Data-driven discovery of relevant information in many-body problems

Roberto Verdel Aranda
ICTP, Trieste

arXiv:2307.10040
arXiv:2308.13636

"Many-Body quantum physics with machine learning"
ECT*, Trento
08/09/2023

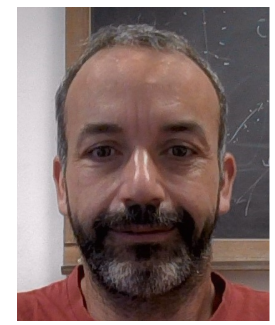
The Abdus Salam International Centre for Theoretical Physics

ECT*
EUROPEAN CENTRE FOR THEORETICAL STUDIES IN NUCLEAR PHYSICS AND RELATED AREAS

Collaborators

R. K. Panda
(ICTP/SISSA)

V. Vitale
(U. Grenoble Alps)



A. Rodriguez
(UniTS)

M. Dalmonte
(ICTP/SISSA)

S. Pedrielli
(UniTS -> TU Berlin)

E. Donkor
(ICTP/SISSA)

H. Sun
(QMU London)

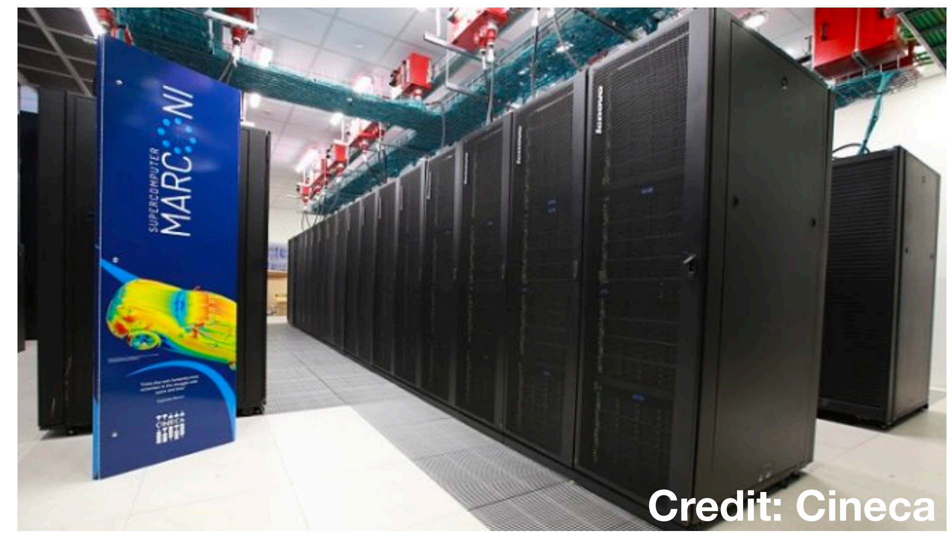
G. Bianconi
(QMU London)

M. Oberthaler's
group
(U. Heidelberg)

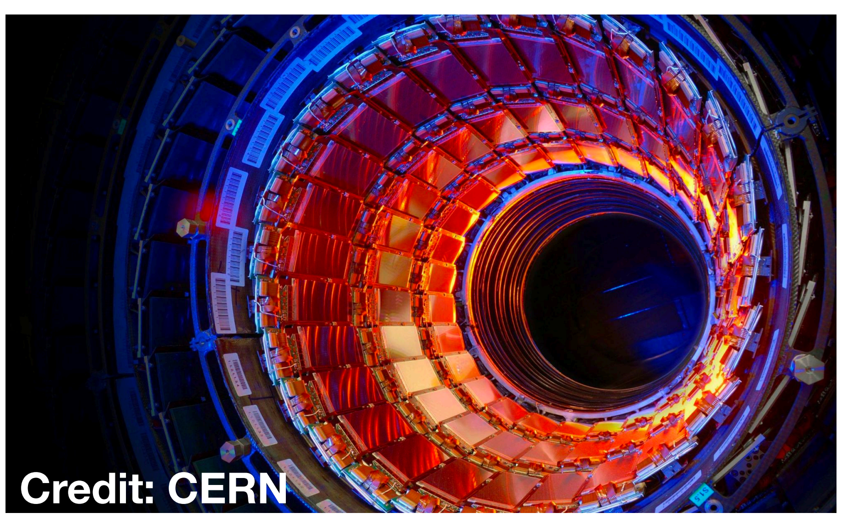
Motivation I: Physics in the age of big data

- Astrophysical observations

- Large-scale classical simulations



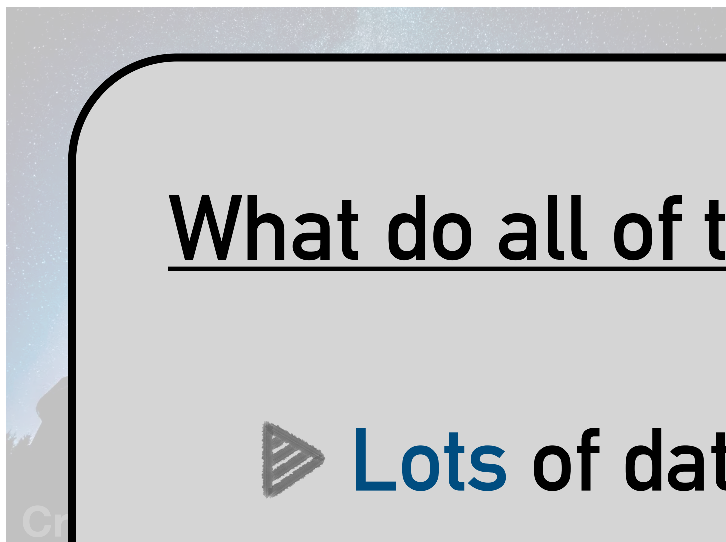
- Particle physics experiments



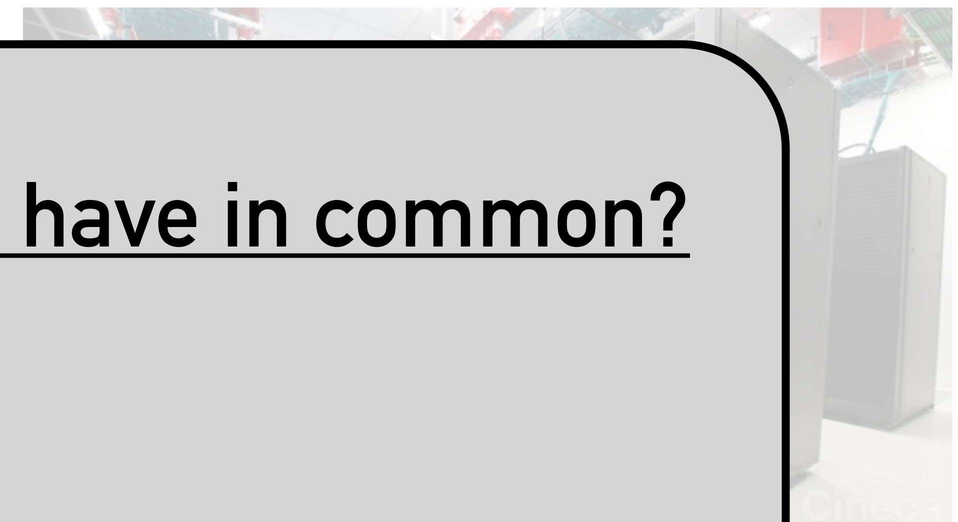
- Quantum simulation

Motivation I: Physics in the age of big data

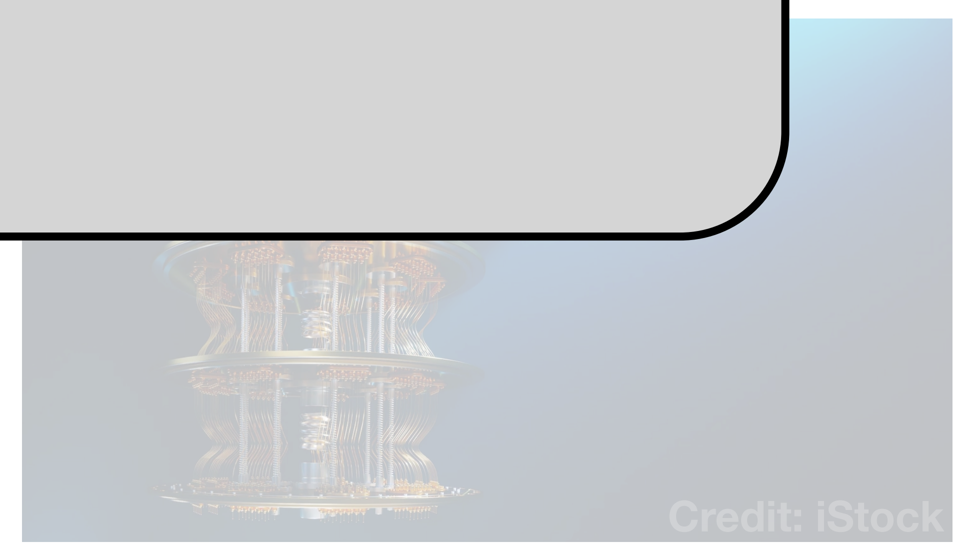
- Astrophysical observations



- Large-scale classical simulations



- Particle



<u>What do all of these fields have in common?</u>

- Lots of data available
- Data mining/ML methods can enable/ facilitate discovery

Motivation II: Quantum technology era

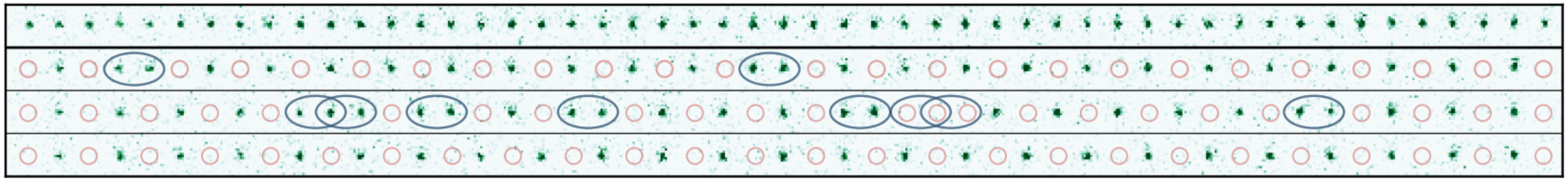
Present-day synthetic quantum devices offer an unparalleled way to probe correlated quantum matter

- Coherent dynamics with control at the individual quantum level
- Capable to produce "wave function snapshots"

Motivation II: Quantum technology era

Present-day synthetic quantum devices offer an unparalleled way to probe correlated quantum matter

- Coherent dynamics with control at the individual quantum level
- Capable to produce "wave function snapshots"



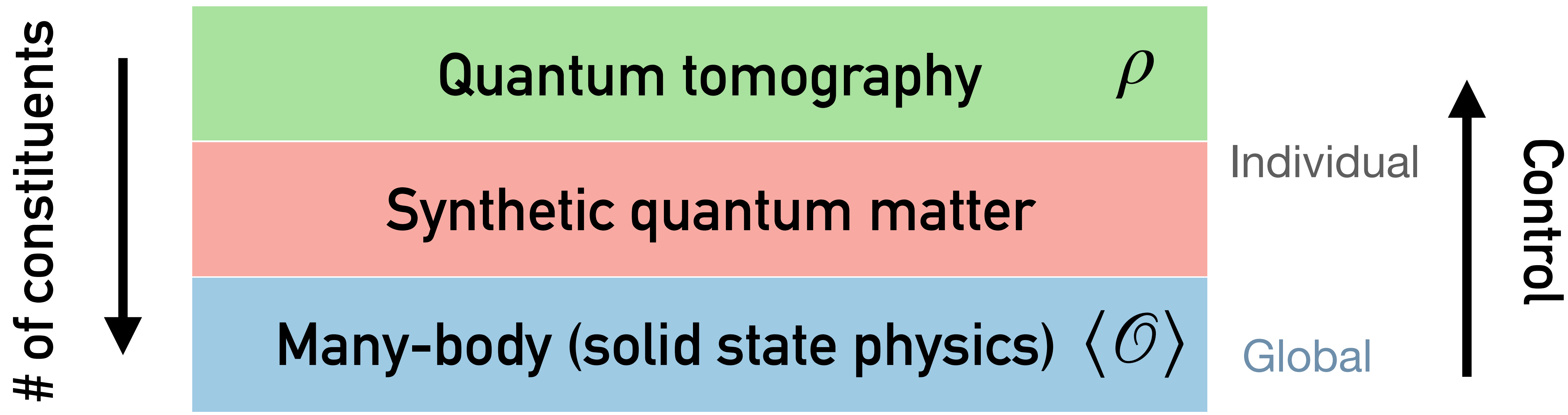
Bernien et al., Nature '17

$$\vec{X}_i = (0,1,0,1,0,1,\ldots) \qquad \mathbf{M} = \{\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_{N_r}\}$$

Motivation II: Quantum technology era

Present-day synthetic quantum devices offer an unparalleled way to probe correlated quantum matter

- Coherent dynamics with control at the individual quantum level
- Capable to produce "wave function snapshots"



Motivation II: Quantum technology era

Present-day synthetic quantum devices offer an unparalleled way to probe correlated quantum matter

How can we deal with the full information content of many-body snapshots provided by synthetic quantum systems?

# of constituents

Synthetic quantum matter

Many-body theory $\langle \mathcal{O} \rangle$

Control

A couple of remarks

- Output of quantum simulations are classical objects. Hence data-wise equivalent to output of classical simulations (the techniques I will discuss today can also be applied to classical simulations).
- We work with limited sampling: $N_s \ll 2^N$.
- Complementary to other techniques such as classical shadows, randomised measurements, etc.

How do we extract relevant information from many-body snapshots?

"Traditional" approaches (stat mech / effective field theory):
compute few-point correlators, for instance:

$$C_{ij}^{(2)} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$$

Allows us to characterize classical/quantum phase transitions, determine "proper vertices" of the quantum effective action, etc.

However, it disregards part of the information content of many-body snapshots	In data science jargon: an "uncontrolled" dimensional reduction

Why we would like to go beyond?

- Unbiased identification the relevant degrees of freedom at play in strongly interacting theories

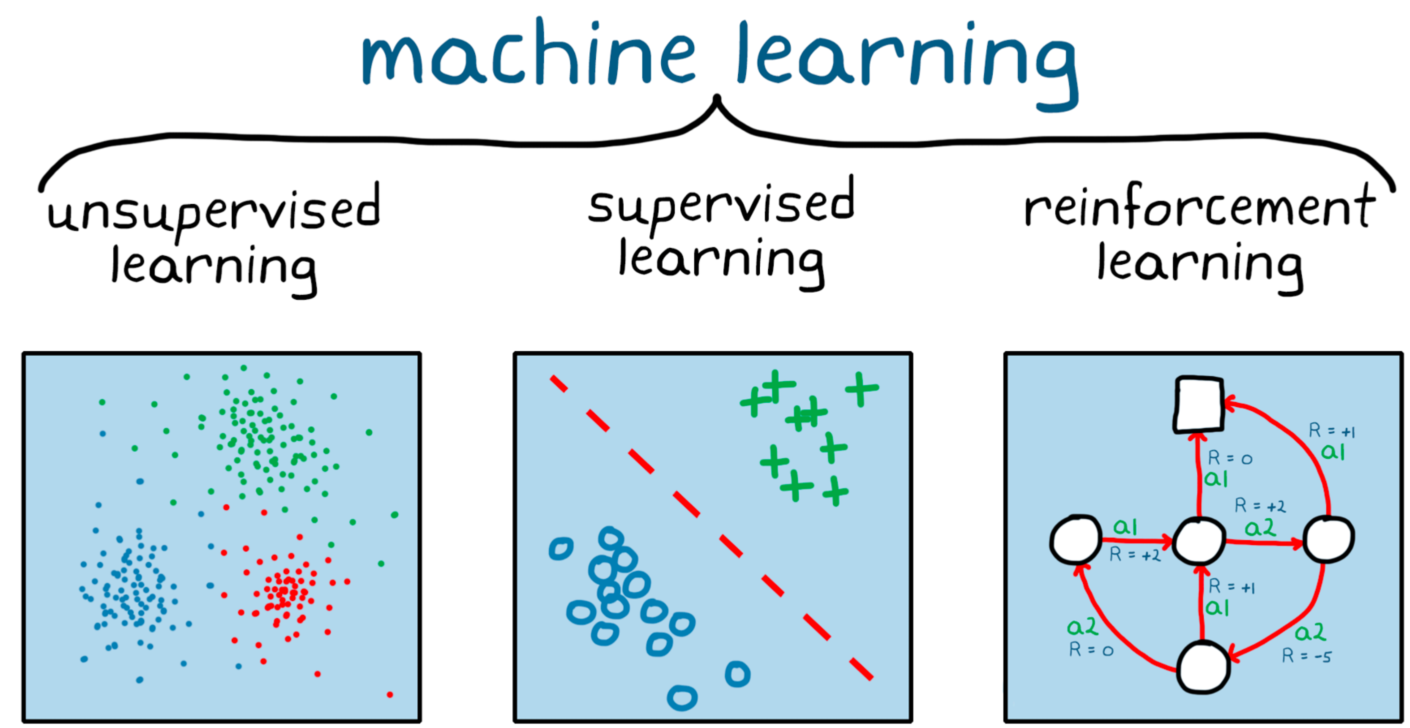
- Understanding the working of quantum computers (e.g. choosing best suited measurement basis, cross-platform verification, noise tomography, etc.)

- Quantifying the complexity of wave functions

- Detect and characterise systems with non-local correlations (topological phases of matter)

Data-driven strategy

Use non-parametric unsupervised approaches to discover and extract relevant information in many-body physics problems by leveraging all available information

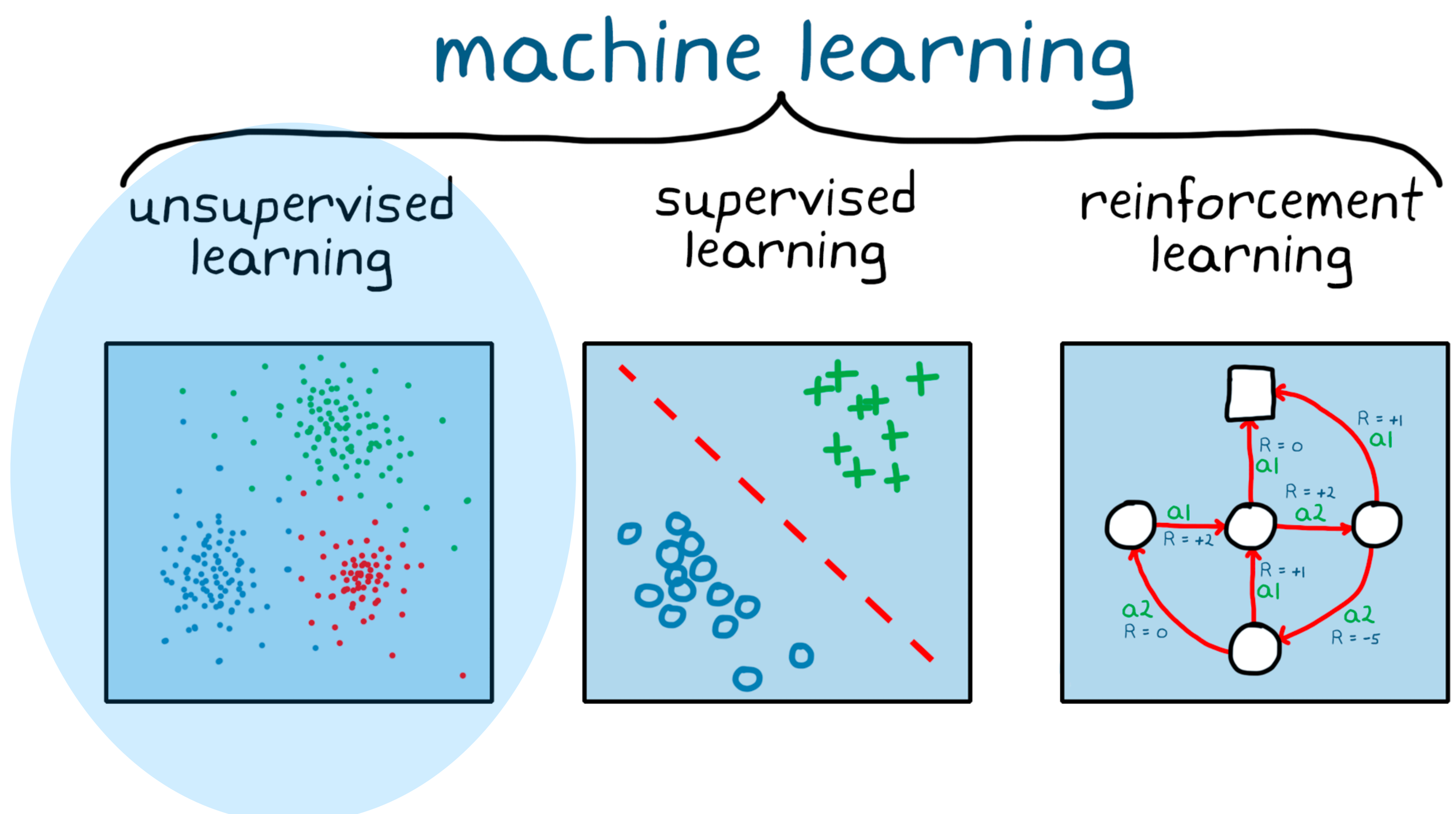


https://www.mathworks.com/discovery/reinforcement-learning.html

Data-driven strategy

Use non-parametric unsupervised approaches to discover and extract relevant information in many-body physics problems by leveraging all available information

Think of methods related to dimensional reduction, feature selection, etc.



https://www.mathworks.com/discovery/reinforcement-learning.html

Technical outline

- Learning critical behaviour in Ising partition functions
  - Intrinsic dimension
  - PCA-based entropy

arXiv:2308.13636

- Relevant information discovery in quantum simulators
  - Ranking operators via PCA entropy and information imbalance*
  - Complexity of universal dynamics far from equilibrium

arXiv:2307.10040

Similar approaches

Stat mech / lattice field theory:

Hu et al. PRE '17
Wetzel PRE '17
Wang & Zhai, PRB '17
Ch'ng et al., PRE '18
Mendes-Santos et al., PRX '21
Sale et al., PRE '22; PRD '23
Sehayek & Melko, PRB '22
Spitz, et al., PRD '23
Vitale et al., arXiv '23

Quantum many-body:

Rodriguez-Nieva & Scheurer, Nat Phys '19
Lidiak & Gong, PRL '20
Mendes-Santos et al., PRX Quantum '21
Bohrdt et al., PRL '21
Spitz, et al., SciPost Phys '21
Tirelli & Costa, PRB '21
Schmitt & Lenarčič, PRB '22
Miles et al., PRR '23
Mendes-Santos et al., arXiv '23

… and many more!

Learning critical behaviour in Ising partition functions

Rajat Panda

arXiv:2308.13636; see also arXiv 2308.13604

Intrinsic dimension

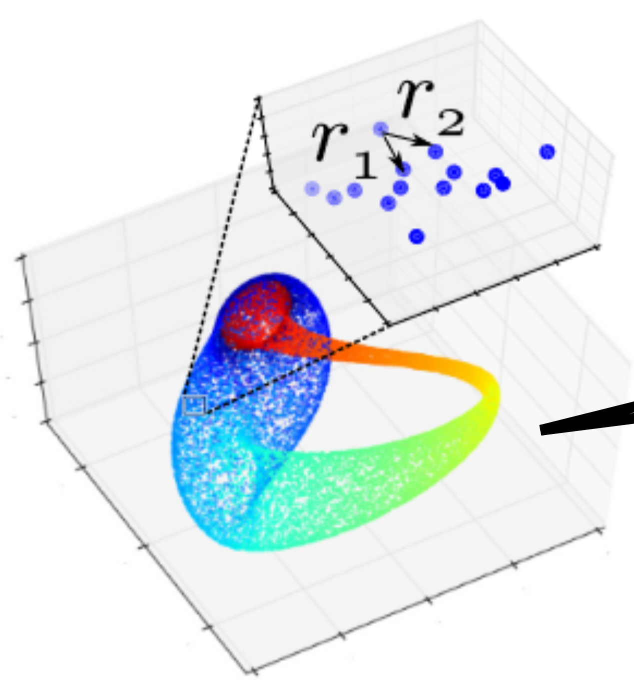
- Basic tool in data mining with multiple applications in chemical and bimolecular science and image analysis

Glielmo et al., Chem. Rev. ‘21

- Quantifies the minimum number of variables needed to describe the data

- Serves as a proxy of the Kolmogorov complexity

$D = 3$



$I_d = 2$

Each point is, for example, a spin configuration

Mendes-Santos et al., PRX ‘21

Intrinsic dimension: TWO-NN

Facco et al., Sci. Rep. '17

Uses statistics of distances between nearest-neighbor (NN) points

Needs a metric (e.g. for spin systems: Hamming distance)

$$d(i,j) := \sum_r | \vec{S}_r^i - \vec{S}_r^j |$$

Intrinsic dimension: TWO-NN

Facco et al., Sci. Rep. '17

Uses statistics of distances between nearest-neighbor (NN) points

Needs a metric (e.g. for spin systems: Hamming distance)

$$d(i,j) := \sum_r |\vec{S}_r^i - \vec{S}_r^j|$$

Example: 3-site system

$$\vec{S}^1 = (0,1,1)$$

$$\vec{S}^2 = (1,1,1)$$

$$\vec{S}^3 = (1,0,1)$$

$$\vec{S}^4 = (0,0,0)$$

$$d(\vec{S}^1, \vec{S}^2) = |0-1| + |1-1| + |1-1| = 1$$

$$d(\vec{S}^1, \vec{S}^3) = 2$$

$$\ldots$$

Intrinsic dimension: TWO-NN

Facco et al., Sci. Rep. '17

Uses statistics of distances between nearest-neighbor (NN) points

Needs a metric (e.g. for spin systems: Hamming distance)

Main assumption: NN points are drawn uniformly from I_d-dim hyperspheres

For each point, compute:

$$\mu = \frac{r_2}{r_1}$$

Distribution function of μ:

$$f(\mu) = \frac{I_d}{\mu^{I_d+1}}$$

Intrinsic dimension: TWO-NN

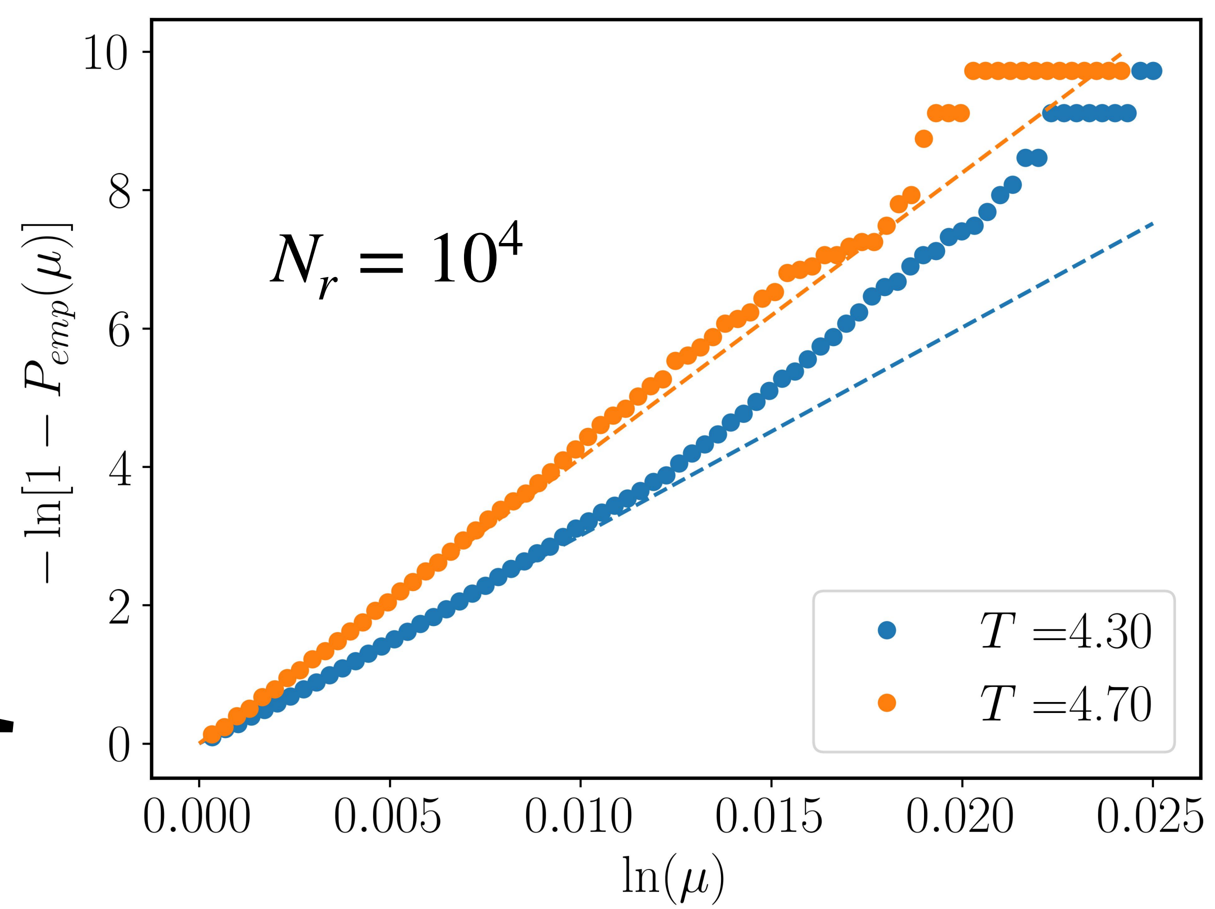
Uses statistics of distances between nearest-neighbor (NN) points

Needs a metric (e.g. for spin systems: Hamming distance)

Main assumption: NN points are drawn uniformly from I_d-dim hyperspheres

$$\mu = \frac{r_2}{r_1} \qquad f(\mu) = \frac{I_d}{\mu^{I_d+1}}$$

Linear fit using cumulative dist. function



Intrinsic dimension: toy example

Toy example: 3-site XY model

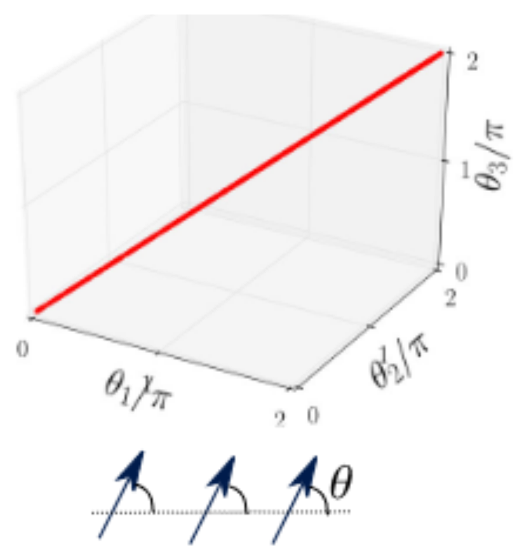
Mendes-Santos et al., PRX '21

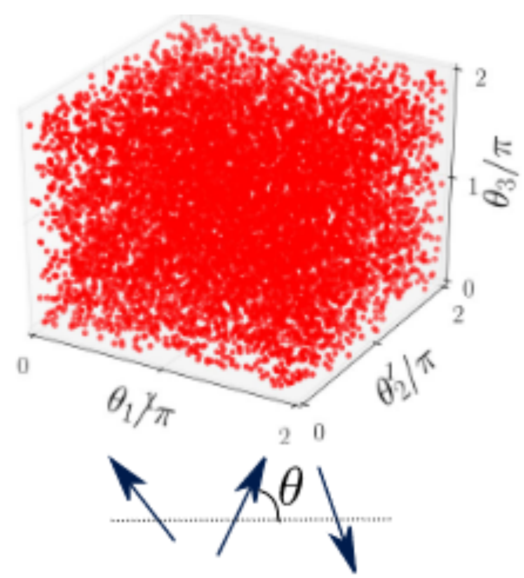
Hamiltonian: $$H = -\sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j)$$

Configurations (data points): $(\theta_1, \theta_2, \theta_3)$

Low temperature

$I_d = 1$





High temperature

$I_d = D = 3$

What about close to a transition point?

Intrinsic dimension: 2D Ising

<u>2D classical Ising model</u>

$$E = -J \sum_{\langle i,j \rangle} S_i S_j$$

Square lattice

Divergent correlation length: do data structures are more complex?

Second-order (conformal) phase transition

$$T_c = \frac{2}{\ln(1+\sqrt{2})} \approx 2.269$$

$$\nu = 1$$

Mendes-Santos et al., PRX '21

Intrinsic dimension: 2D Ising

2D classical Ising model

$$E = -J \sum_{\langle i,j \rangle} S_i S_j$$

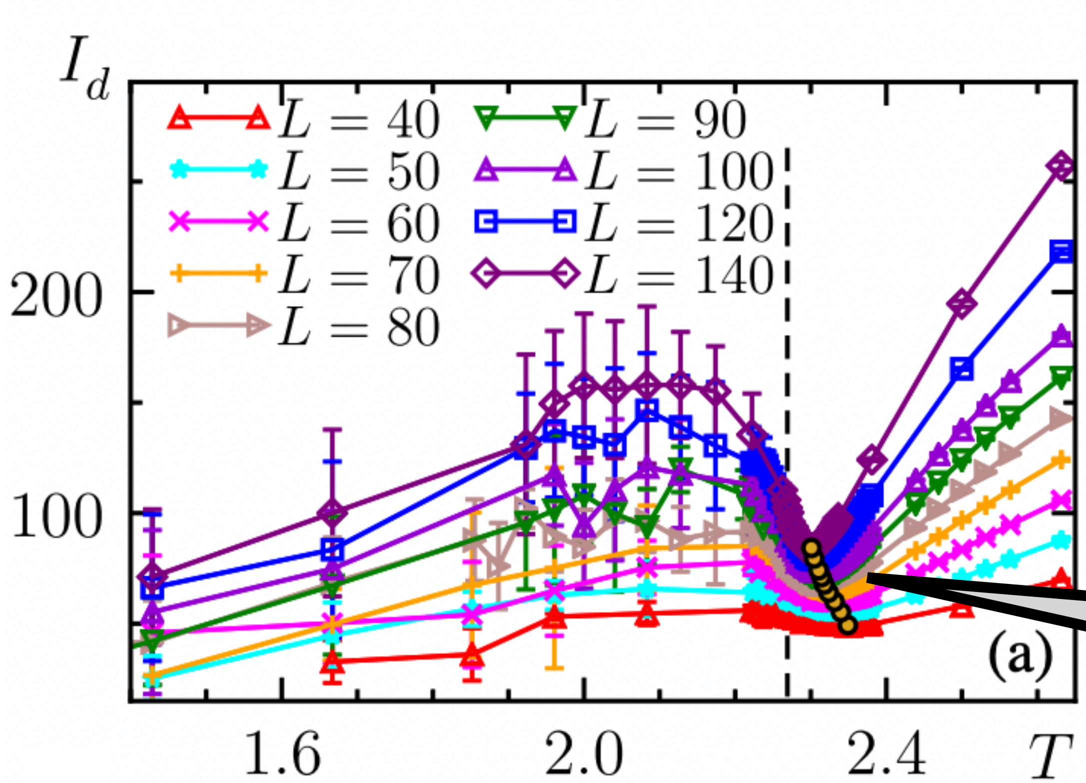
Square lattice

Second-order (conformal) phase transition

$$T_c = \frac{2}{\ln(1+\sqrt{2})} \approx 2.269$$

$$\nu = 1$$

Divergent correlation length: do data structures are more complex?



Manifold simplifies at the transition!

Intuition: universality

Mendes-Santos et al., PRX '21

Intrinsic dimension: 2D Ising

<u>2D classical Ising model</u>

$$E = -J \sum_{\langle i,j \rangle} S_i S_j$$

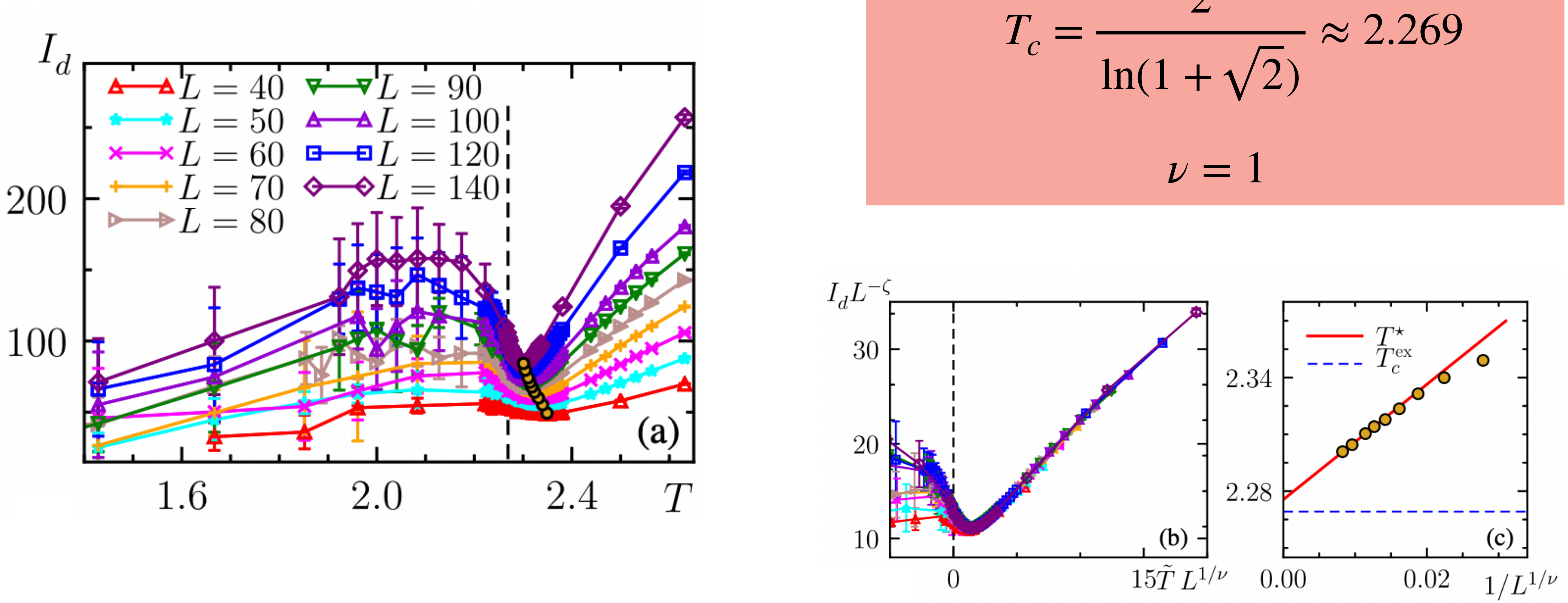
Square lattice

Divergent correlation length: do data structures are more complex?

Second-order (conformal) phase transition

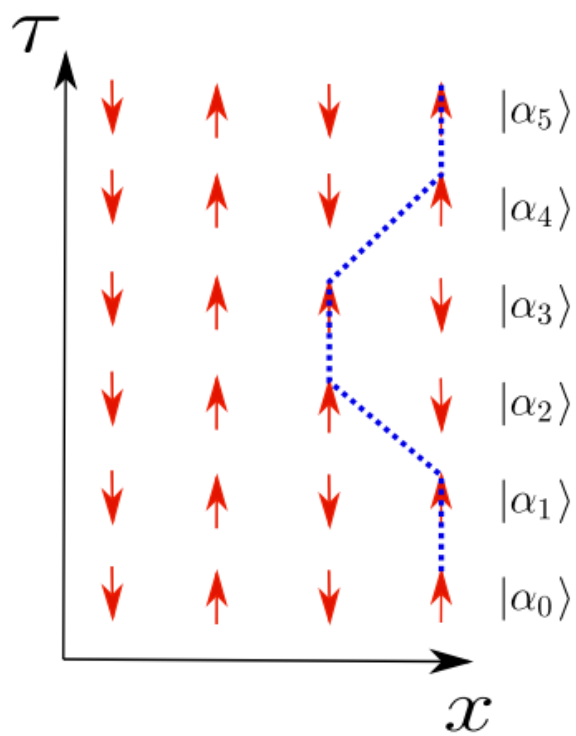
$$T_c = \frac{2}{\ln(1+\sqrt{2})} \approx 2.269$$

$$\nu = 1$$

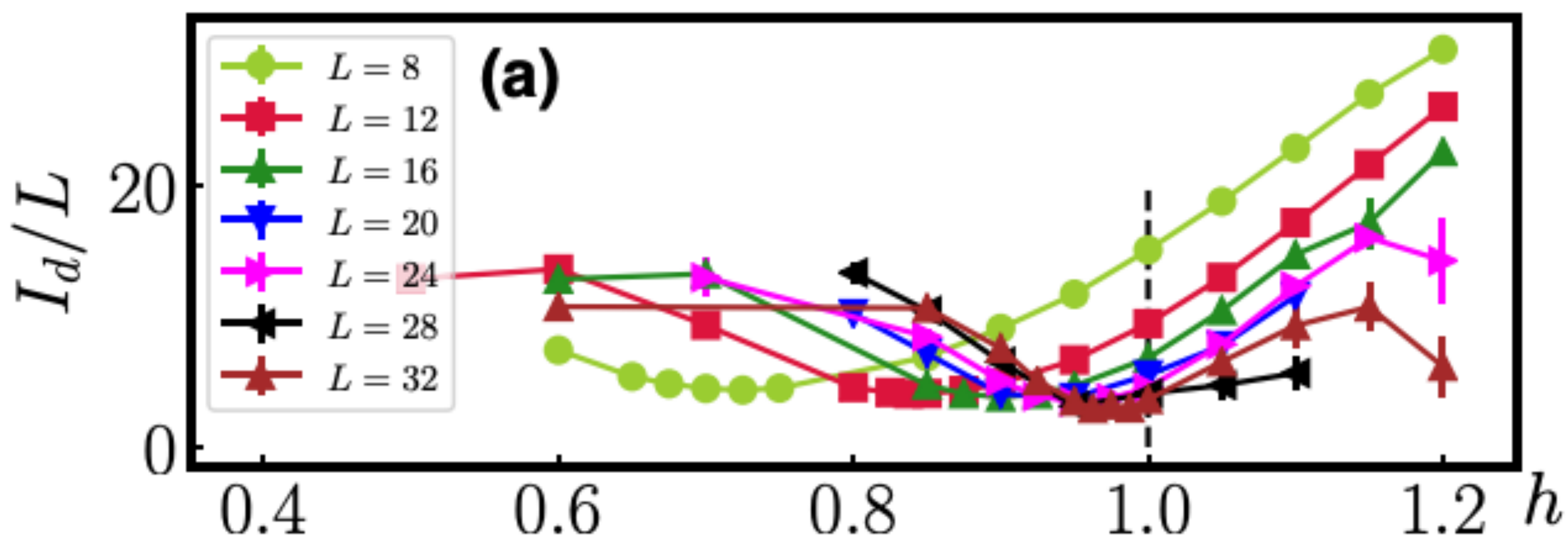


Mendes-Santos et al., PRX '21

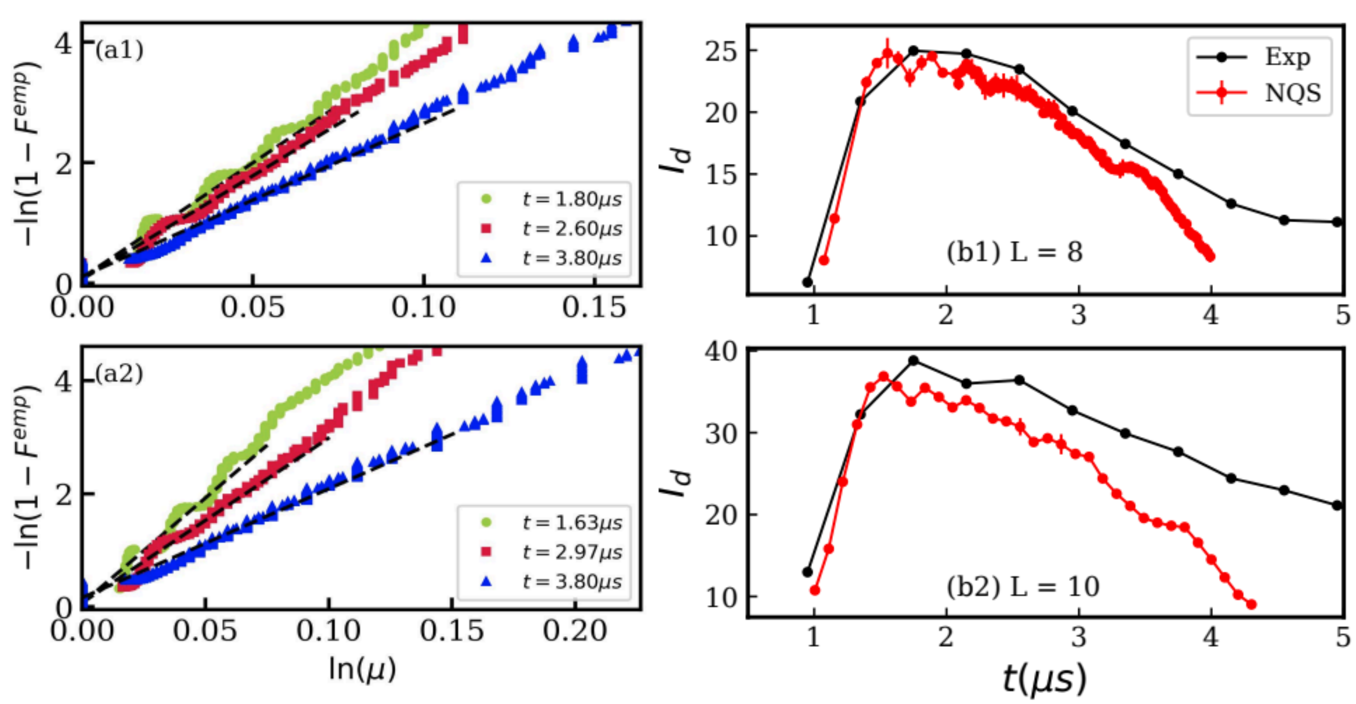
Intrinsic dimension in quantum systems



Quantum Ising chain



Mendes-Santos et al., PRX Quantum '21



Kibble-Zurek in a Rydberg quantum simulation

Mendes-Santos, Schmitt, et al., arXiv:2301.13216

Role of the physical dimension

How does volume affects the data structure and the intrinsic dimension?

3D Ising model $$E = -J \sum_{\langle i,j \rangle} S_i S_j$$

- No analytical solution known so far
- Continuous phase transition at $T_c \approx 4.51$ (believed to be conformal)
- Dual to a $\mathbb{Z}_2$ lattice gauge theory
- QCD critical point expected to belong to the 3D Ising universality class
  [Stephanov et al., PRL '98; Gavin et al., PRD '94; …]

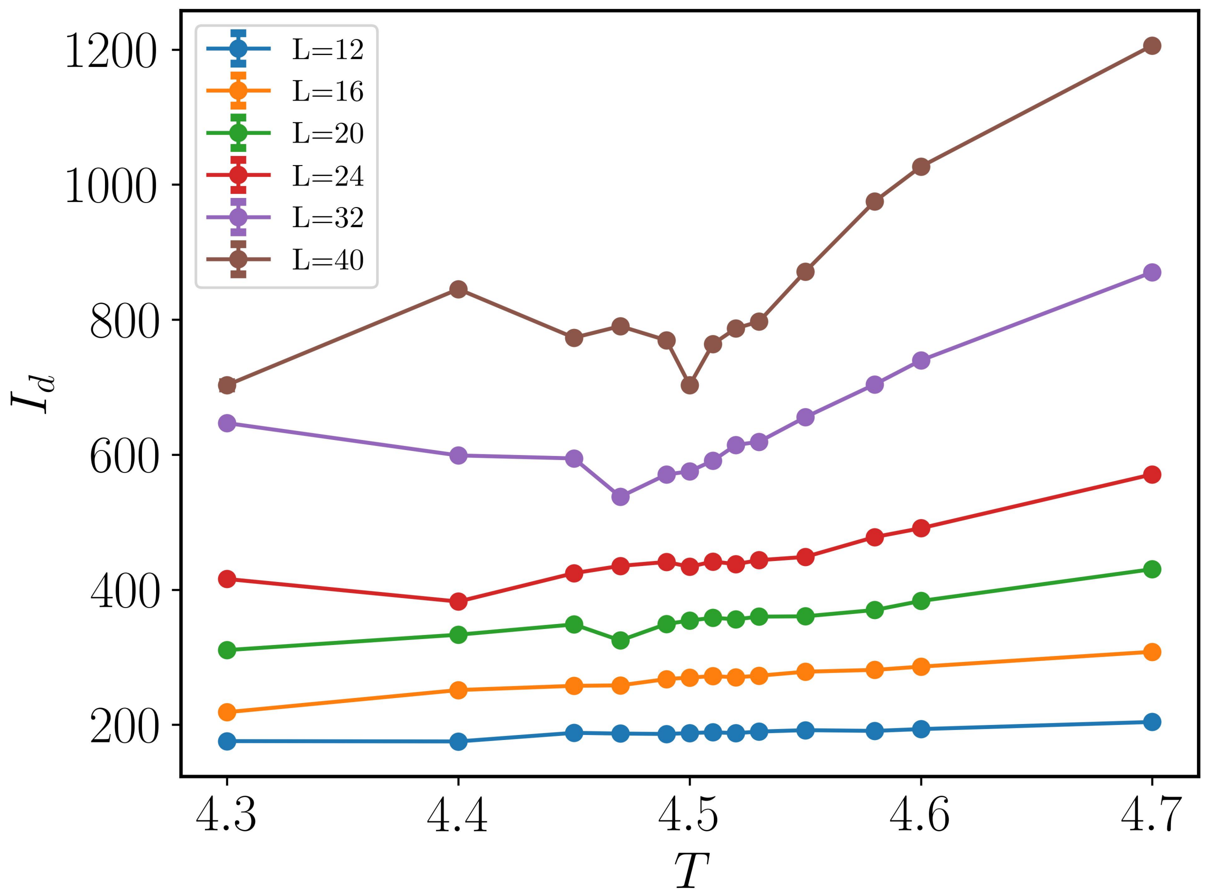
Panda, RV, et al., arXiv:2308.13636

Volume effects on I_d: TWO-NN

How does volume affects the data structure and the intrinsic dimension?

3D Ising model

$$E = -J \sum_{\langle i,j \rangle} S_i S_j$$



- Very high I_d (results must be taken warily)
- Minimum not so clear at the transition (TWO-NN estimator)
- In general, harder to extract information through I_d

Panda, RV, et al., arXiv:2308.13636

PCA entropy

Can we use complementary statistical tests to still be able to extract relevant information?

PCA entropy

Can we use complementary statistical tests to still be able to extract relevant information?

Principal Component Analysis (PCA)

Transformation of the coordinate system to find high-variance directions

It amounts to diagonalizing the covariance matrix $\mathbf{\Sigma} = \mathbf{X}^T\mathbf{X}/(N_r - 1)$:

$$\mathbf{\Sigma}\lambda_n = \lambda_n \overrightarrow{w}_n$$

See e.g. Jolliffe (2005)

PCA entropy

Can we use complementary statistical tests to still be able to extract relevant information?

Principal Component Analysis (PCA)

Transformation of the coordinate system to find high-variance directions

It amounts to diagonalizing the covariance matrix $\mathbf{\Sigma} = \mathbf{X}^T\mathbf{X}/(N_r - 1)$:

$$\mathbf{\Sigma}\lambda_n = \lambda_n \overrightarrow{w}_n$$

See e.g. Jolliffe (2005)

Normalized eigenvalues:

$$\tilde{\lambda}_n = \frac{\lambda_n}{\sum_m \lambda_m}$$

By construction: $\tilde{\lambda}_n \geq 0, \ \sum_n \tilde{\lambda}_n = 1$

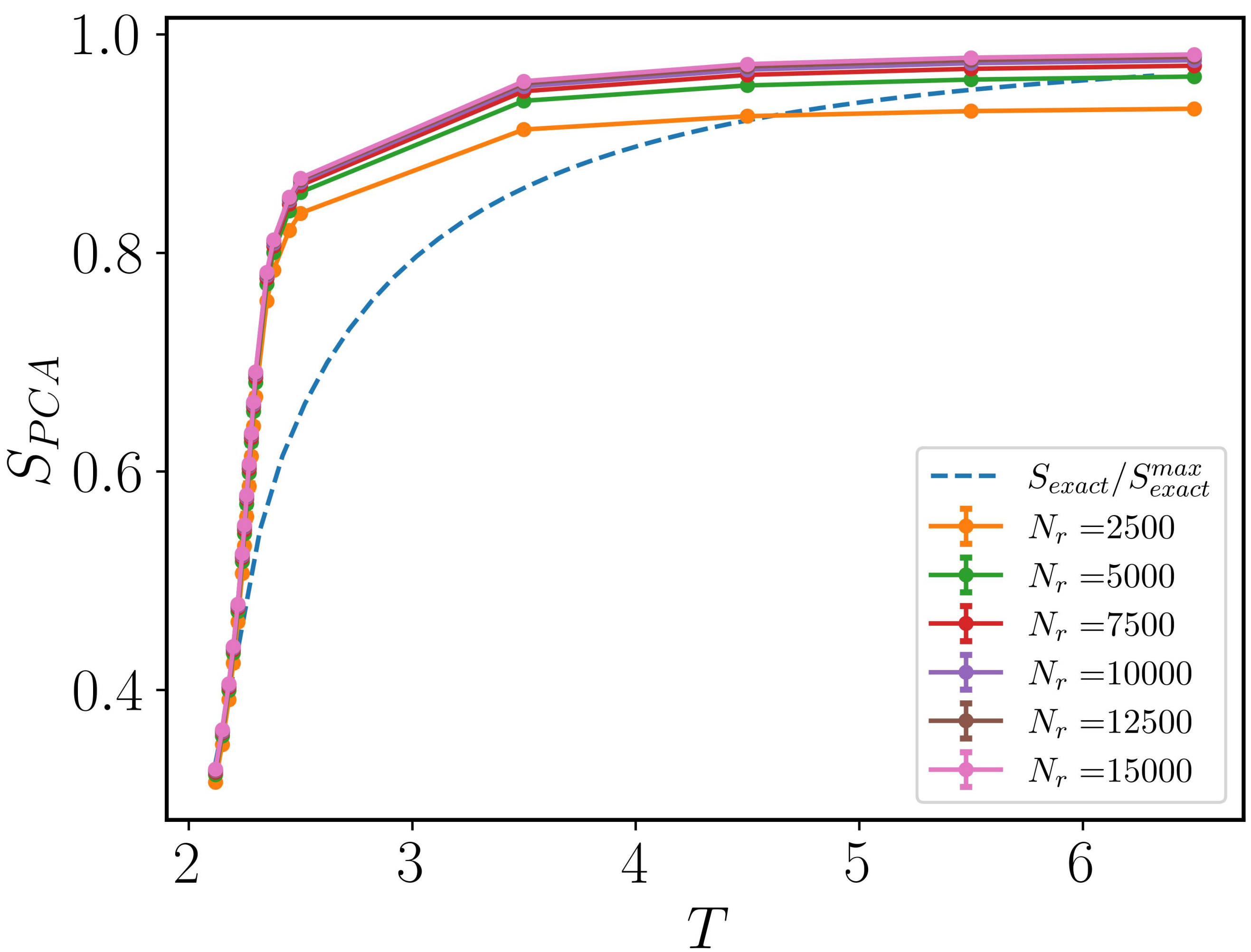
("Shannon") PCA entropy

$$S_{\mathrm{PCA}} = -\sum_n \tilde{\lambda}_n \ln(\tilde{\lambda}_n)$$

Alter et al., PNAS (2000), …

PCA entropy: 2D Ising

Striking qualitative similarity to the thermodynamic entropy!

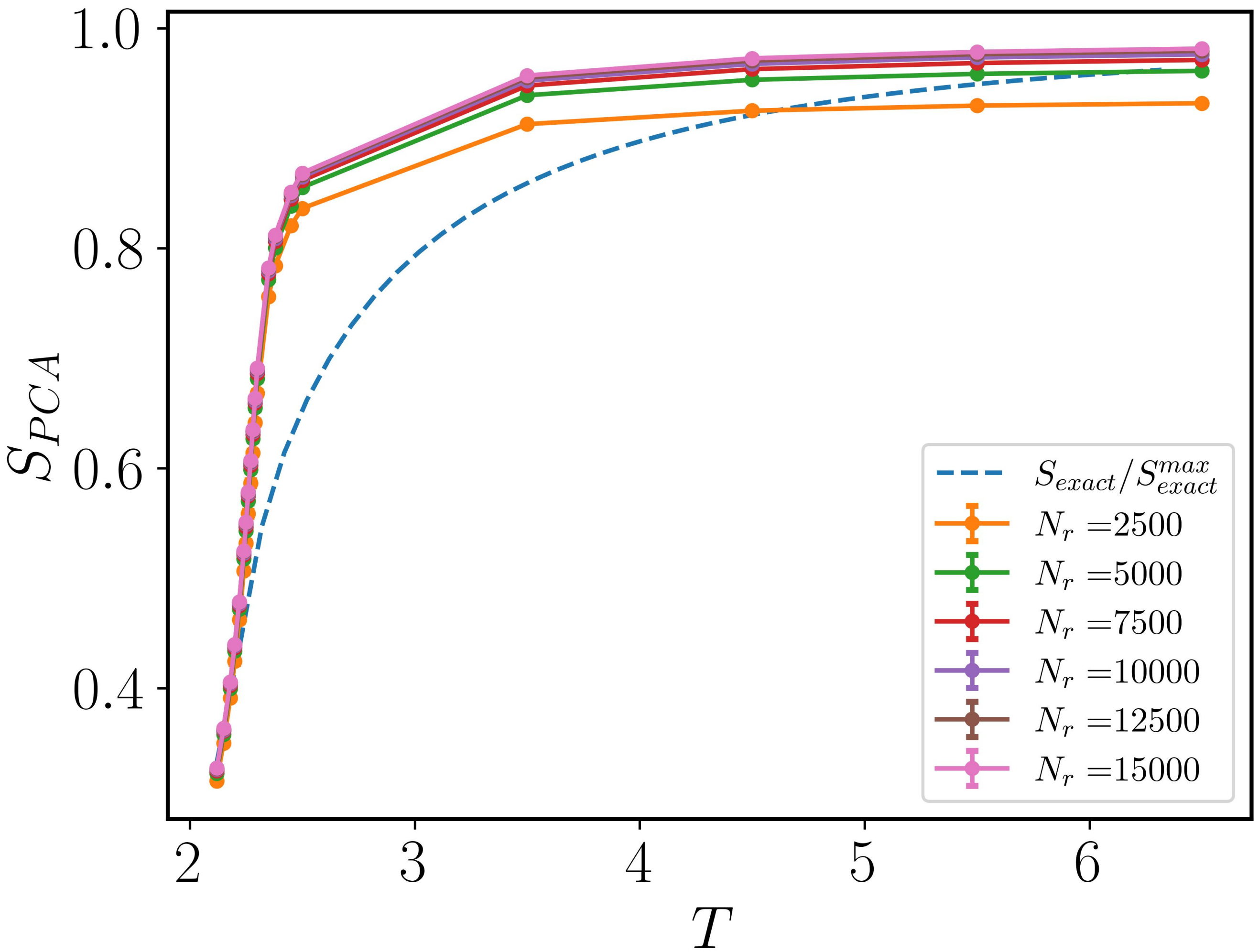


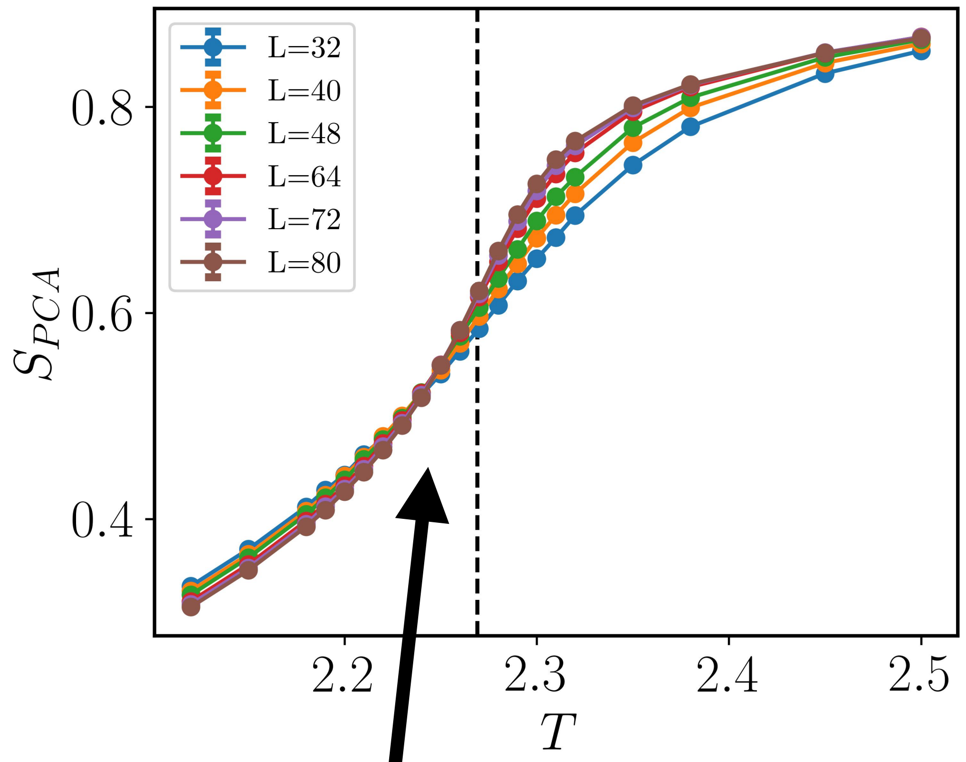
- Suggests a direct link between the thermodynamic entropy and the (easy-to-compute) PCA entropy
- See also alternative approaches using ML or compression algorithms, eg. Nir et al., PNAS '20; Avinery et al., PRL '19; etc.

Panda, RV, et al., arXiv:2308.13636

PCA entropy: 2D Ising

Striking qualitative similarity to the thermodynamic entropy!



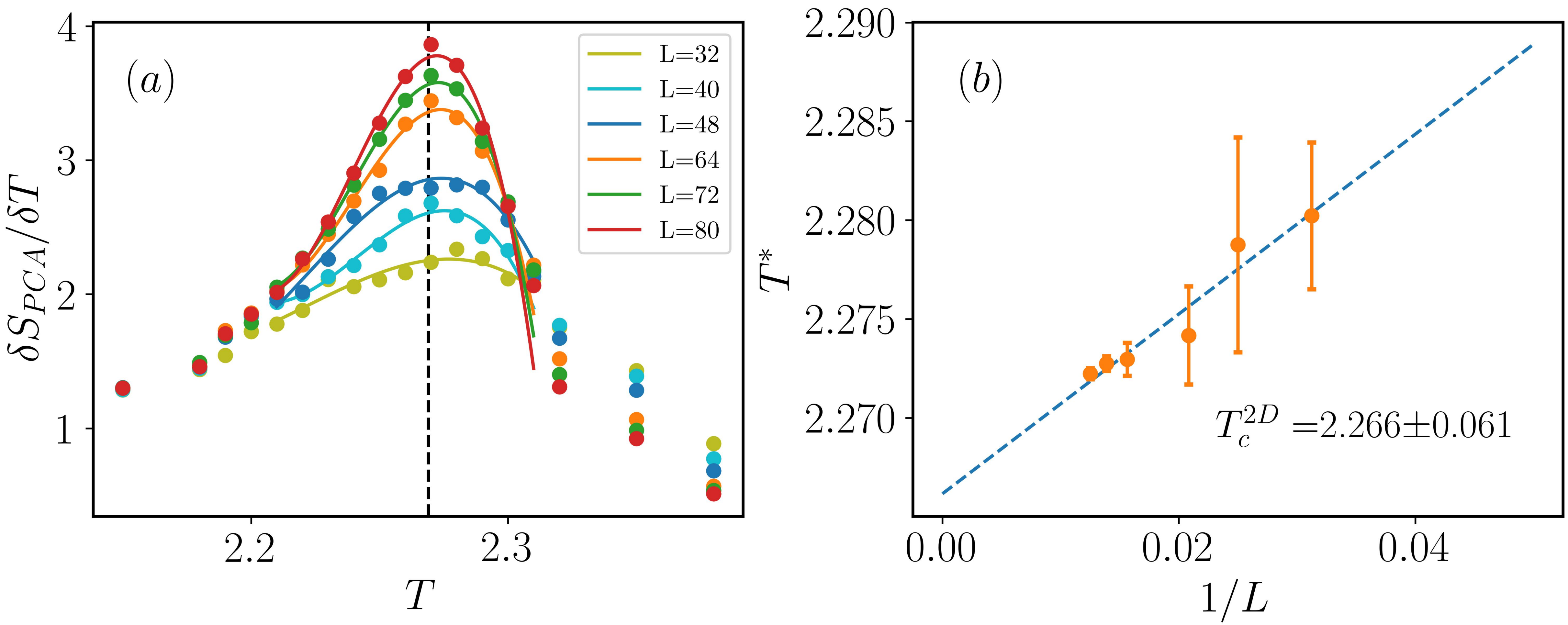


Flex very close to the transition point

Panda, RV, et al., arXiv:2308.13636

PCA entropy: 2D Ising

Quantitative prediction of T_c via a linear finite-size scaling analysis

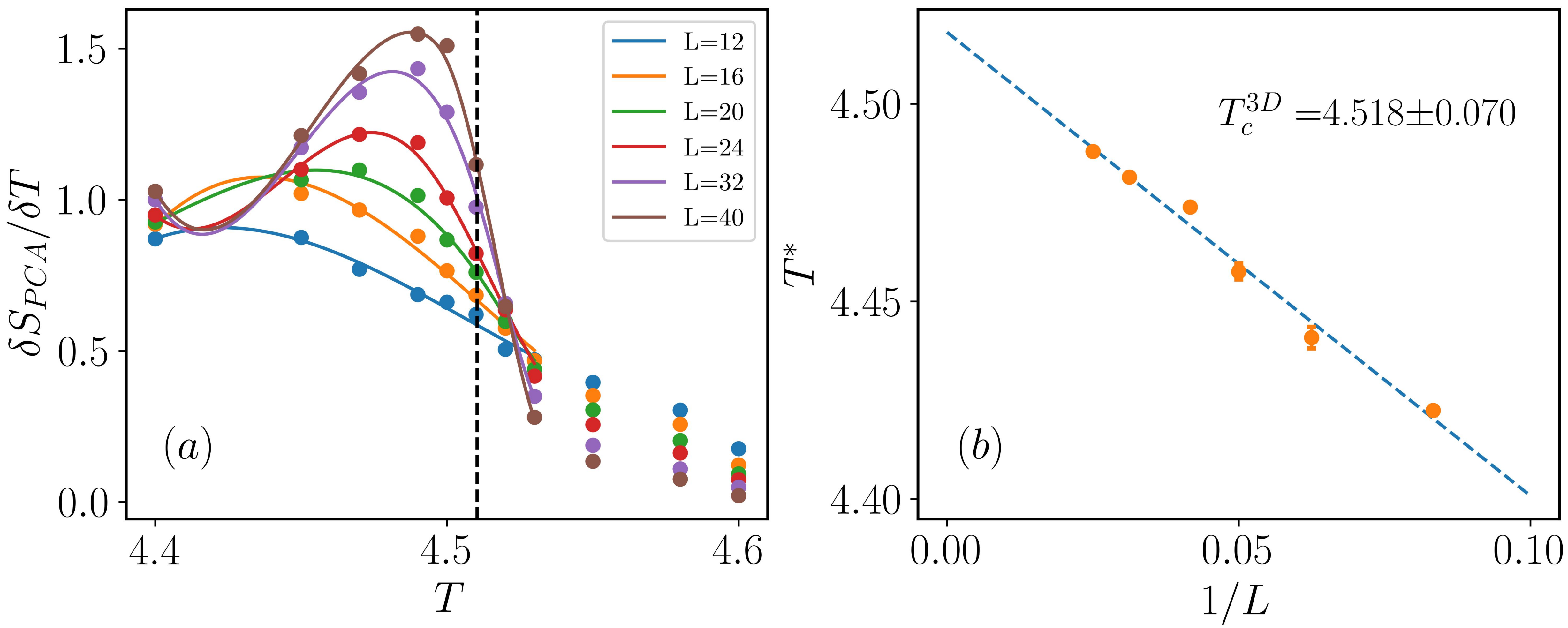


Allows to estimate T_c with less than 1% error

Panda, RV, et al., arXiv:2308.13636

PCA entropy: 3D Ising

Also works nicely for the 3D model!



Panda, RV, et al., arXiv:2308.13636

Data-driven discovery of relevant information in quantum simulation

RV et al., arXiv:2307.10040

In collaboration with M. Oberthaler's group

Experiments

LETTERS

https://doi.org/10.1038/s41567-020-0933-6

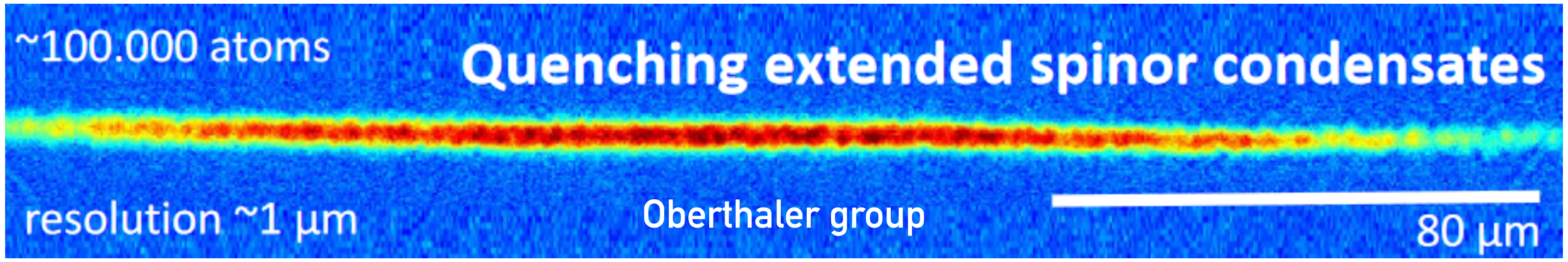
nature physics

Check for updates

Experimental extraction of the quantum effective action for a non-equilibrium many-body system

Maximilian Prüfer[1,3], Torsten V. Zache[2,3], Philipp Kunkel[1], Stefan Lannig[1], Alexis Bonnin[1], Helmut Strobel[1], Jürgen Berges[2] and Markus K. Oberthaler[1]

In collaboration with M. Oberthaler's group



$$\Gamma_t[\Phi] = \sum_{n=1}^{\infty} \frac{1}{n!} \Gamma_t^{\alpha_1,\dots,\alpha_n}(y_1,\dots,y_n)\, \Phi^{\alpha_1}(y_1)\cdots\Phi^{\alpha_n}(y_n)$$

$\downarrow$

What are the relevant operators to determine the proper vertices?

Experiments

LETTERS
https://doi.org/10.1038/s41567-020-0933-6
nature physics
Check for updates

Experimental extraction of the quantum effective action for a non-equilibrium many-body system

Maximilian Prüfer[1,3], Torsten V. Zache[2,3], Philipp Kunkel[1], Stefan Lannig[1], Alexis Bonnin[1], Helmut Strobel[1], Jürgen Berges[2] and Markus K. Oberthaler[1]

In collaboration with M. Oberthaler's group

$$\Gamma_t[\Phi] = \sum_{n=1}^{\infty} \frac{1}{n!} \Gamma_t^{\alpha_1,\dots,\alpha_n}(y_1,\dots,y_n)\, \Phi^{\alpha_1}(y_1)\cdots\Phi^{\alpha_n}(y_n)$$

Obtained from irreducible parts of correlators of the transverse spin

$$F_\perp(y) = F_x(y) + iF_y(y) = |F_\perp(y)|e^{i\varphi(y)}$$

See e.g. Kawaguchi & Ueda, Phys. Rep. '12

Experiments

LETTERS
https://doi.org/10.1038/s41567-020-0933-6
nature physics
Check for updates

Experimental extraction of the quantum effective action for a non-equilibrium many-body system

Maximilian Prüfer[1,3], Torsten V. Zache[2,3], Philipp Kunkel[1], Stefan Lannig[1], Alexis Bonnin[1], Helmut Strobel[1], Jürgen Berges[2] and Markus K. Oberthaler[1]

In collaboration with M. Oberthaler's group

$$\Gamma_t[\Phi] = \sum_{n=1}^{\infty} \frac{1}{n!} \Gamma_t^{\alpha_1,\dots,\alpha_n}(y_1,\dots,y_n)\, \Phi^{\alpha_1}(y_1)\cdots\Phi^{\alpha_n}(y_n)$$

Obtained from irreducible parts of correlators of the transverse spin

$$F_\perp(y) = F_x(y) + iF_y(y) = |F_\perp(y)|e^{i\varphi(y)}$$

See e.g. Kawaguchi & Ueda, Phys. Rep. '12

Determined by particular combinations of populations

$$F_x(y) = \left(N_{+2}^{F=2}(y) - N_{-2}^{F=2}(y)\right)/N_{\mathrm{tot}}^{F=2}(y)$$

$$F_y(y) = \left(N_{+1}^{F=1}(y) - N_{-1}^{F=1}(y)\right)/N_{\mathrm{tot}}^{F=1}(y)$$

Ranking observables: PCA entropy

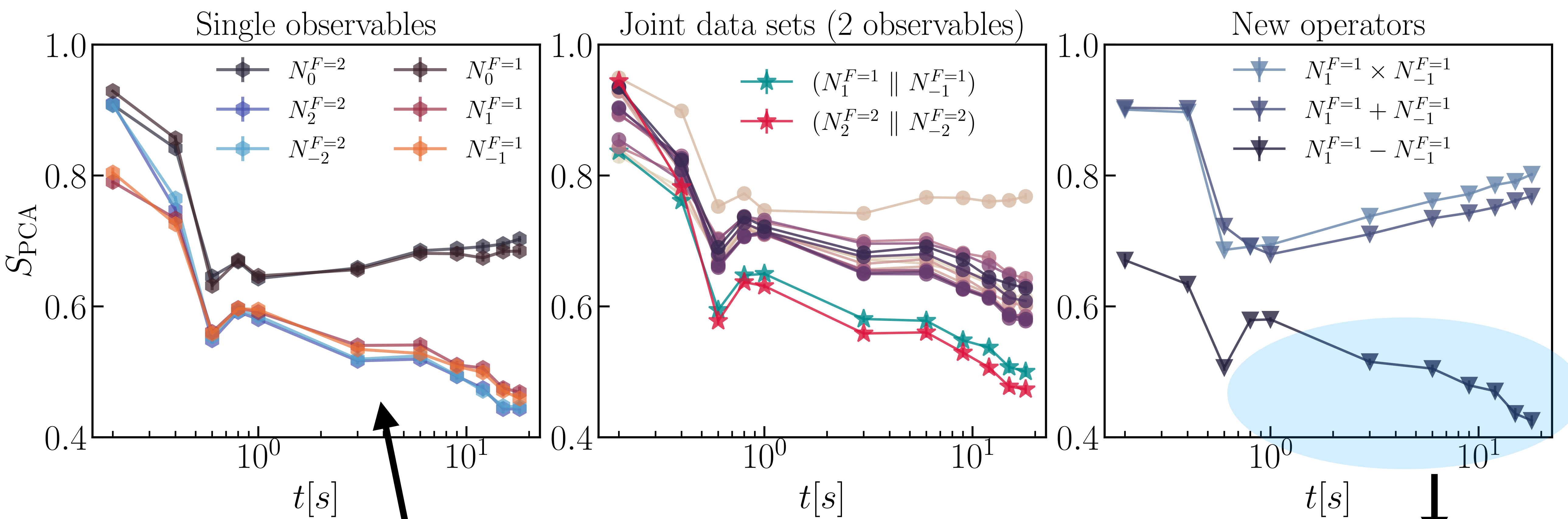
<u>Remember</u>: PCA entropy quantifies how 'messy' a data set is.

In that sense, PCA entropy can be used as a metric to rank different observations according to their relevance

The lower S_{PCA}, the stronger the correlations captured by a given observation

The higher S_{PCA}, the more 'randomness' (less predictive power)

Ranking observables: PCA entropy



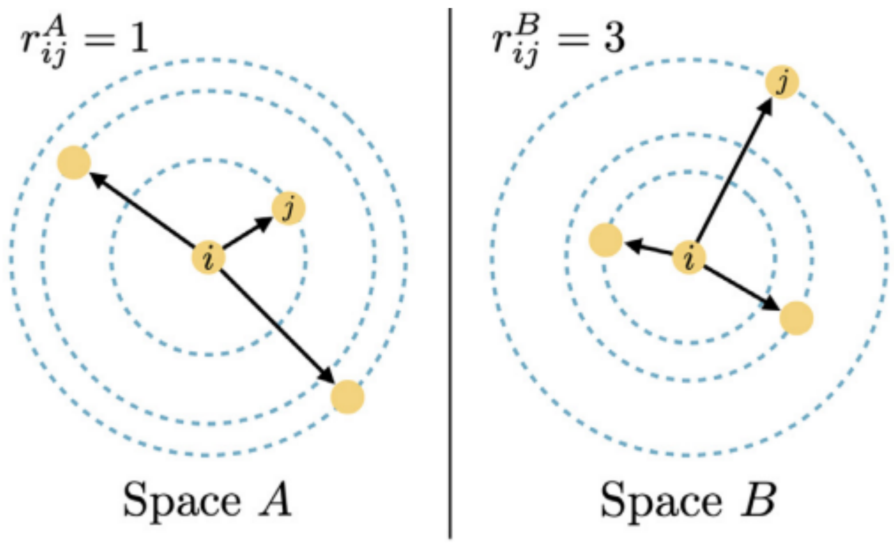
The curves correspond to observations in different basis

$$F_x(y) = \left(N_{+2}^{F=2}(y) - N_{-2}^{F=2}(y)\right)/N_{\mathrm{tot}}^{F=2}(y)$$

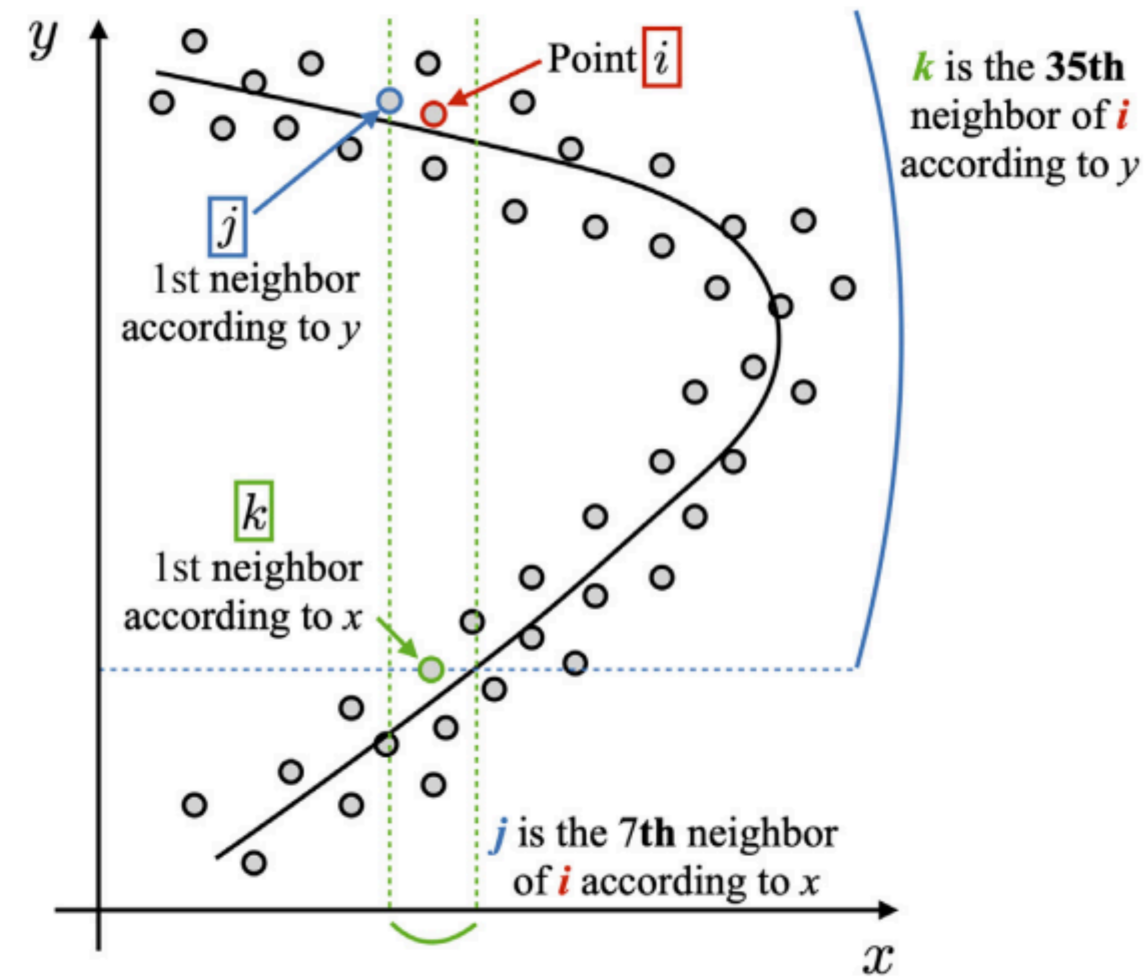
$$F_y(y) = \left(N_{+1}^{F=1}(y) - N_{-1}^{F=1}(y)\right)/N_{\mathrm{tot}}^{F=1}(y)$$

Ranking observables: information imbalance

Recently developed ML technique to quantify the relative amount of information between different types of variables

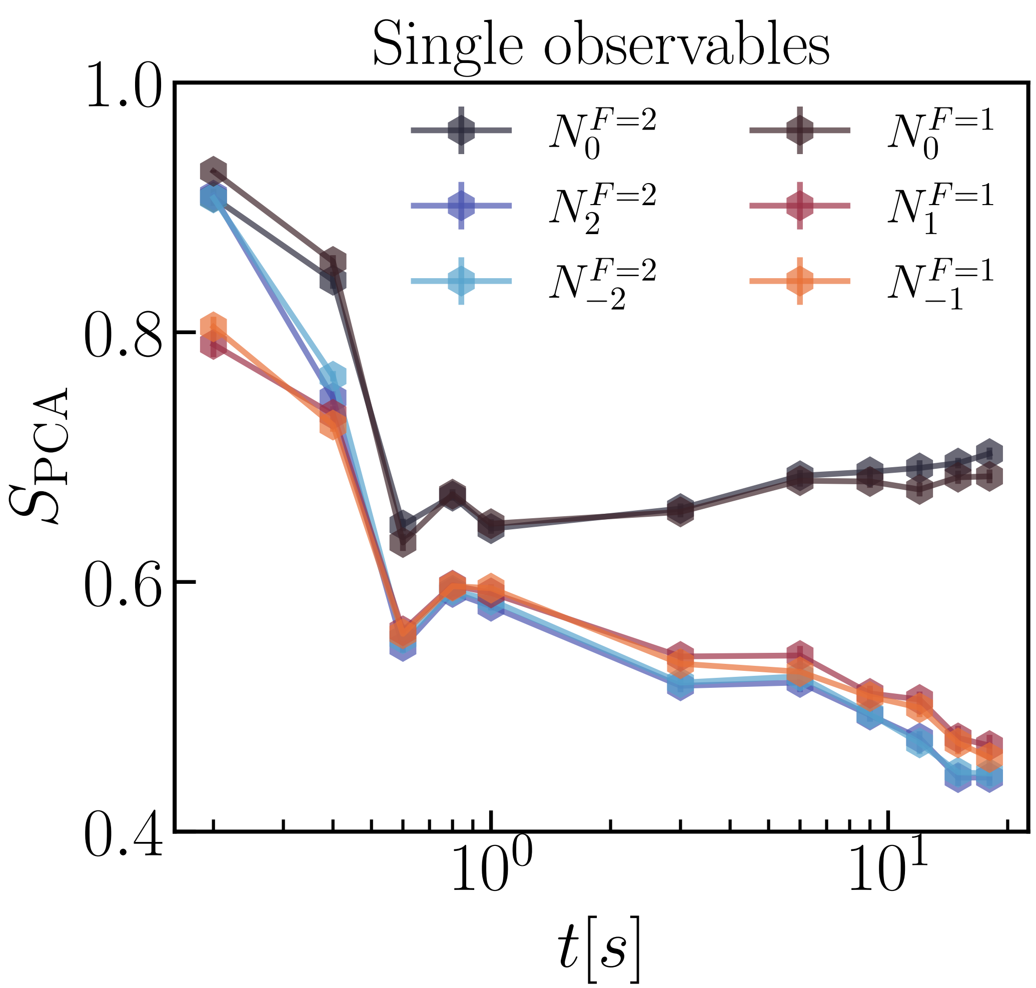


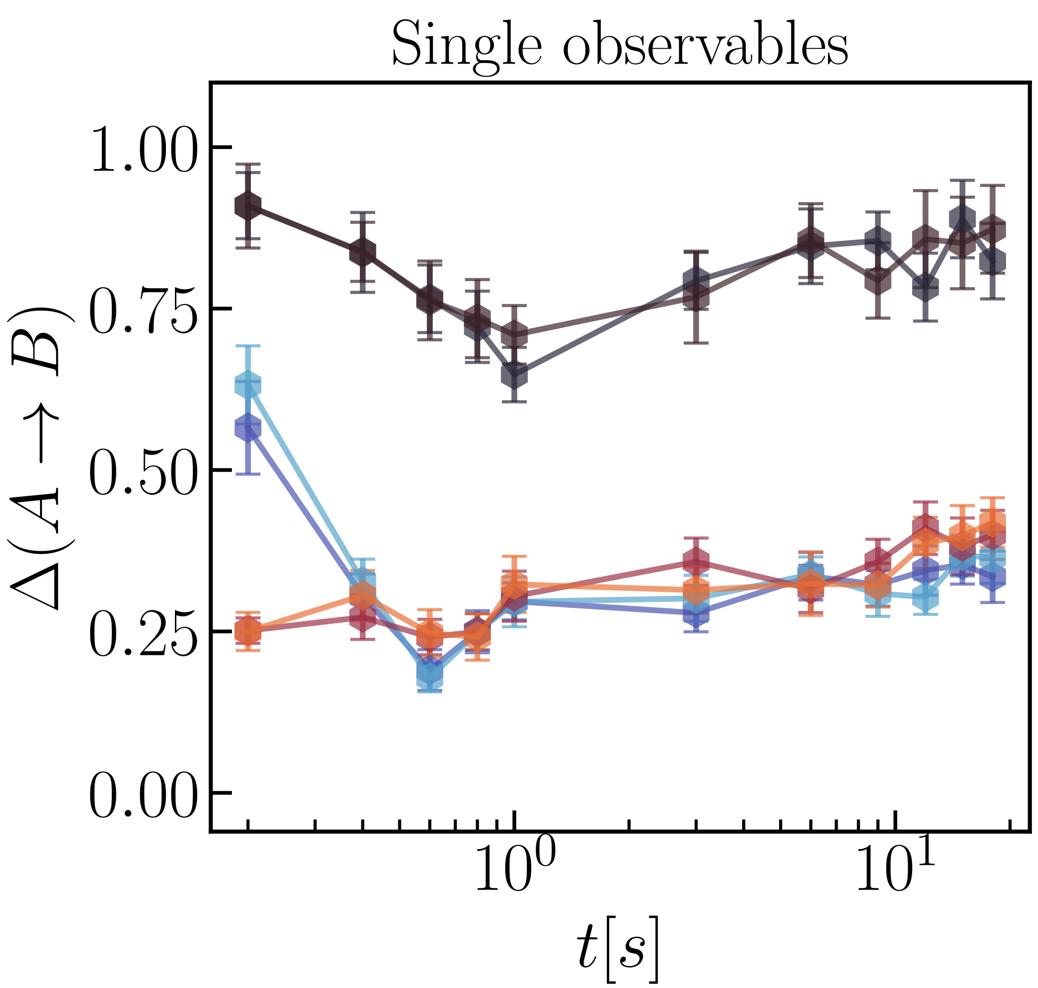
$$\Delta(A \to B) = \frac{2}{N_r^2} \sum_{i,j:r_{ij}^A=1} r_{ij}^B$$



Glielmo et al., PNAS Nexus '22

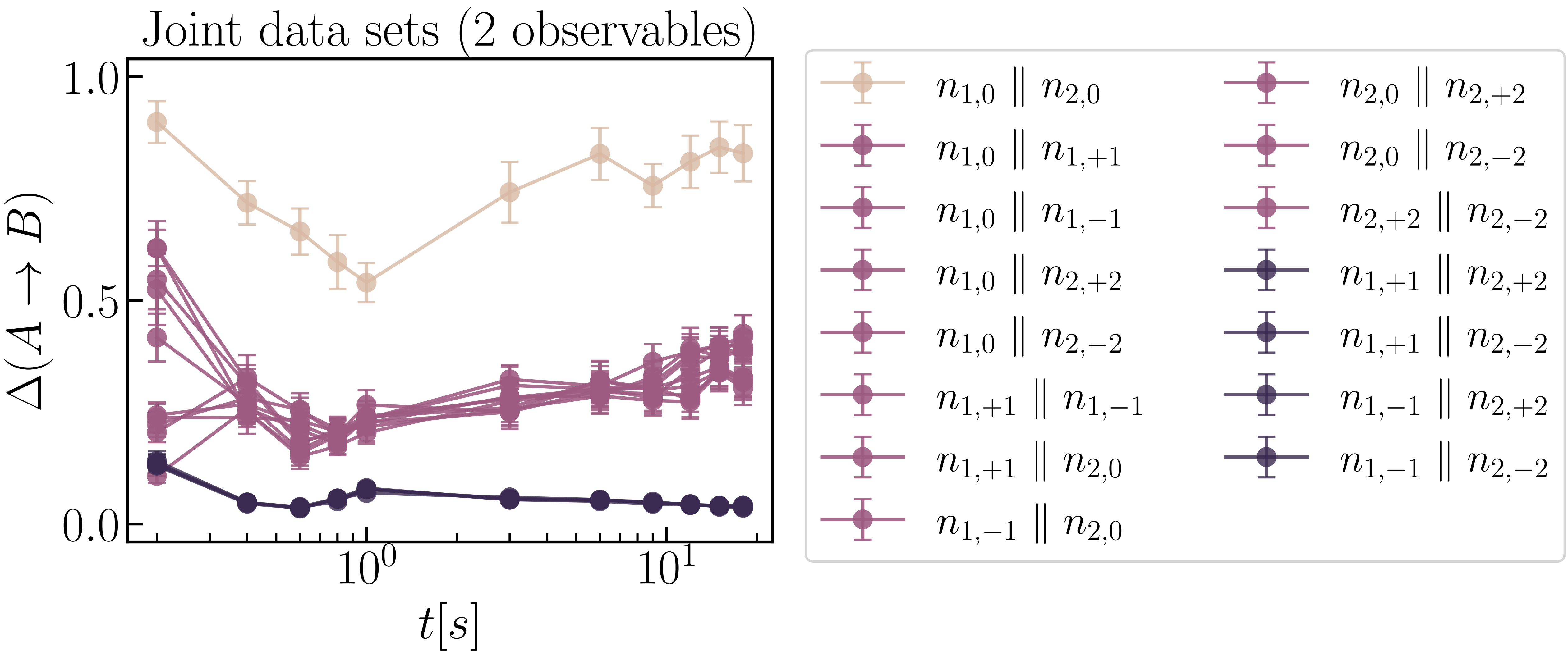
Ranking observables: information imbalance





Cross-verifies ranking obtained with PCA entropy

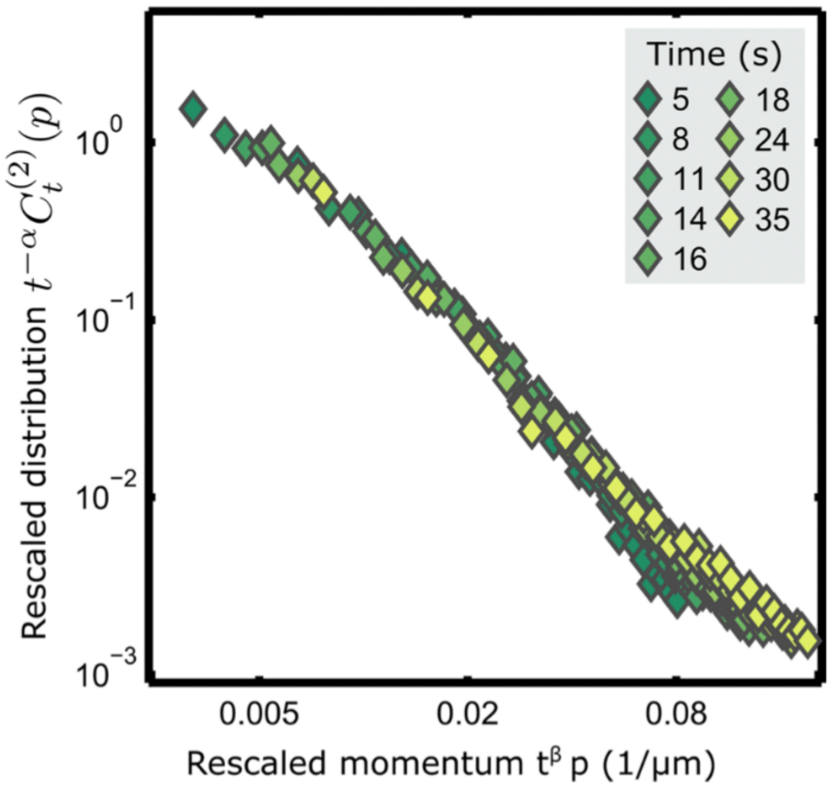
Ranking observables: information imbalance



Complementary characterisation of relevance: needs to combine observables from two relevant pairs to describe the full space of observations

Agnostic bound on universal scaling regime

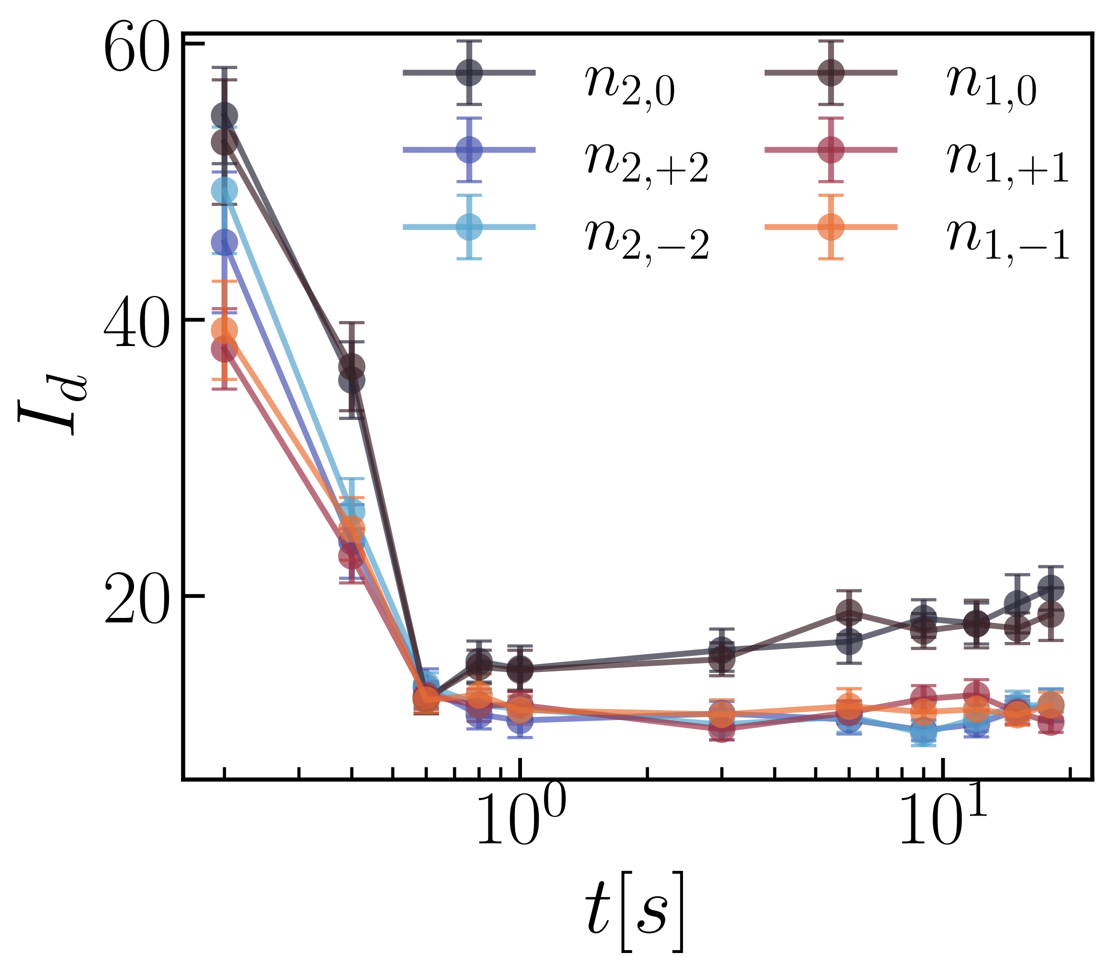
Correlation functions of the transverse spin exhibit self-similar dynamics



Prüfer et al., Nat. Phys. '20

Theo: Berges et al., PRL '08, …

Intrinsic dimension features long, stable plateaus in strong agreement with universal behavior



Conclusions and Outlook

- Non-parametric unsupervised methods provide powerful tools to enable assumption-free discoveries in many-body physics!
- Widely applicable methods: classical/quantum, in and out of equilibrium, and working with limited sampling
- Insights on lattice gauge theory and topological matter (on-going)
- Interesting connections to the entropy and measures of complexity (e.g. Kolmogorov complexity, Shannon entropy)

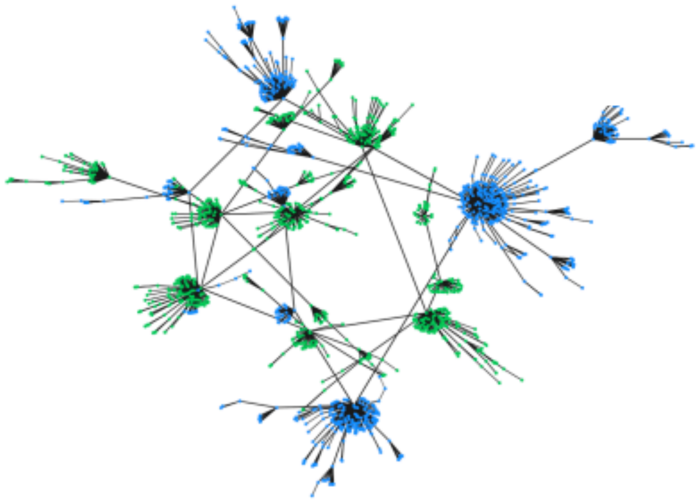
Thank you!

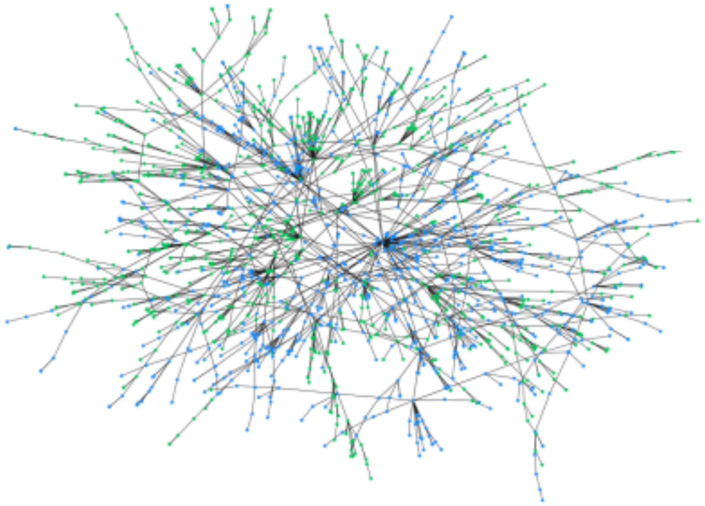
What I didn't talk about

Complex network based 'data mining'

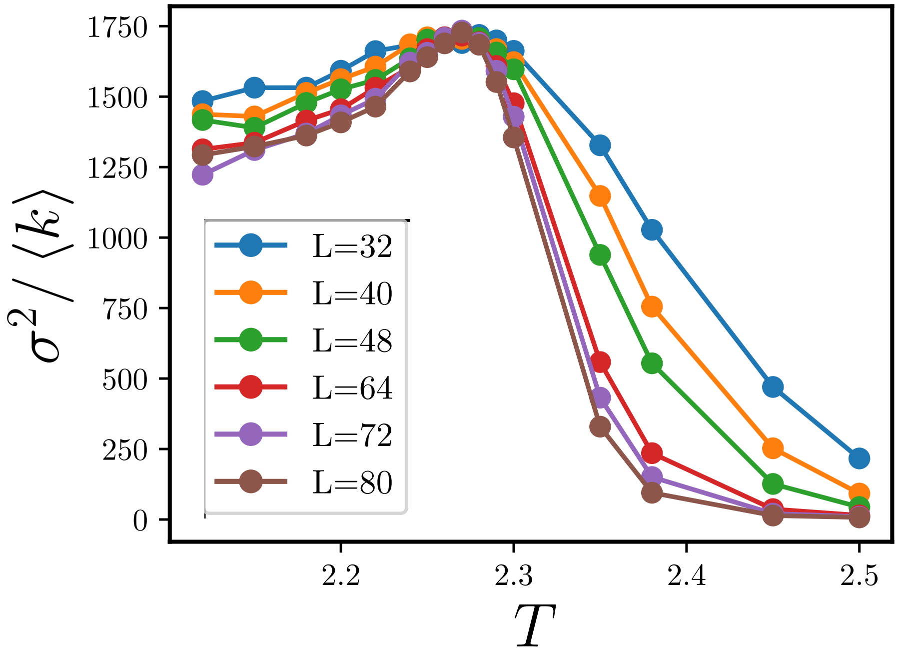
H. Sun

G. Bianconi





- Plethora of tools from network science that provide an in-depth statistical, combinatorial, geometrical and topological analysis of data sets
- Nice complementary tools for unsupervised approaches



Sun, RV, et al., arXiv:2308.13604 (see also arXiv:2301:13216)

Extra material

More about intrinsic dimension

- Lower bound of complexity in data sets (e.g. relation to bottleneck in autoencoders [Ansuini et al., NearIPS 2019])
- Crucial dependence on the chosen scale

Mendes-Santos et al., PRX '21

- Related to the Kolmogorov complexity

How long shall a classical computer code be to reproduce a given string?

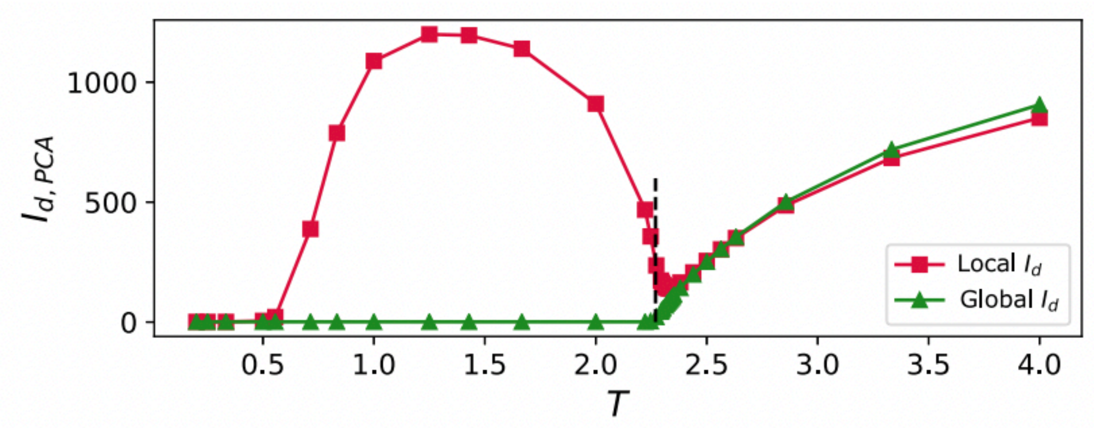
'11111111…'	'10011010…'
print '1' n times (lower complexity)	print '10011010…' (higher complexity)

I_d estimation: PCA

See e.g. Jolliffe (2005)

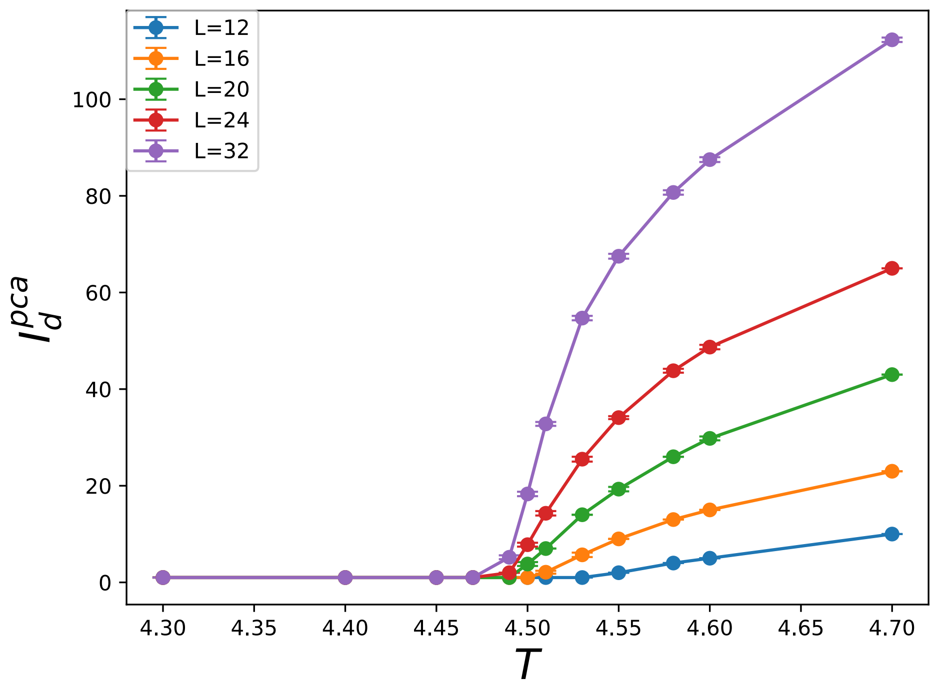
- Based on a ad-hoc cutoff parameter in the integrated spectrum of the covariance matrix $\sum_{n=1}^{I_d} \tilde{\lambda}_n \approx \zeta$
- Bad estimate for curved manifolds

2D Ising



Mendes-Santos et al., PRX '21

3D Ising



Panda, RV, et al., arXiv:2308.13636