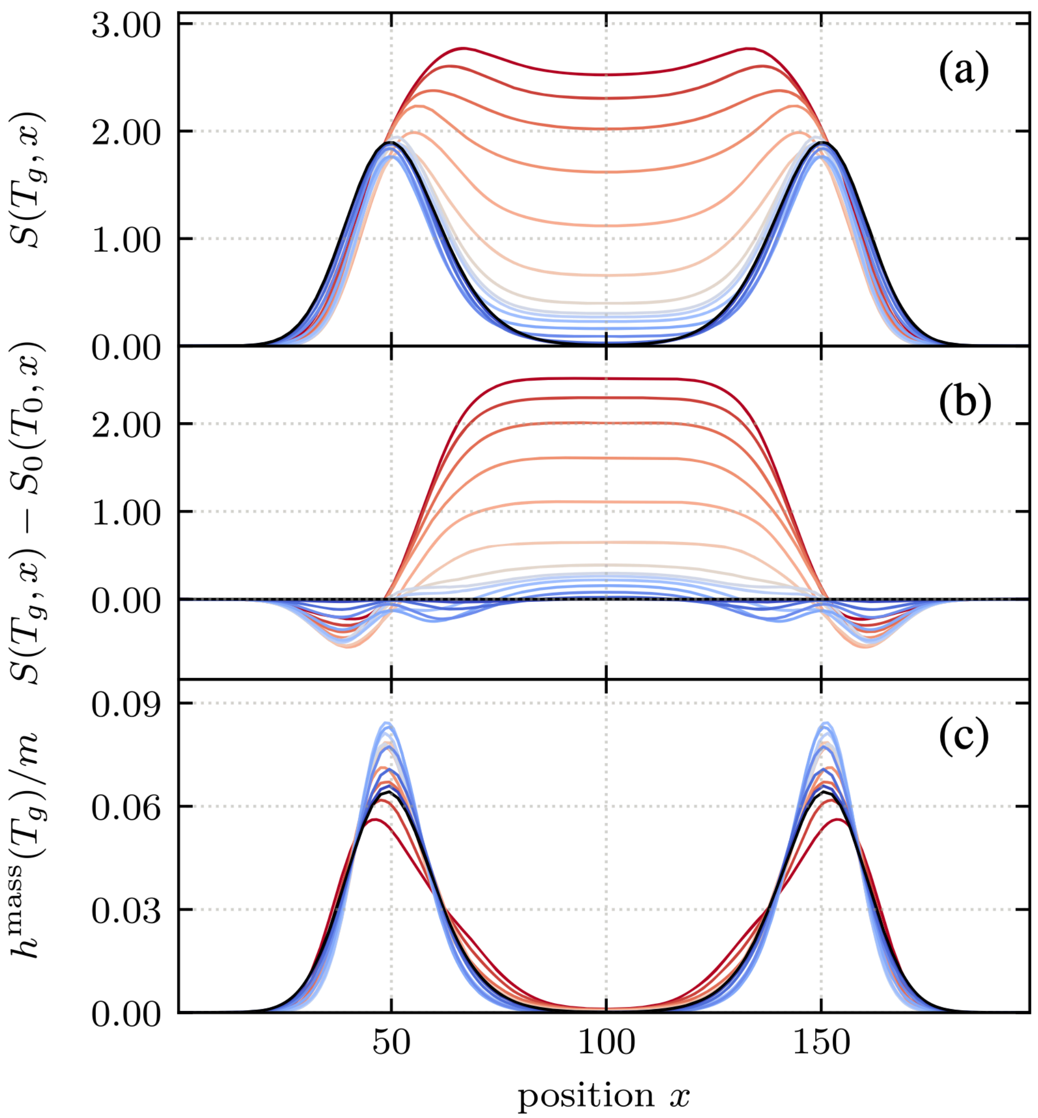
Space

Real time

MESONS SCATTERING

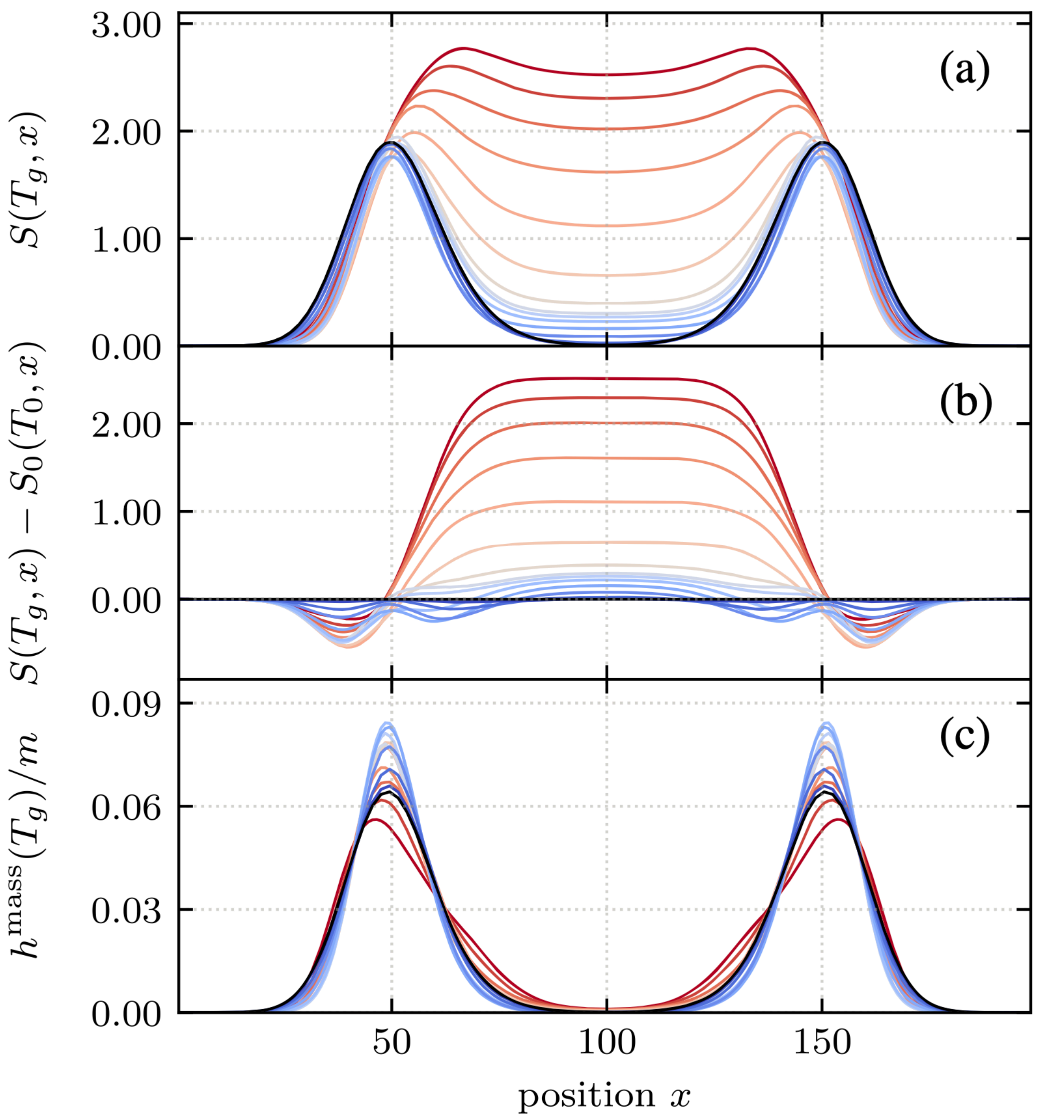
T. Pichler, E. Rico, M. Dalmonte, P. Zoller, and SM, PRX (2016)

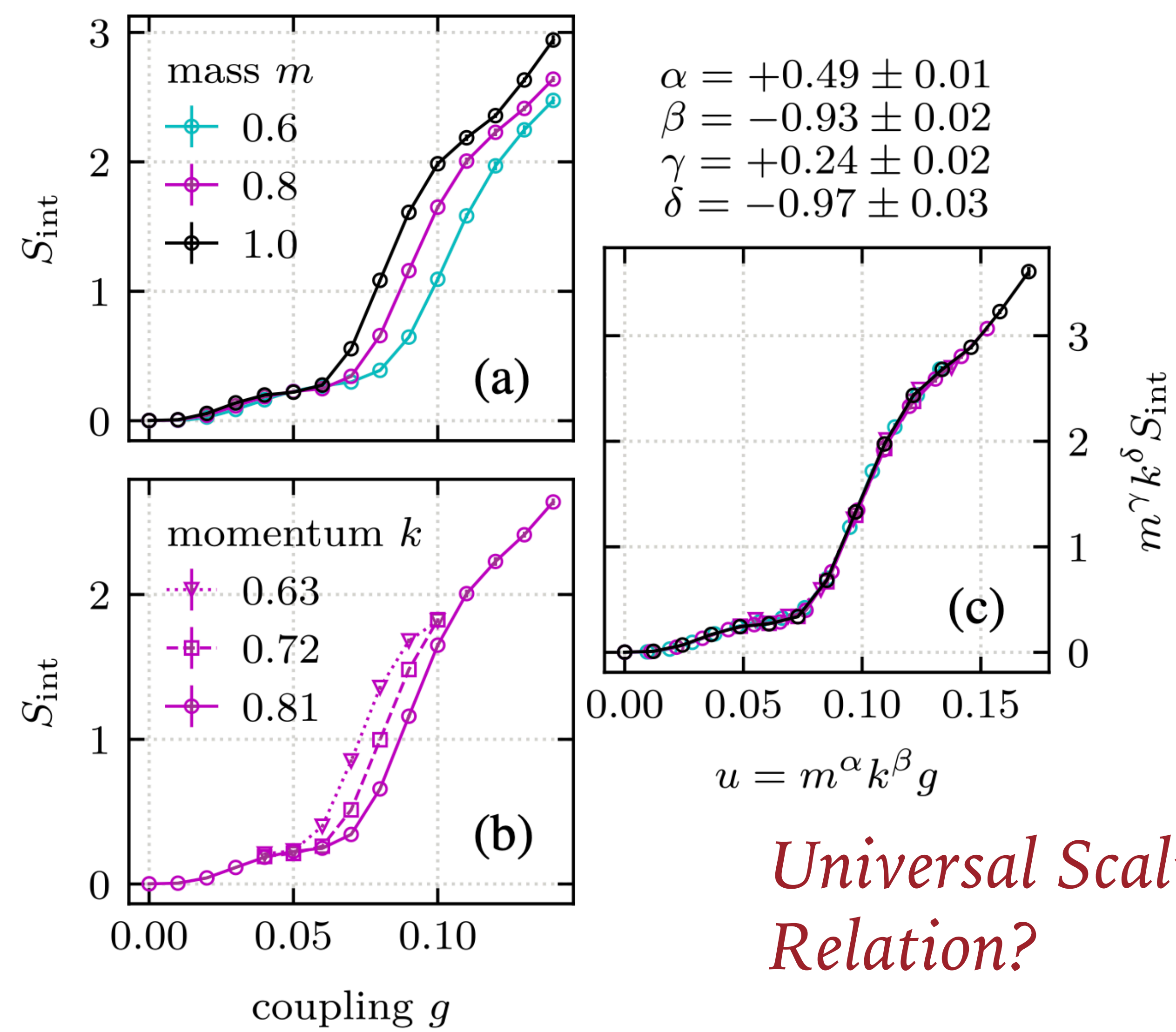
ENTANGLEMENT GENERATION IN QED SCATTERING PROCESSES



M. Rigobello, S. Notarnicola, G. Magnifico, and S. Montangero, Phys. Rev. D 104, 114501 (2021).

ENTANGLEMENT GENERATION IN QED SCATTERING PROCESSES

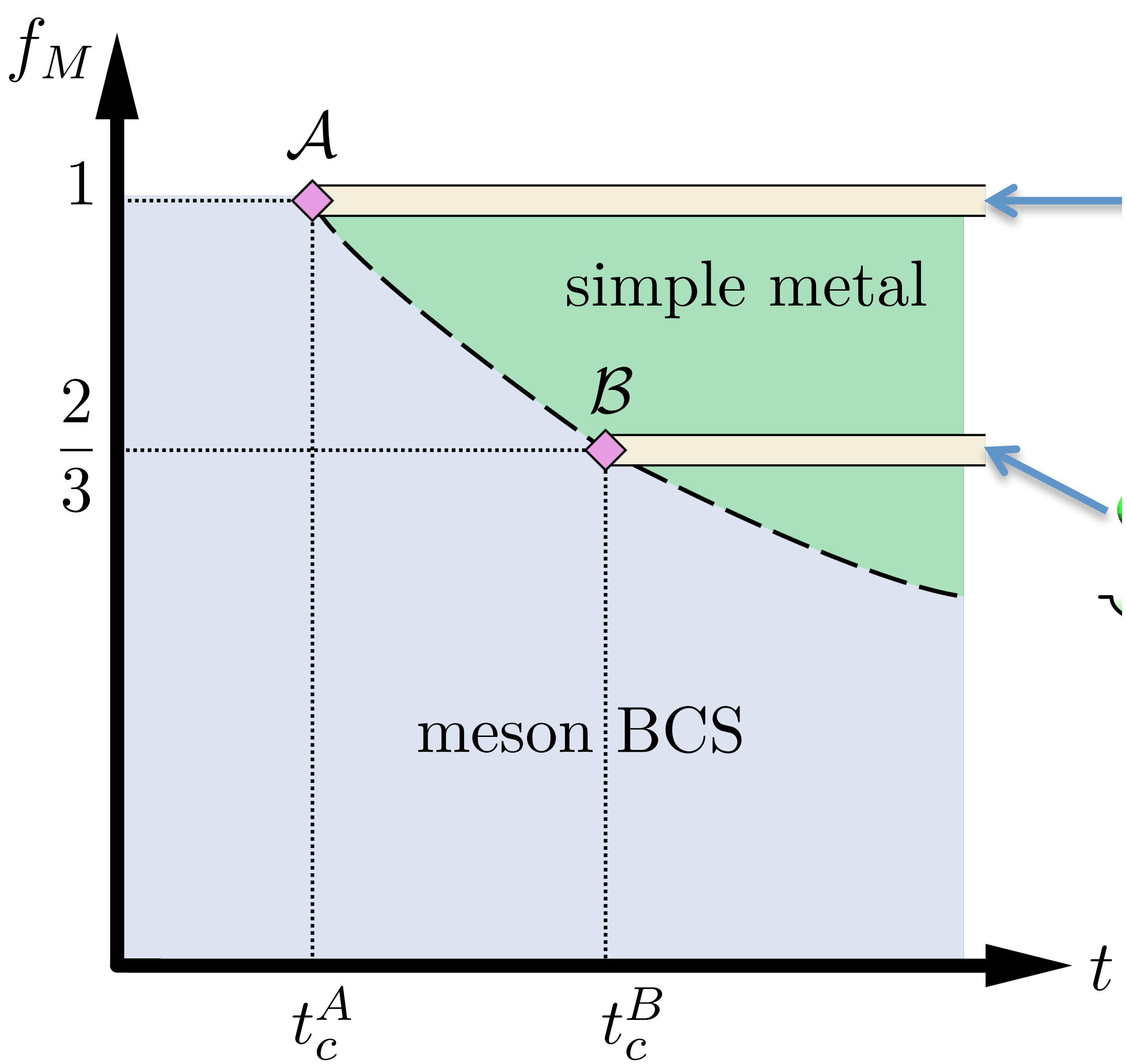




Universal Scaling Relation?

M. Rigobello, S. Notarnicola, G. Magnifico, and S. Montangero, Phys. Rev. D 104, 114501 (2021).

SU(2) LATTICE GAUGE THEORY IN 1+1D



$$H = H_{\text{coupl}} + H_{\text{free}} + H_{\text{break}}$$

$$H_{\text{coupl}} = t \sum_{j=1}^{\mathrm{L}-1} \sum_{s,s'=\uparrow,\downarrow} c_{j,s}^{[M]\dagger} U_{j,j+1;s,s'} c_{j+1,s'}^{[M]} + \text{h.c.}$$

$$H_{\text{free}} = \frac{g_0^2}{2} \sum_{j=1}^{\mathrm{L}} \left[\vec{J}_{j-1,j}^{[R]}\right]^2 + \left[\vec{J}_{j,j+1}^{[L]}\right]^2$$

Phase diagram at

finite chemical potential

P. Silvi et al Quantum (2017)

SU(3) 1+1D study in P. Silvi, et al. PRD (2019)

Simulating Lattice Gauge Theories within Quantum Technologies

M.C. Bañuls[1,2], R. Blatt[3,4], J. Catani[5,6,7], A. Celi[3,8], J.I. Cirac[1,2], M. Dalmonte[9,10], L. Fallani[5,6,7], K. Jansen[11], M. Lewenstein[8,12,13], S. Montangero[7,14 a], C.A. Muschik[3], B. Reznik[15], E. Rico[16,17 b], L. Tagliacozzo[18], K. Van Acoleyen[19], F. Verstraete[19,20], U.-J. Wiese[21], M. Wingate[22], J. Zakrzewski[23,24], and P. Zoller[3]

EPJD (2020)

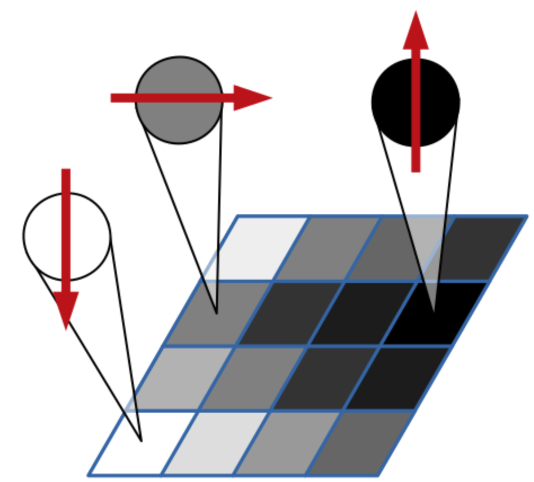
QUANTERA

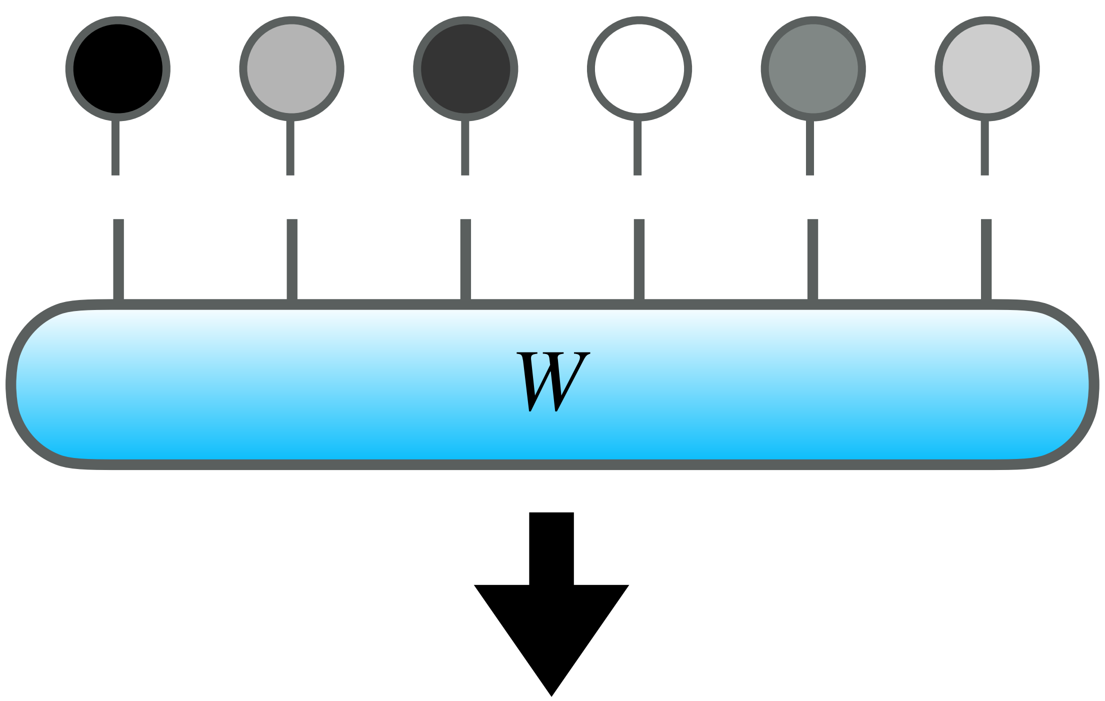
MACHINE LEARNING WITH TENSOR NETWORKS

Raw data

Map to "Spins"

$$\Phi(x_j) = \left[\cos\left(\frac{\pi}{2}x_j\right), \sin\left(\frac{\pi}{2}x_j\right)\right]$$





f(x)

$$f^{\ell}(\bar{x}) = \sum_{\mathbf{s}} W^{\ell}_{s_1 s_2 \ldots s_N} \phi(x_1)^{s_1} \phi(x_2)^{s_2} \ldots \phi(x_N)^{s_N}$$

W : weight tensor

f(x) : decision function

Stoudenmire, Advances in Neural IPS 29, 4799 (2016), arXiv:1605.05775

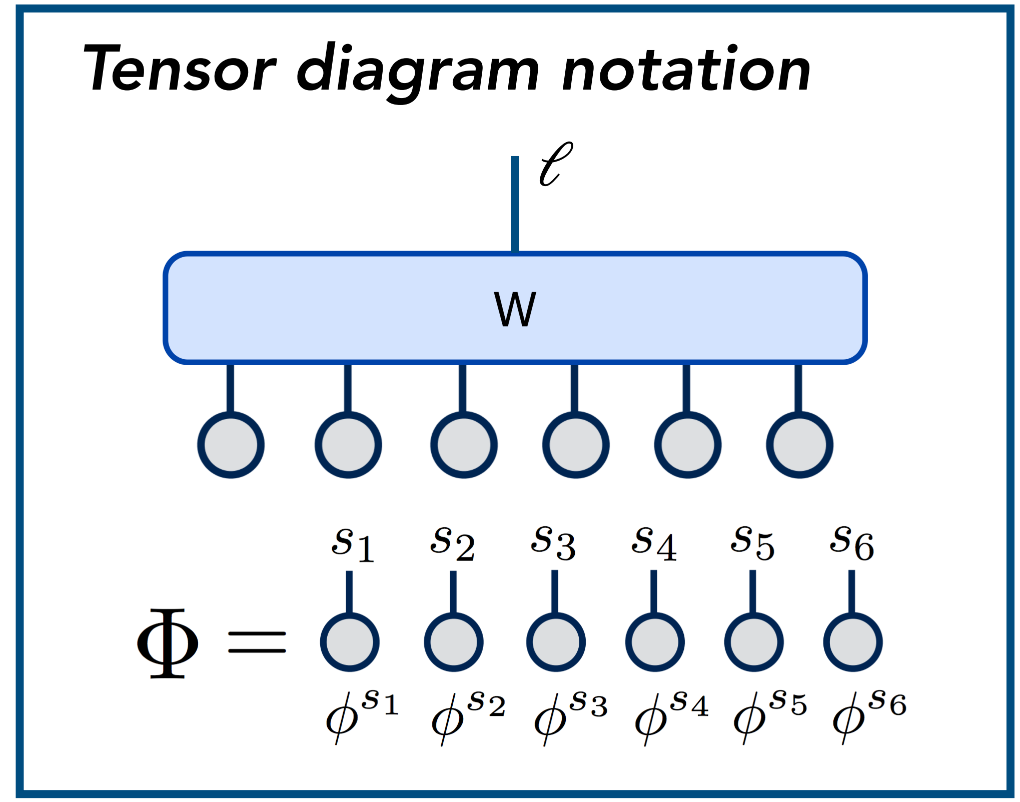
TN MACHINE LEARNING OF HEP DATA

Hypothesis class: $f^{\ell}(\bar{x}) = \boldsymbol{W}^{\ell} \cdot \Phi(\bar{x})$

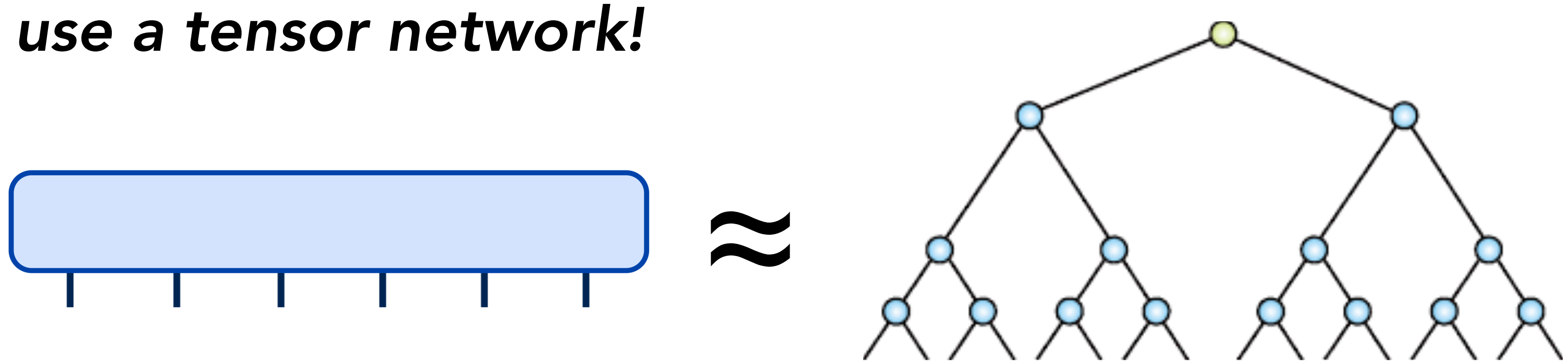
$$f^{\ell}(\bar{x}) = \sum_{\mathbf{s}} W^{\ell}_{s_1 s_2 \ldots s_N} \phi(x_1)^{s_1} \phi(x_2)^{s_2} \ldots \phi(x_N)^{s_N}$$

f^{ℓ} *map input data to the space of labels*

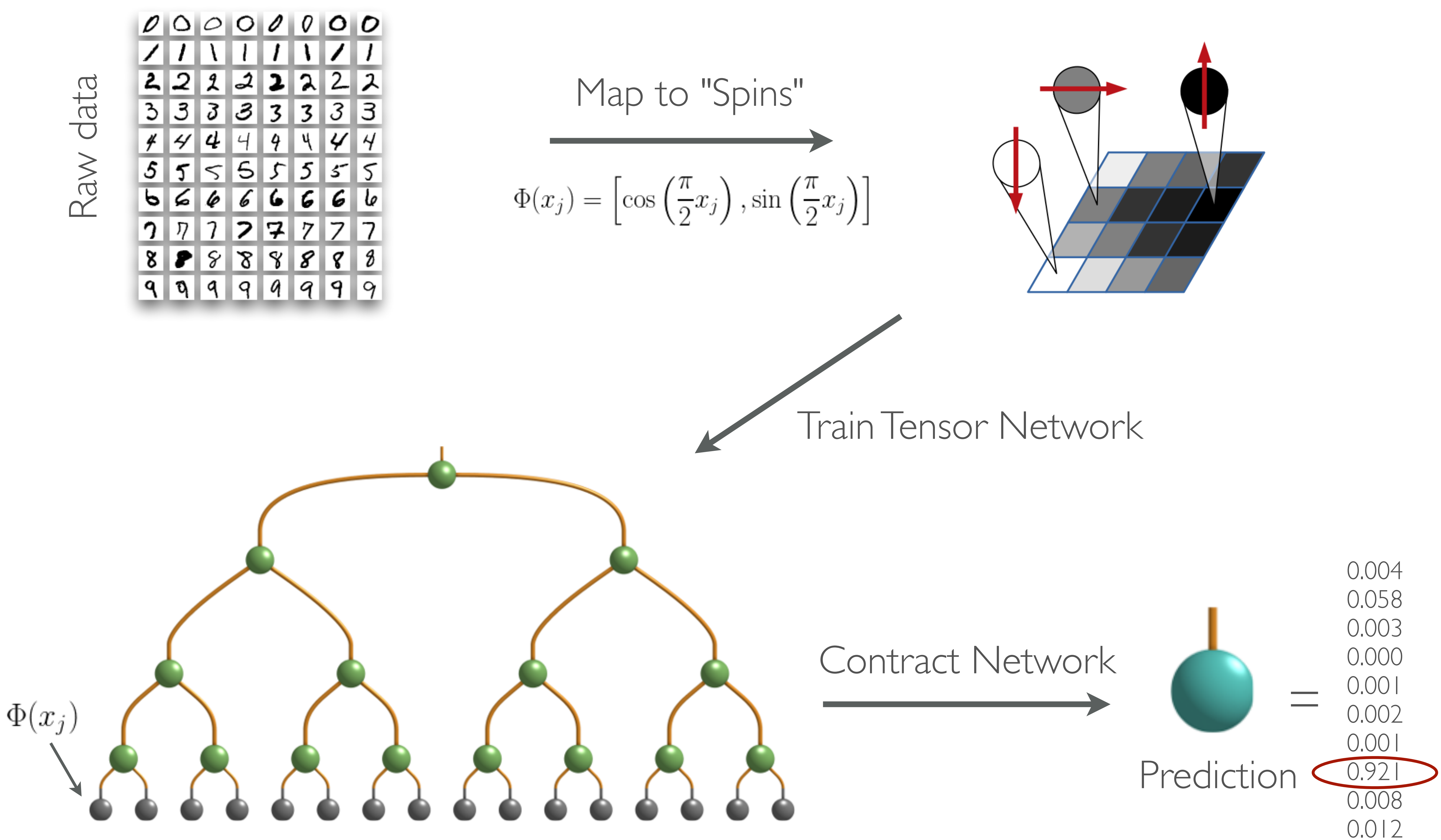
PROBLEM: *W is a N+1 order tensor that grows exponentially with the input data*



SOLUTION: *use a tensor network!*

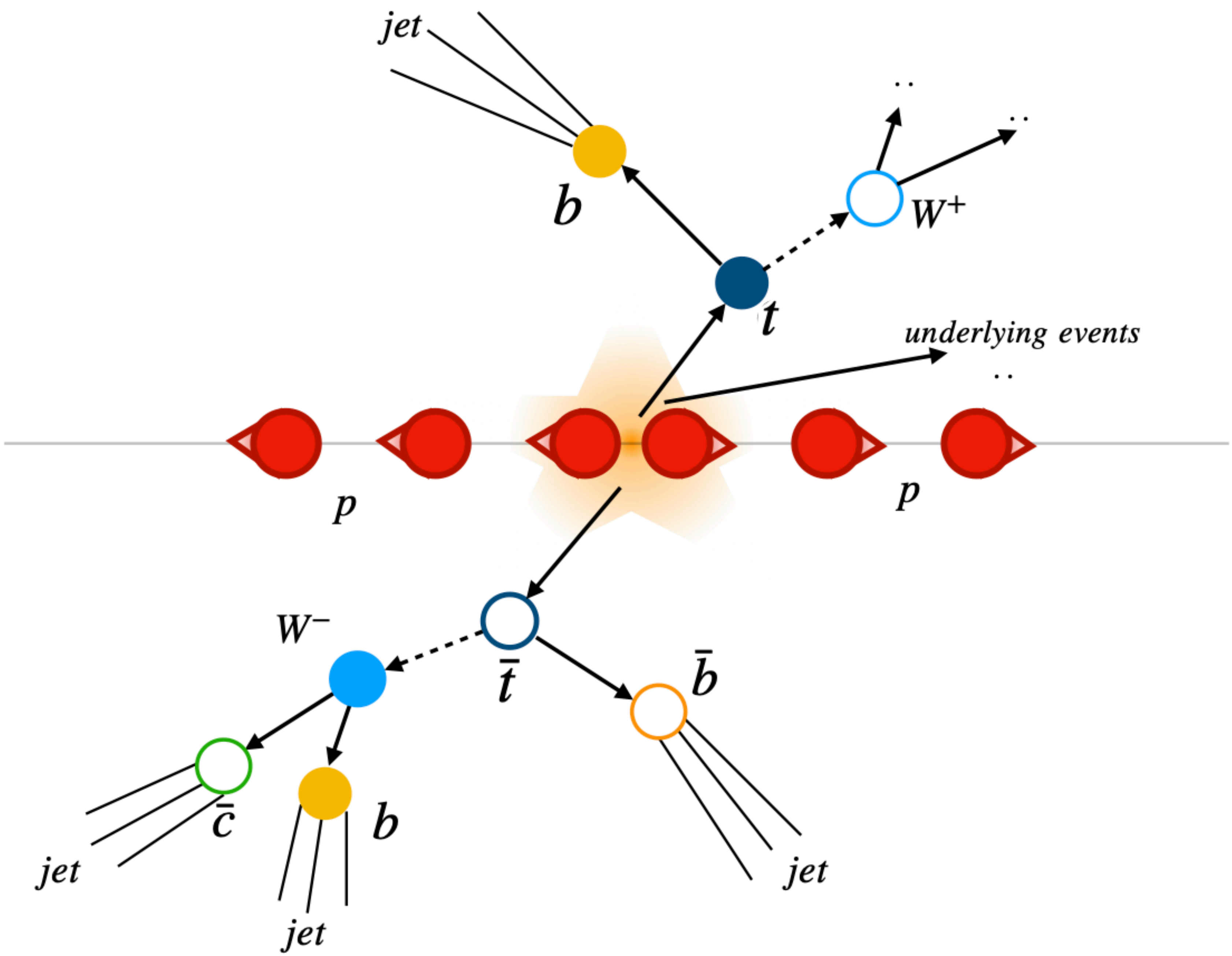


MACHINE LEARNING WITH TREE TENSOR NETWORKS



P-P SCATTERING

Typical event in LHC



Quarks

$$q \begin{cases} u \\ d \\ s \end{cases}$$

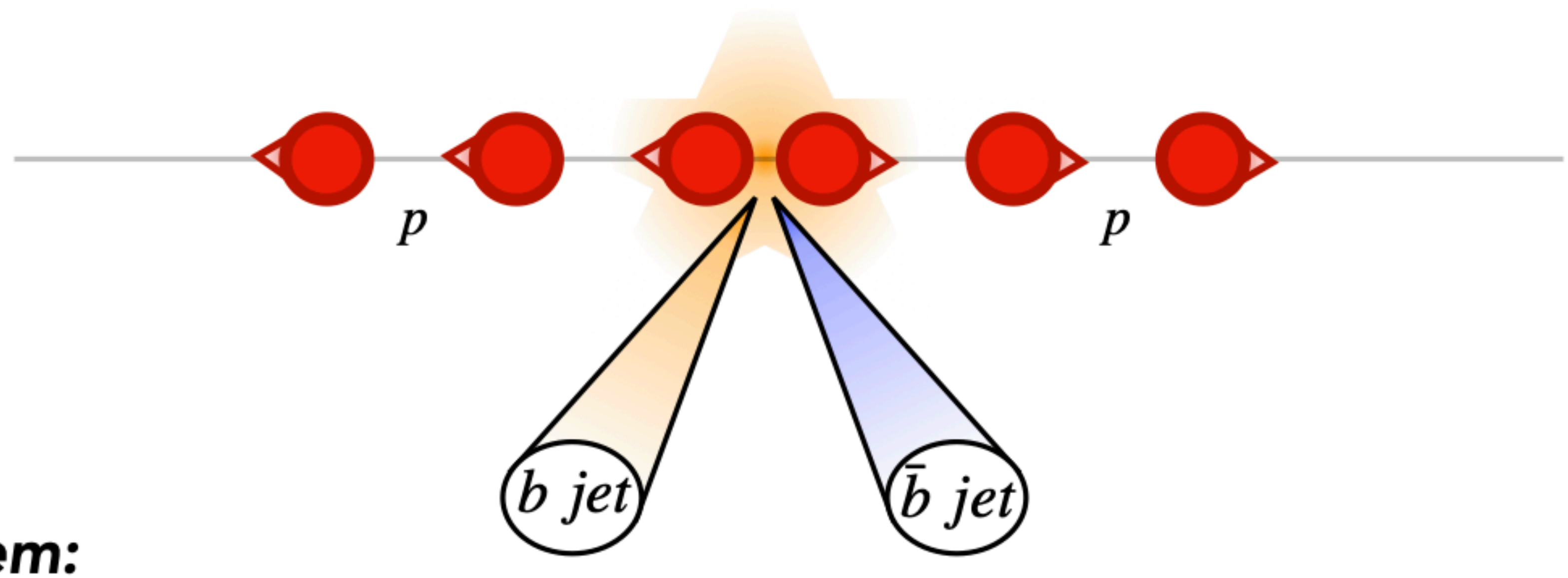
c

b

t

BINARY B BBAR CLASSIFICATION

This kind of events are used to measure **asymmetries** between the charge of b and $\bar{b}$.



Easier problem:

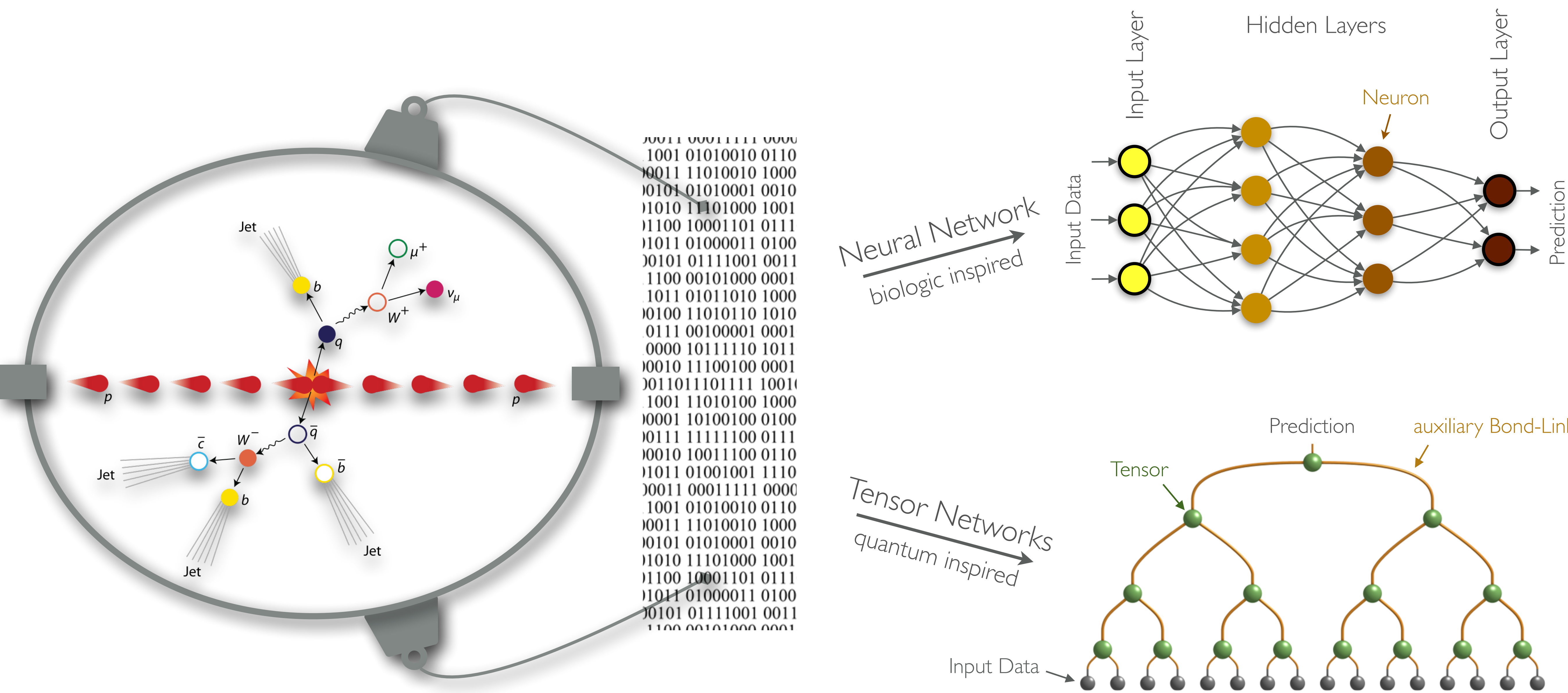
- ***16 selected features (most physically relevant)***
- ***~10^6 data samples***

μ **k** **π** **e** **p**

charge **momenta** **direction**

MACHINE LEARNING BASED CLASSIFICATION

arXiv: 2004.13747

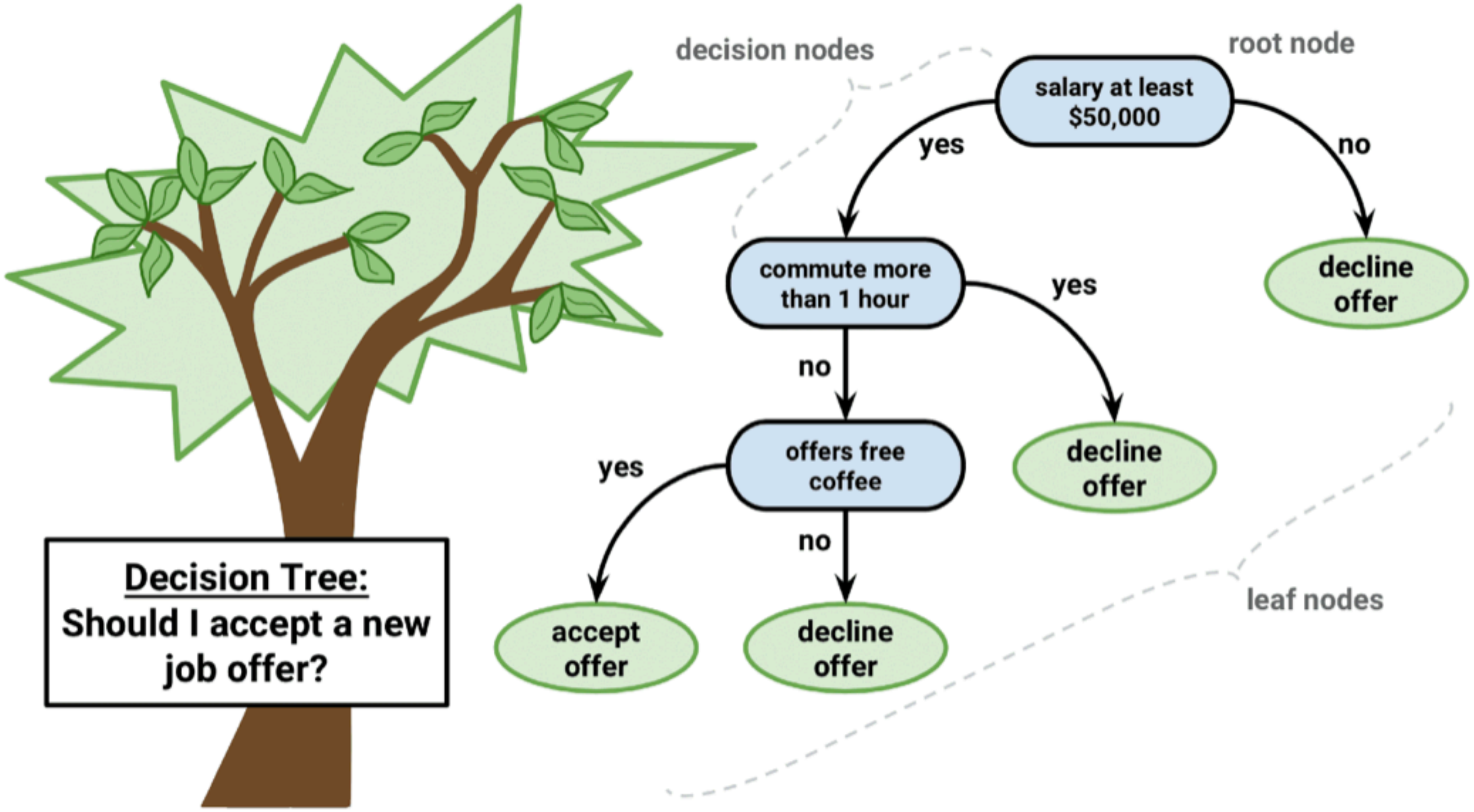


T. Felser et al. Npj quantum inf. (2021)

in collaborato with L. Sestini, A. Gianelle, D. Zuliani, D. Lucchesi

BINARY CLASSIFICATION

Until now, Boosted Decision trees:

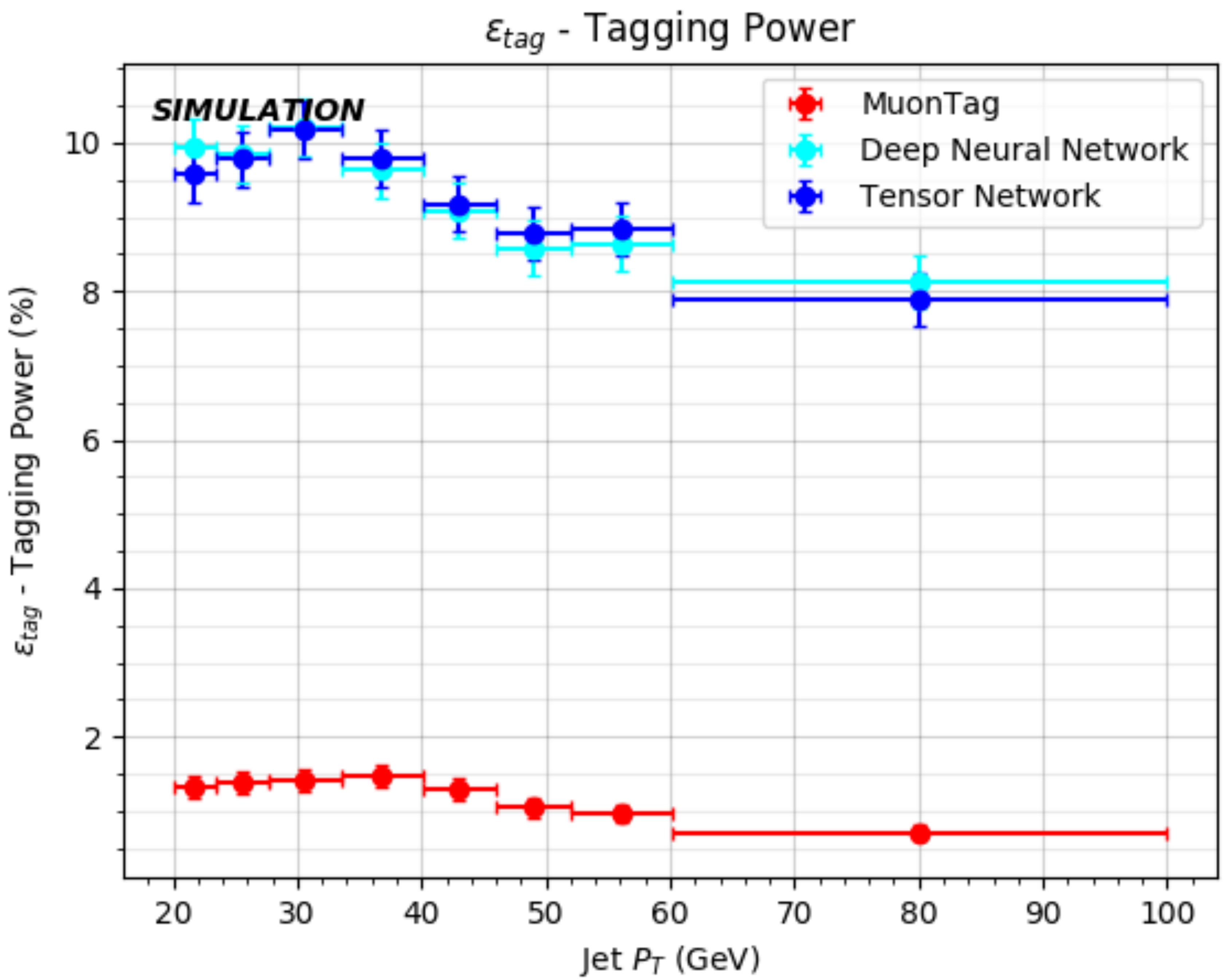


.. giving only a 6% of identification efficiency on processes like H -> $c\bar{c}$.

Article:

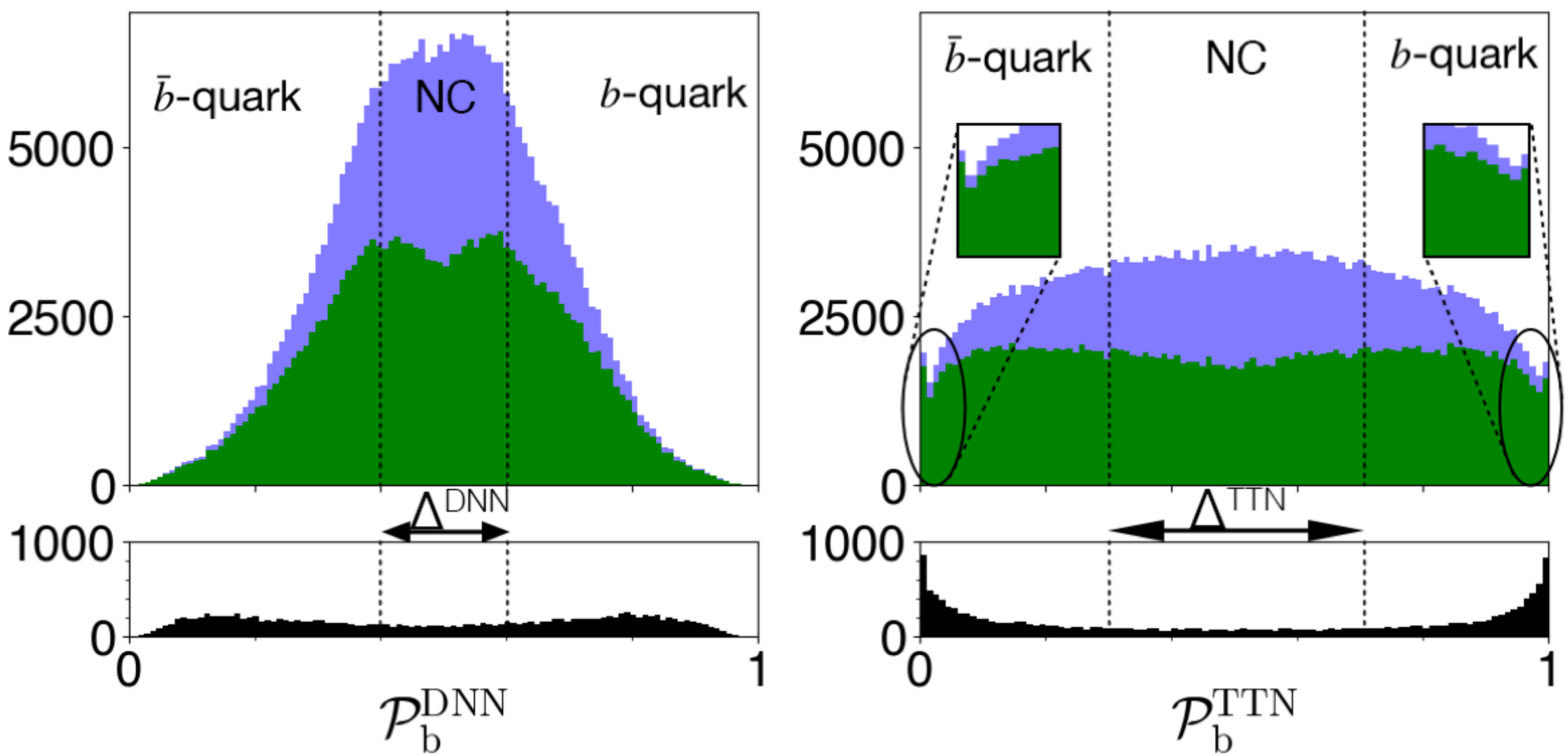
Identification of beauty and charm quark jets at LHCb, The LHCb collaboration

LHCB SIMULATED DATA ANALYSIS



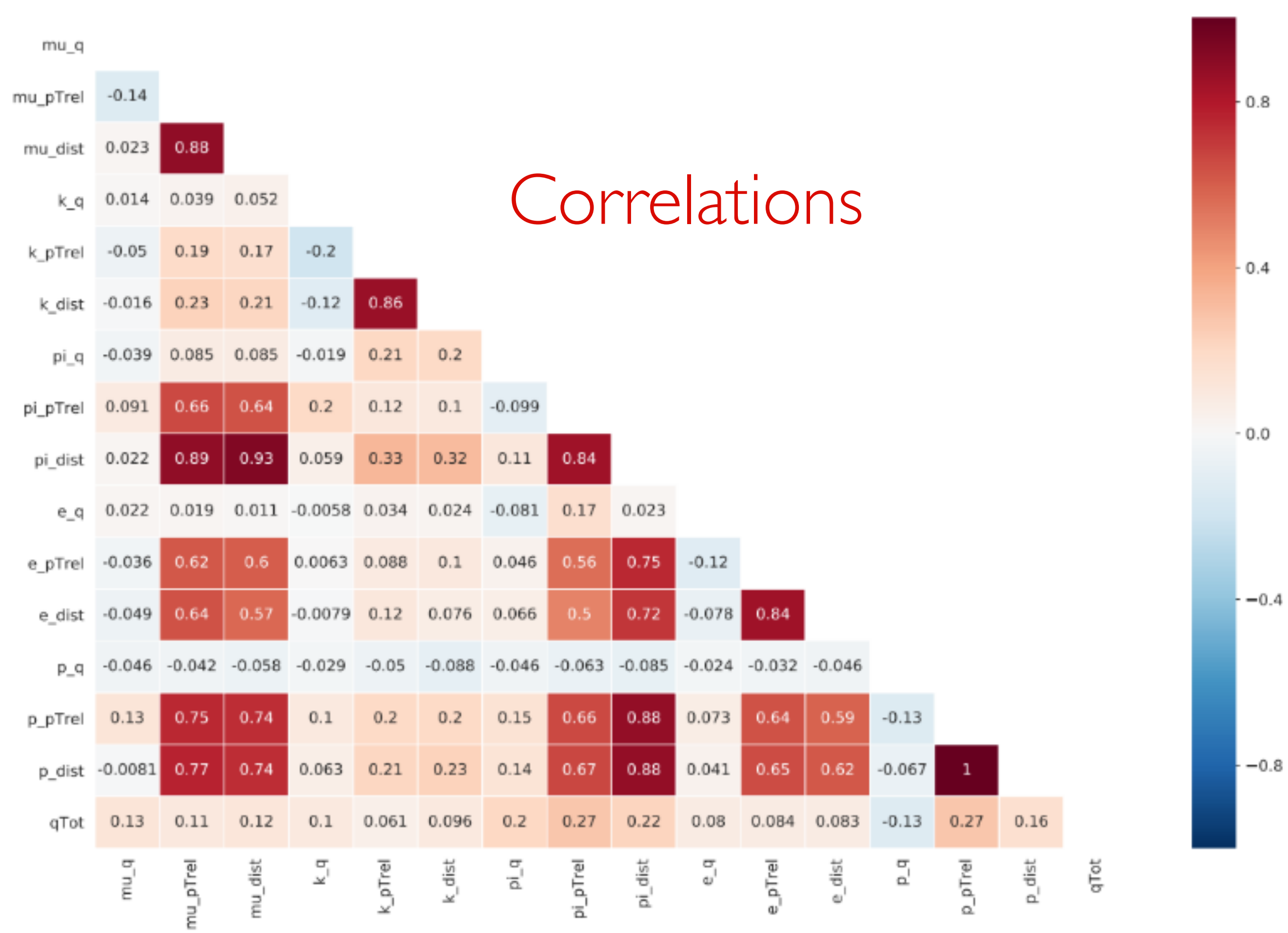
CLASSIFICATION

correctly classified

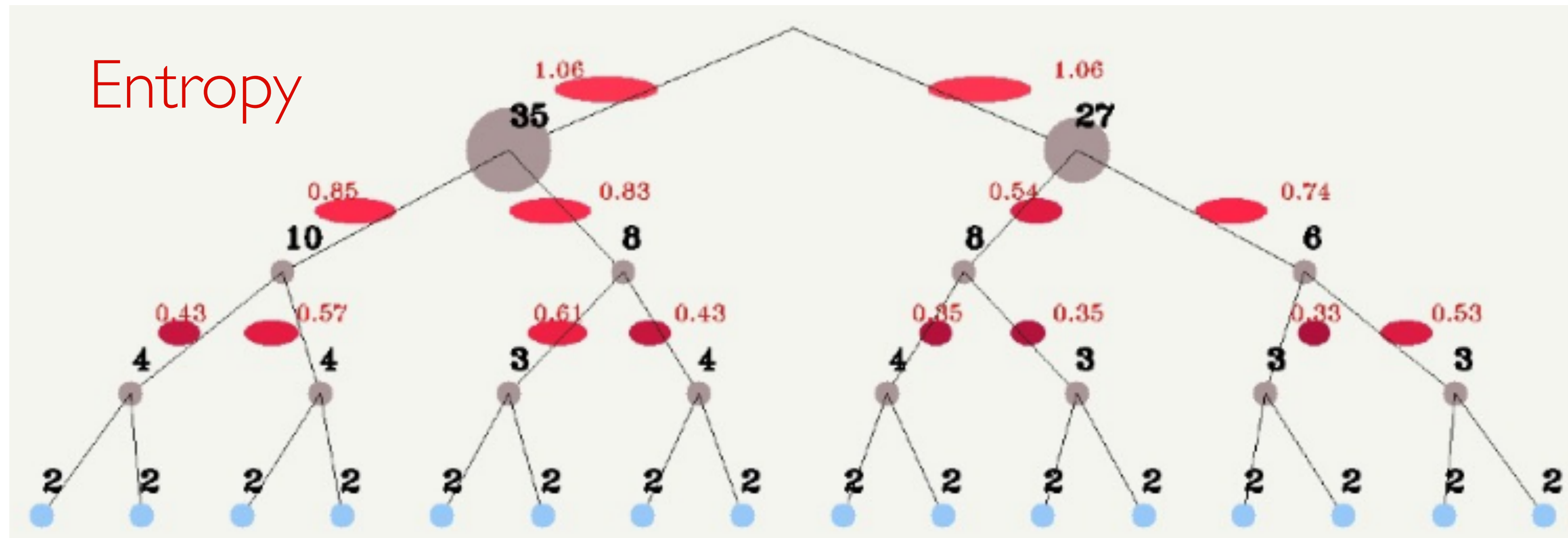


muon

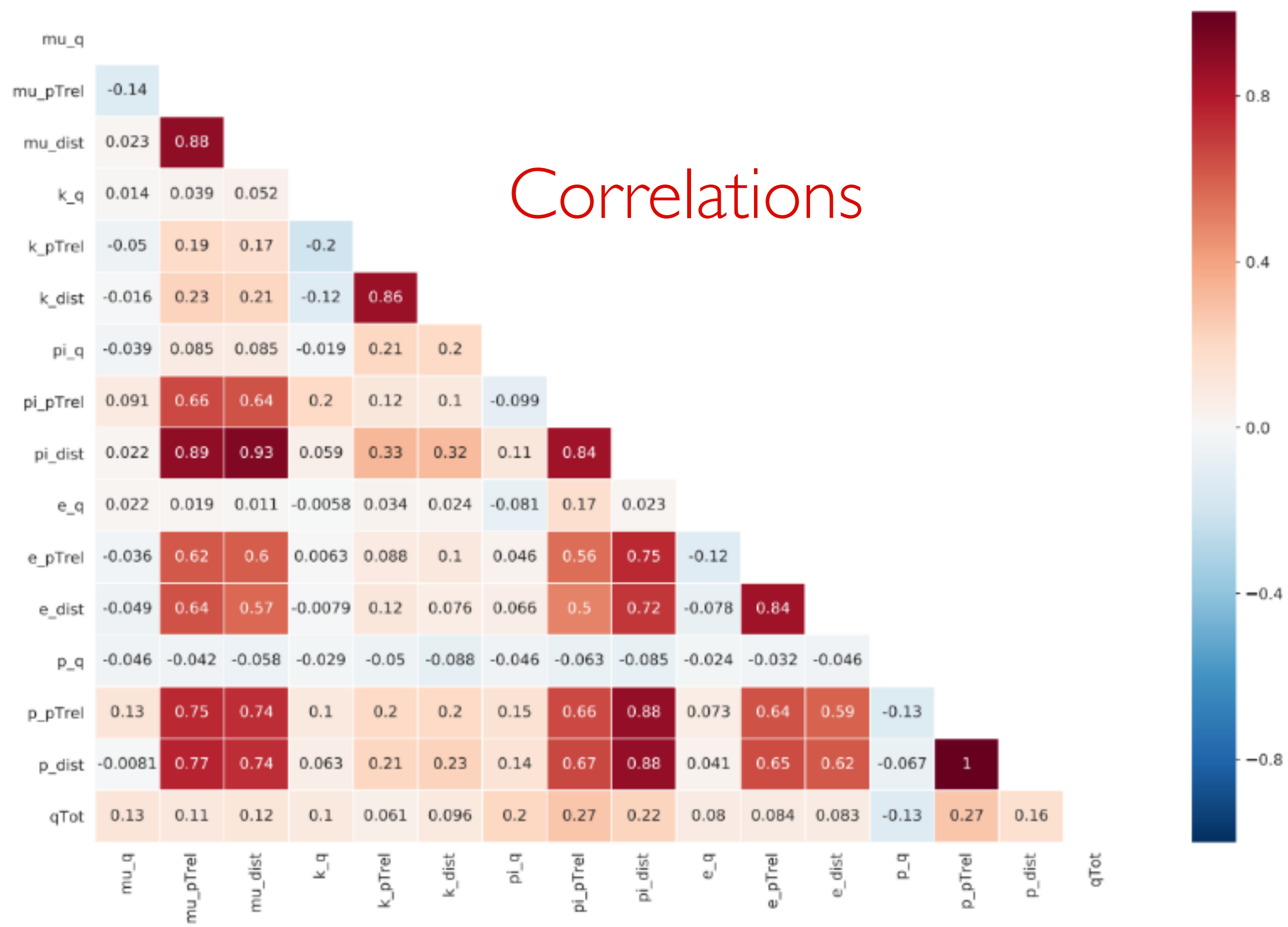
CORRELATIONS



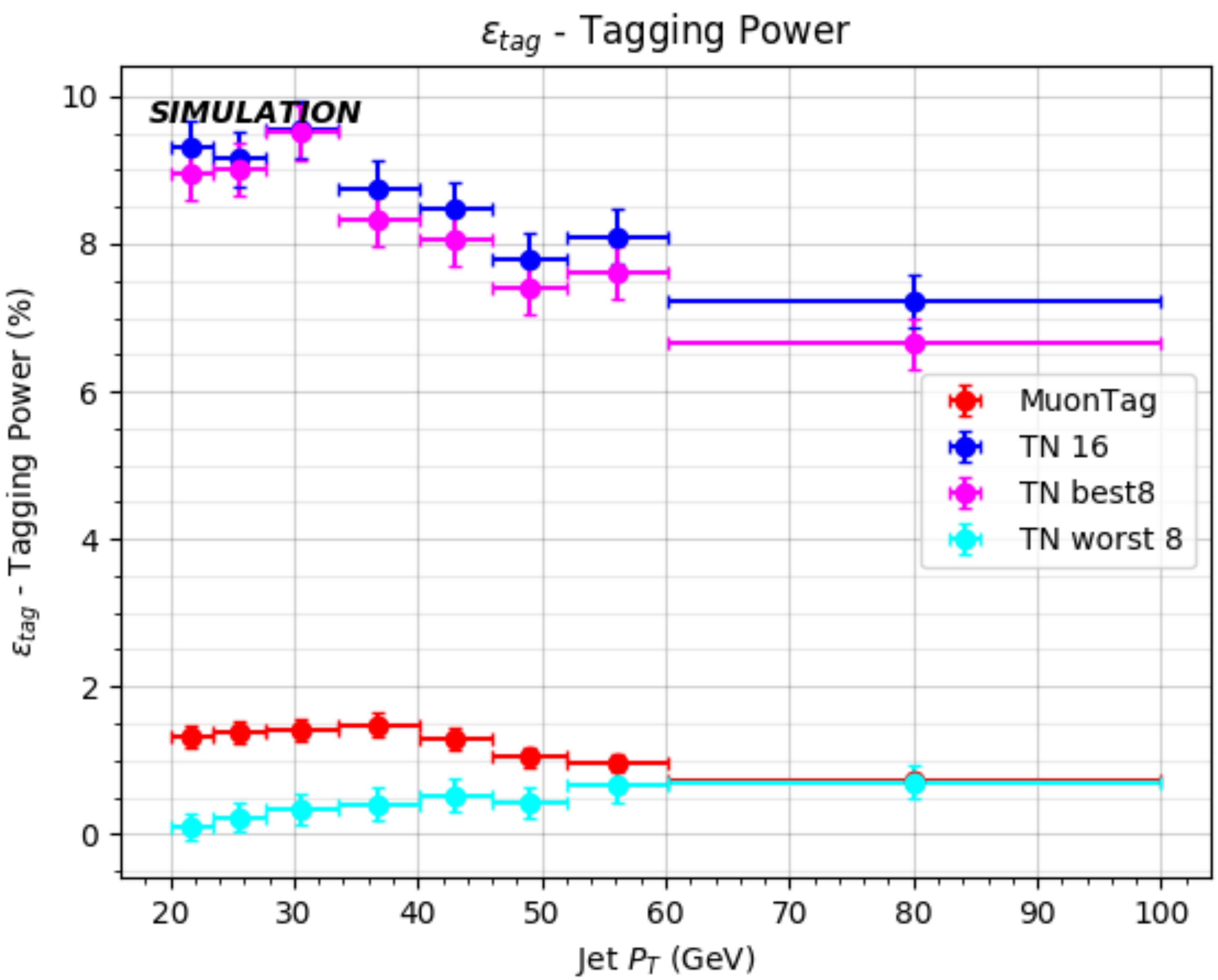
CORRELATIONS



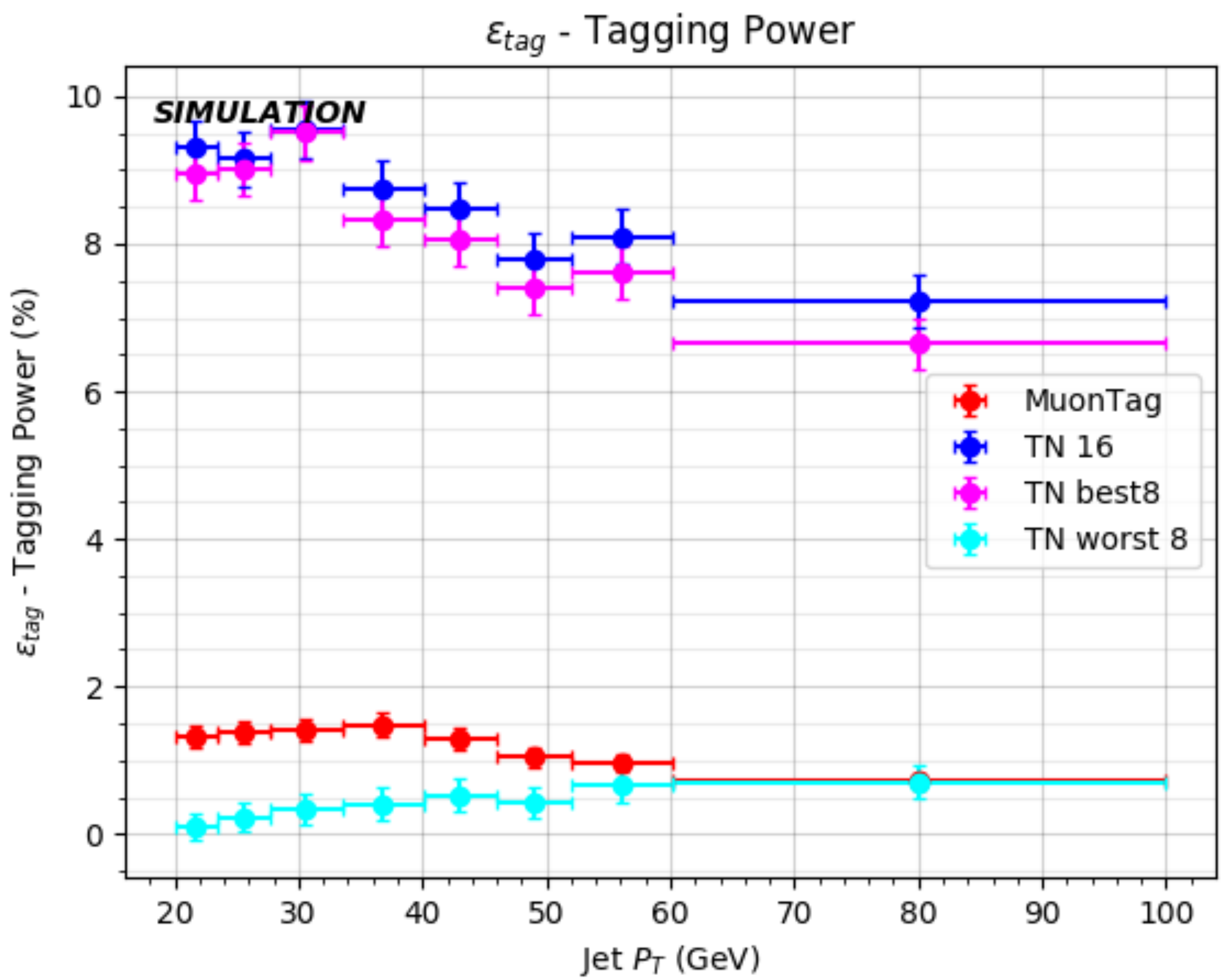
Correlations



FINAL RESULT



FINAL RESULT



- 18 times faster
- Only 0.8% less precise

FINAL RESULT

χ	Model M_{16} (incl. all 16 features) Prediction time	Accuracy	Free parameters	Model B_8 (best 8 features determined by QuIPS) Prediction time	Accuracy	Free parameters
200	345 μs	70.27 % (63.45 %)	51501	-	-	-
100	178 μs	70.34 % (63.47 %)	25968	-	-	-
50	105 μs	70.26 % (63.47 %)	13214	-	-	-
20	62 μs	70.31 % (63.46 %)	5576	-	-	-
16	-	-	-	19 μs	69.10 % (62.78 %)	264
10	40 μs	70.36 % (63.44 %)	1311	19 μs	69.01 % (62.78 %)	171
5	37 μs	69.84 % (62.01 %)	303	19 μs	69.05 % (62.76 %)	95

COMPARISON WITH MACHINE LEARNING

$$\psi_w(\mathbf{s}) = \prod_i \cosh\left(b_i + \sum_j w_{ij}s_j\right)$$

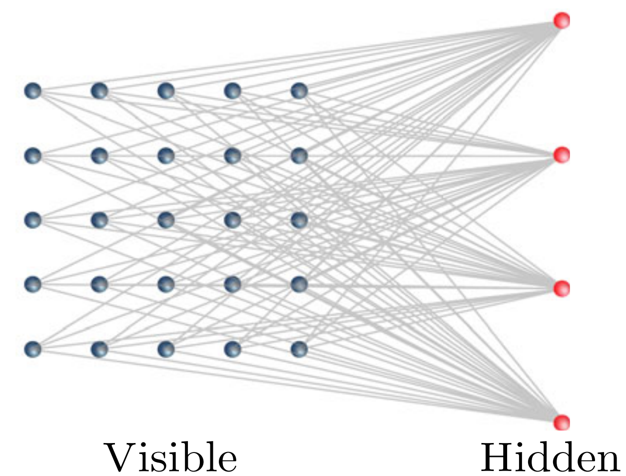
$$\propto \prod_i (e^{b_i+\sum_j w_{ij}s_j} + e^{-b_i-\sum_j w_{ij}s_j})$$

$$\propto \prod_i \mathrm{Tr}\begin{pmatrix} e^{b_i+\sum_j w_{ij}s_j} & 0 \\ 0 & e^{-b_i-\sum_j w_{ij}s_j} \end{pmatrix}$$

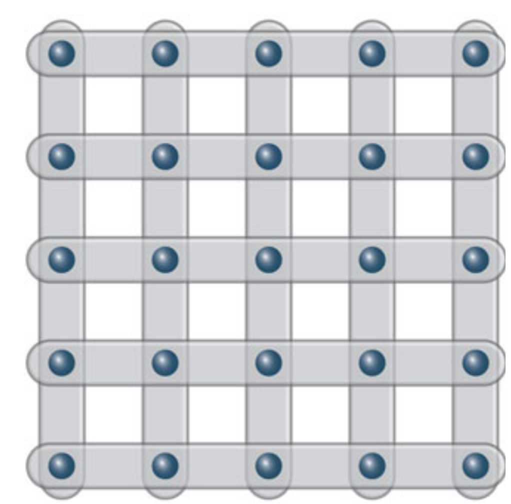
$$\propto \prod_i \mathrm{Tr}\left(\prod_{j\in i} A_{i,j}^{s_j}\right),$$

where

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



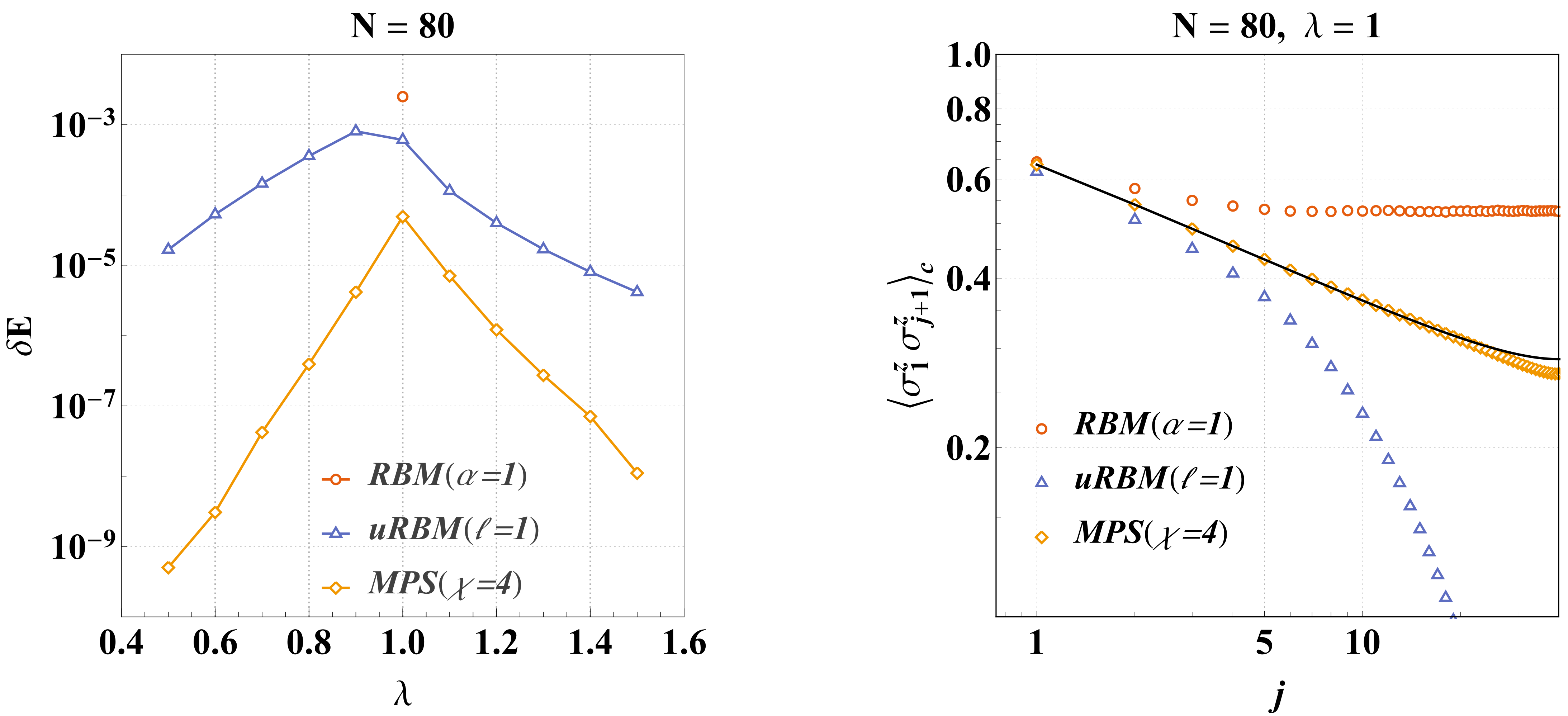
(b) Restricted Boltzmann machine in 2D



(d) SBS

Ivan Glasser et al PRX (2018)

COMPARISON WITH MACHINE LEARNING



G. Carleo, M. Troyer Science (2017)

M. Collura et al., SciPost Phys. Core (2021)

TAKE HOME MESSAGES

- Tensor network algorithms can be used to benchmark, verify, support and guide quantum simulations/computations
- High-dimensional tensor network simulations are becoming available
- Scalability to HPC is necessary to produce relevant results
- Interaction with HEP is becoming more and more relevant
- Interesting developments also in other directions (classical optimisers/annealers)
- Tensor network machine learning is competitive with DNN

quantum.dfa.unipd.it

qtech.unipd.it

Thank you for your attention!

Simone Montangero
Pietro Silvi
Giuseppe Magnifico
Simone Notarnicola
Daniel Jaschke
Luca Arceci
Timo Felser
Phila Rembold
Marco Rossignolo
Marco Rigobello
Samuele Cavinato
Giovanni Cataldi
Marco Ballarin
Alice Pagano

Quantum Flagship
QuantERA
INFN Istituto Nazionale di Fisica Nucleare
DFG Deutsche Forschungsgemeinschaft
QuSCo
PASQuanS
MIUR
QTFLAG
QuantHEP
Fondazione Cassa di Risparmio di Padova e Rovigo
Google
CINECA
bwGRiD

quantum.dfa.unipd.it

qtech.unipd.it