計算科学における情報圧縮

Information Compression in Computational Science

2021.10.7

#1:現代物理学における巨大なデータ

Huge data in modern physics

理学系研究科　大久保　毅

Graduate School of Science, **Tsuyoshi Okubo**

Outline

- Background
  - Background of the lectures
  - Computational Science Alliance, The University of Tokyo
  - Tentative lecture schedule
  - Evaluation
- Introduction: Huge data in physics
  - Computational science and data science
  - Why we need information compression?
  - Examples of information compression

Background of lecturer

大久保　毅 (OKUBO Tsuyoshi)

Project Associate Professor,
(Quantum software endowed chair)
Institute for Physics of Intelligence
Sci. Bldg. #1 950

Research:

Statistical Physics, Condensed matter physics, Magnetism,
(Computational Physics)

- Random packing of disks
- Sciophysics (analysis of hierarchical society...)
- Ordering of (classical) frustrated spin system
  - Topological excitations, Skyrmion, ...
- Deconfined quantum criticality
- **Tensor network methods**
  - Quantum spin systems
  - Quantum computer applications
  - Data science applications
- ….

TeNeS

https://www.pasums.issp.u-tokyo.ac.jp/tenes/en

Lecturer

From Yamaji-sensei's slide

山地 洋平 YAMAJI, Youhei
YAMAJI.Youhei@nims.go.jp

Senior Researcher
IFCS Group, GREEN,
National Institute for Materials Science

Research:
Theoretical condensed matter physics
Computational method of many-body quantum systems

Developer of open source codes for supercomputers

Quantum lattice model solver HΦ

http://ma.cms-initiative.jp/ja/index/ja/listapps/hphi

Computational Science Alliance, The University of Tokyo

Computational Science Alliance

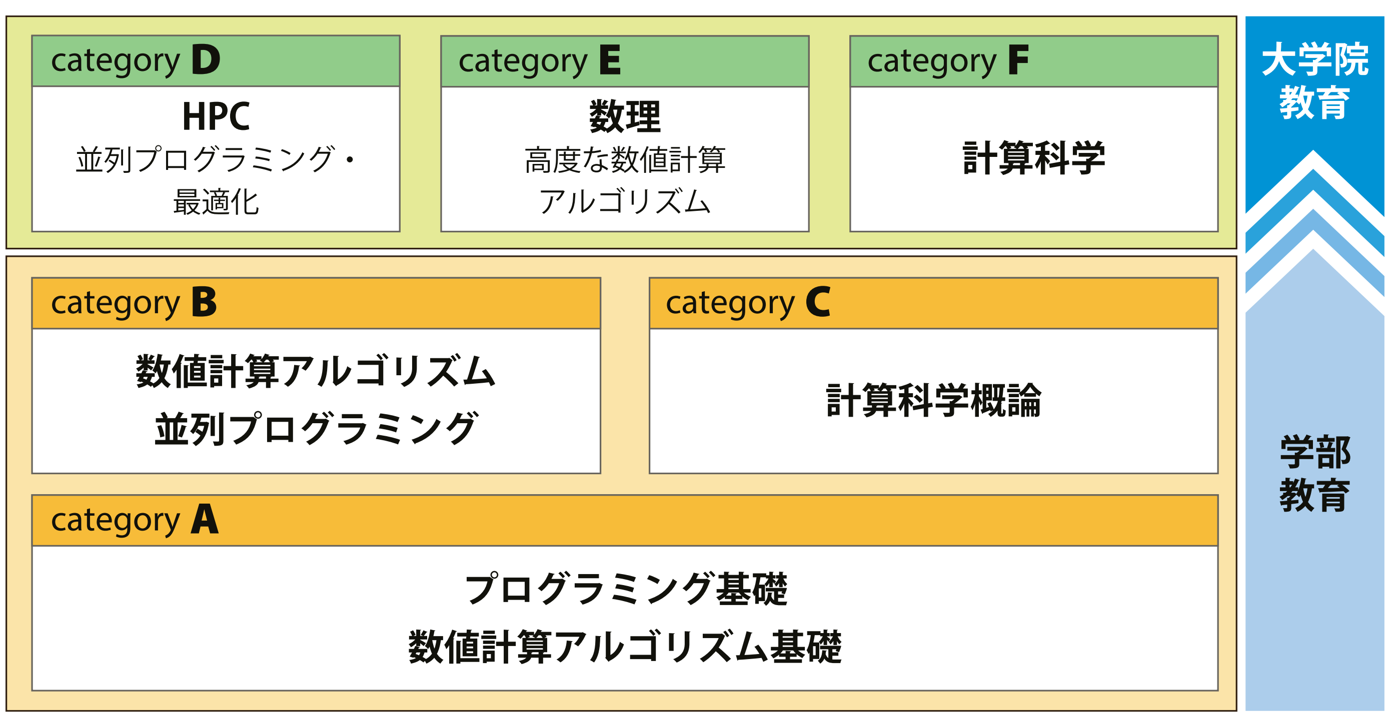
The University of Tokyo

https://www.compsci-alliance.jp

Train experts for computational and computer sciences.

計算科学アライアンス認定講義

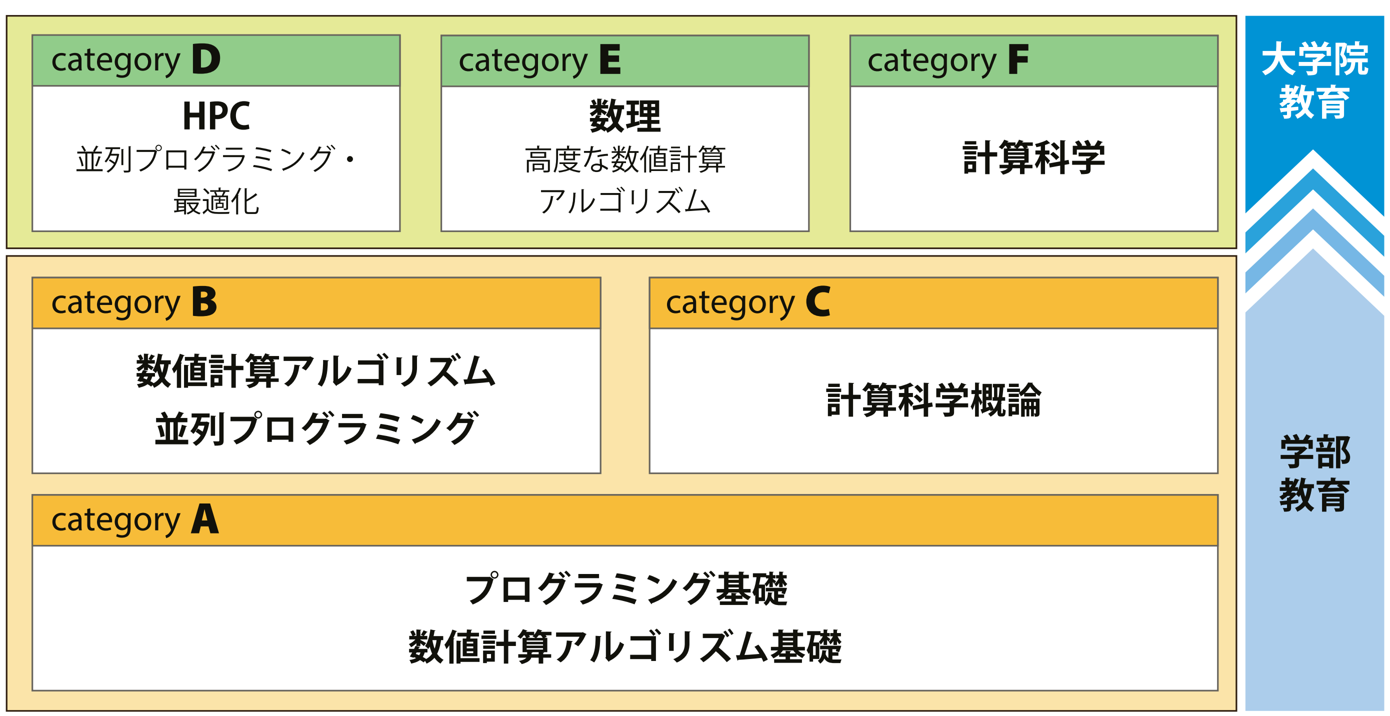
- 平成２９年度から実習にも力点をおいた新しい講義を立ち上げ
- 計算科学・計算機科学に関する80以上の学部・大学院講義とあわせ、「計算科学アライアンス認定講義」として体系化
- 認定講義を内容に応じて6つのカテゴリに分類
- 所定の単位を取得した学生には「修了認定証」を発行
- この講義はカテゴリE



- 学部
  - カテゴリA,B,Cからそれぞれ1.5単位以上
- 大学院
  - カテゴリD,E,Fのうち2つのカテゴリを選択しそれぞれから2単位以上

計算科学アライアンス認定講義：大学院

- カテゴリD - 最先端のスーパーコンピュータを駆使するのに必要とされる技術。種々の並列アルゴリズム、MPI並列やOpenMP並列などの並列プログラミング、メモリアクセス最適化などのチューニング
  - 例：**計算科学アライアンス特別講義I、II**
- カテゴリE - 最先端の数値計算アルゴリズムとその数理的基礎付け。差分法・有限要素法・有限体積法、特異値分解、最適化問題などの手法とその応用
  - 例：**計算科学における情報圧縮**
- カテゴリF - 各分野におけるシミュレーション手法とその研究成果。電子状態計算、分子動力学、量子多体計算、数値流体力学、構造計算、ゲノム解析など
  - 例：**多体問題の計算科学**



Lecture schedule

Okubo	1st: Huge data in modern physics (Today)
	2nd: Information compression in modern physics (+review of linear algebra)
	3rd: Review of linear algebra (+ singular value decomposition)
	4th: Singular value decomposition and low rank approximation
Yamaji	5th: Basics of sparse modeling
	6th: Basics of Krylov subspace methods
	7th: Information compression in materials science
	8th: Accelerating data analysis: Application of sparse modeling
	9th: Data compression: Appication of Krylov subspace method
Okubo	10th: Entanglement of information and matrix product states
	11th: Application of MPS to eigenvalue problems
	12th: Tensor network representation
	13th: Information compression by tensor network renormalization

Additional lectures related to quantum algorithms

Recently, school of science has established
Quantum Software Endowed Chair
（量子ソフトウェア寄付講座）

https://qsw.phys.s.u-tokyo.ac.jp

- Around from December 2021 to February 2022, we will organize lectures related to quantum algorithms.
  - The lectures will be given by Prof. Todo and Okubo.
  - They will be given in Japanese, although we will prepare lecture materials in English.

Tentative topics (90 min each)

- Tensor network renormalizaiton
- Quantum computers and simulations
- Quantum-classical hybrid algorithms and tensor network
- Quantum error corrections

Important infomations

Style: This course will be offered online via zoom.

URL is announced in **UTAS** and **ITC-LMS.**

(It is from 13:15 to 14:45 (90 min. lecture).)

Evaluation: Based on 2 reports:

Exercises include algorithms and computer simulations.

(We will probably provide python codes (jupyter notebooks).)

Notice!

The grade will be evaluated based on the sum of scores of two reports. (So, if you will miss one of them, it will be big disadvantage.)

Slides: The lecture slides will be uploaded to

- ITC-LMS (Information Technology Center Learning Management System)
- https://github.com/compsci-alliance/information-compression

Introduction: Huge data in physics

Computer science and data science

1. Experimental Science
2. Theoretical Science
3. Computational Science
4. Data Science
5. AI?

"Information compression in computational science"
is related to 3rd and 4th sciences.
(It might be slightly related to AI.)

Huge data in physics

Many-body problems in physics

- Celestial movement （天体運動）
- Gases, Liquids
- Molecules, Polymers (eg. Proteins), ...
- Electrons in molecules and solids
- Elemental particles (Quantum Chromo Dynamics)
（量子色力学）

In these problems, "systems" contain huge degrees of freedom:

$6N$-dimensional phase space for classical mechanics

$O(e^N)$-dimensional Hilbert space for quantum system

Complex particle system

Eg. Poliovirus capsid in electrolyte solution

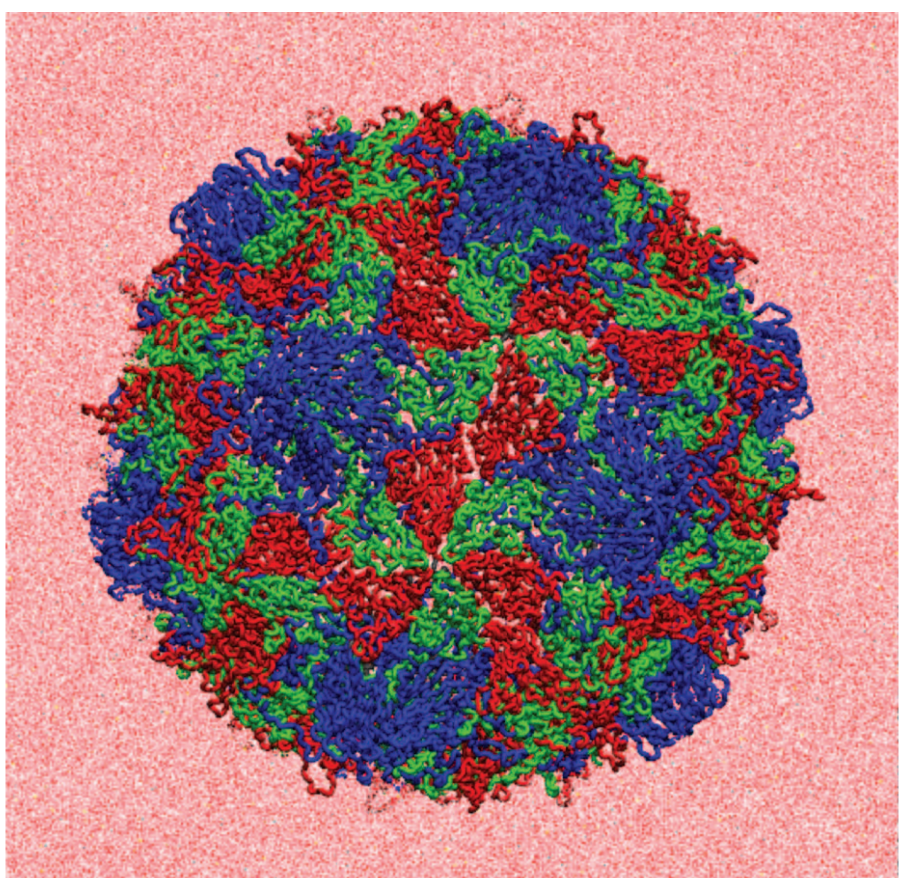
（ポリオウイルス　カプシド）　　（電解質溶液）

Y. Ando et al, J. Chem. Phys. **141**, 165101(2014).

Long-range coulomb interaction

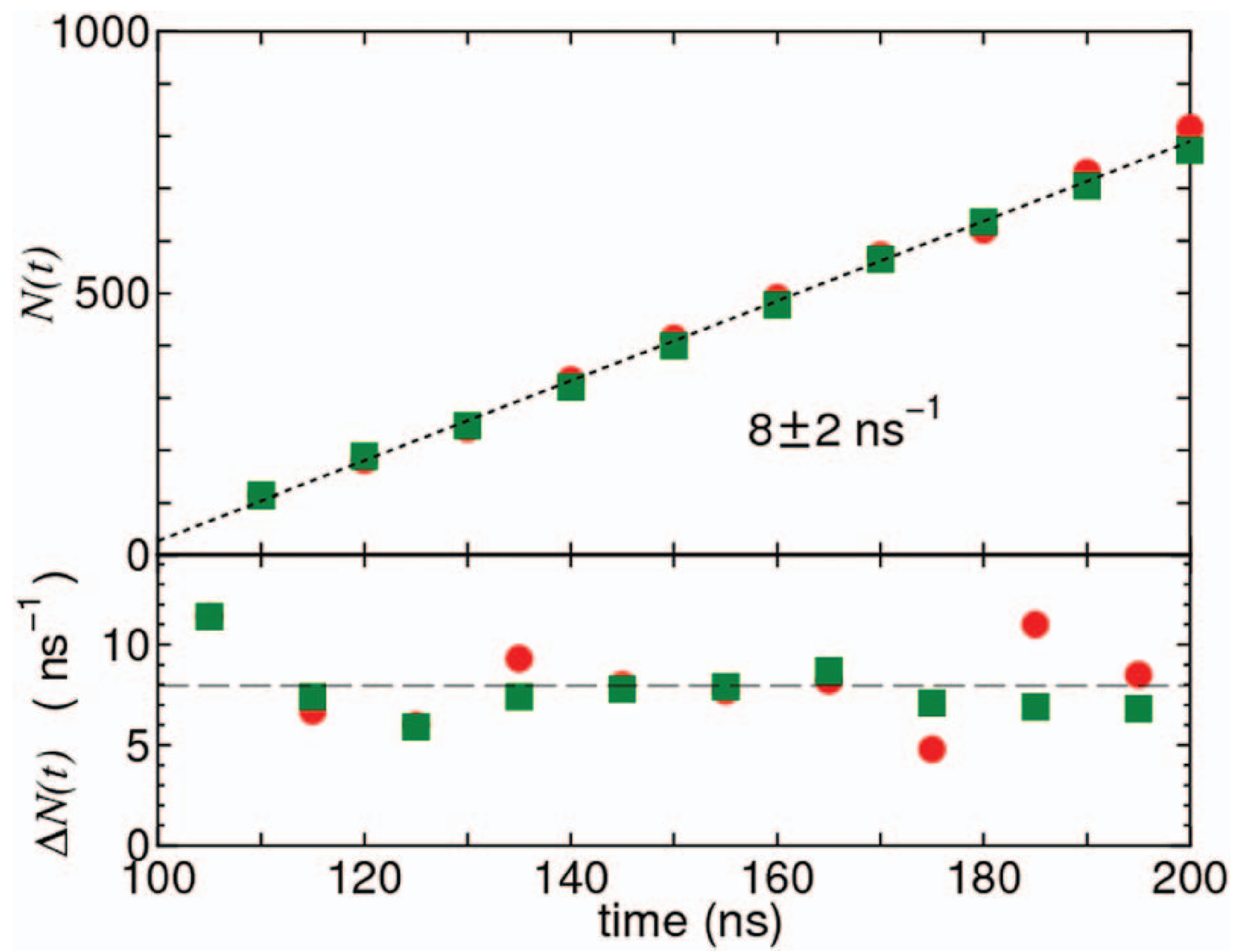
（クーロン相互作用）

Poliovirus capsid



6.5 million atoms

Dynamics of water molecules



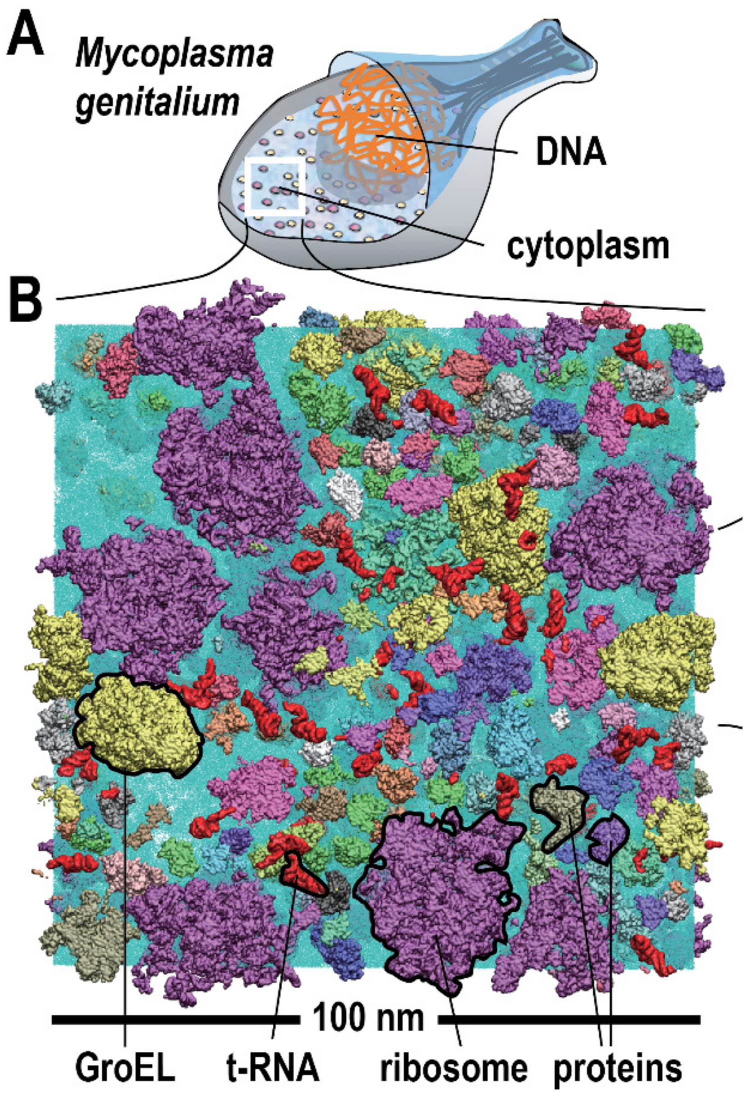
Biological system

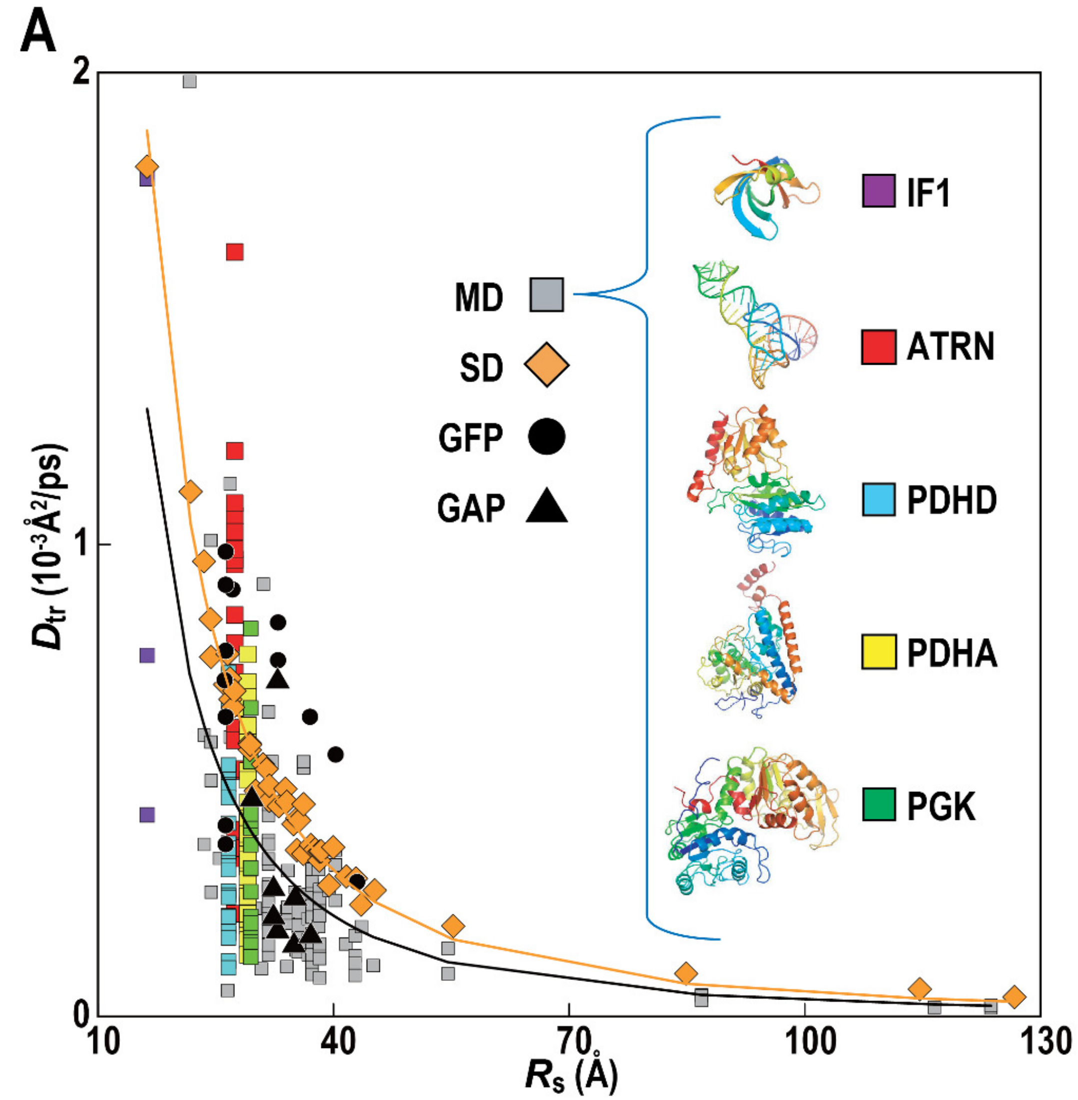
Atomistic model of bacterial cytoplasm (バクテリアの細胞質)

➡ **100 million atoms**

I. Yu, et al., elife **5**, e19274 (2016).

Diffusion constant of various molecules





Diffusion constant of various molecules

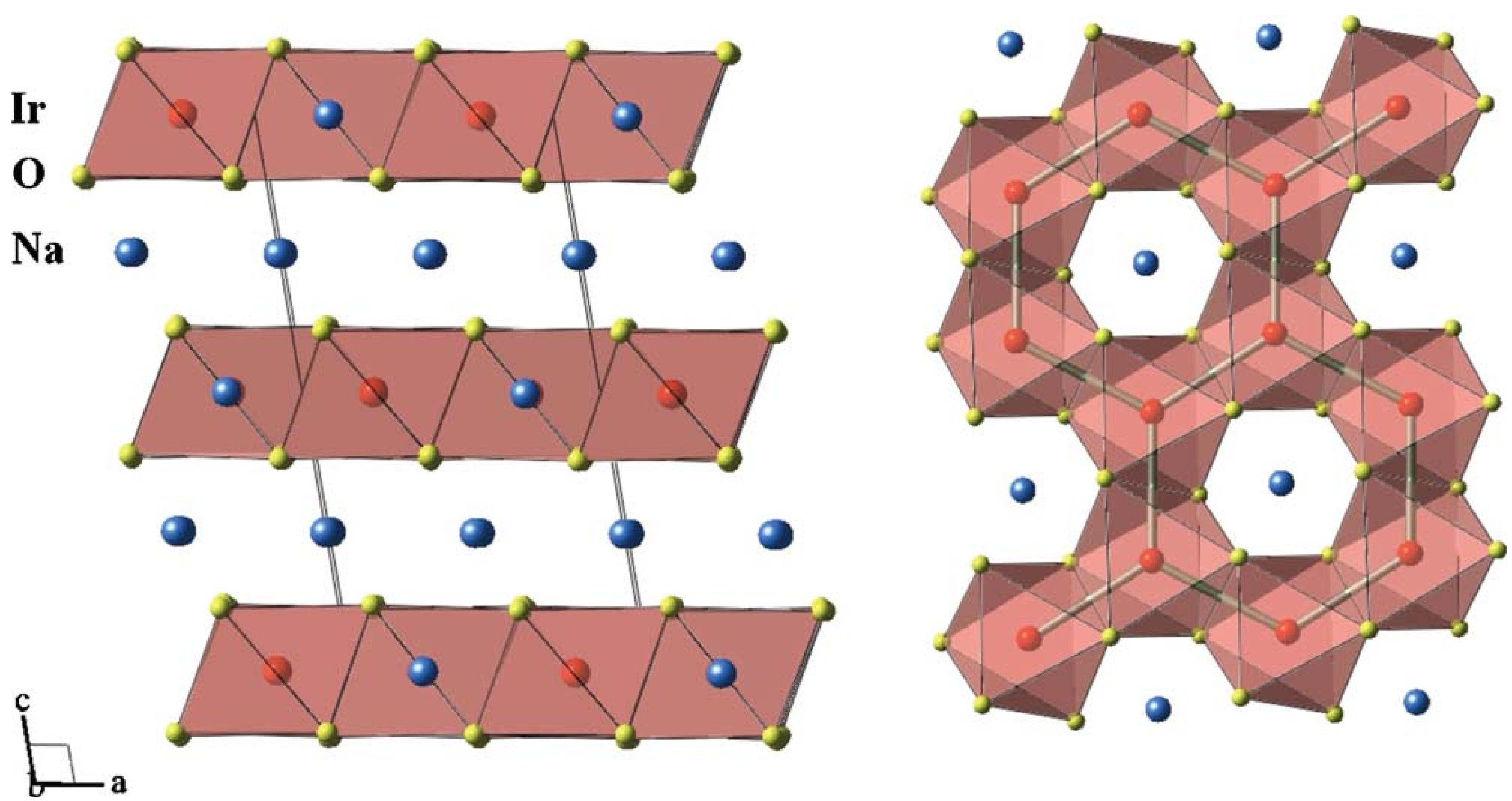
Localized electrons as quantum spin systems

Eg. Antiferromagnetic Mott insulator Na_2IrO_3
（反強磁性）　（モット絶縁体）

Y. Singh and P. Gegenwart, Physical Review B **82**, 064412 (2010)

$$\mathcal{H} = \sum_{i,j} J_{ij} S_i S_j$$

S_i : spin operator



(a) (b)

Why we need information compression?

1. We cannot understand huge information directly.

→ We try to characterize "systems" thorough a few parameters.

Examples:

Thermodynamics:

Systems are characterized by thermodynamic quantities,

Internal energy, Entropy, Pressure, Volume, Particle number,...

Critical phenomena:

Critical systems are characterized by a few critical exponents. (臨界指数)

Related topics in 13th lecture: "Renormalization（繰り込み）"

Why we need information compression?

1. We cannot understand huge information directly.

→ We try to characterize "systems" thorough a few parameters.

William of Ockham

(from wikipedia)

Modeling:

We want to determine "essential" parameters to explain observed phenomena.

Occam's razor:
In order to understand the essence,
we should not assume too much things.

Related topics in 5th and 8th lectures: "Sparse Modeling"

Example of sparse modeling (Compressed sensing)

Reconstruction of MRI image
from smaller samplings

M. Lustig et al, Magnetic Resonance in Medicine **58**, 1182 (2007).

Experiment: Random sampling in k-space

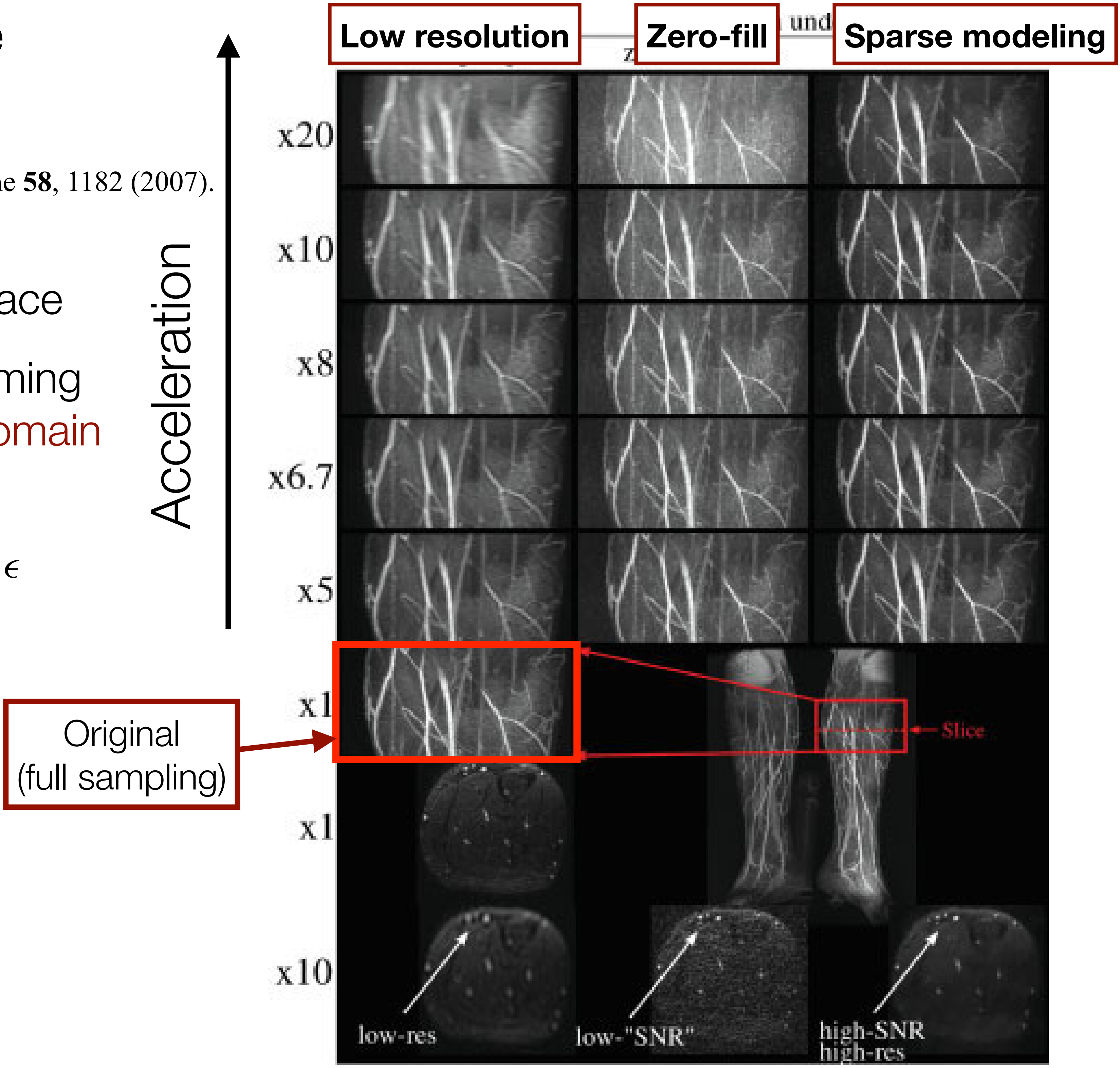
Image reconstruction by assuming
sparsity in wavelet transform domain

$$\begin{aligned} minimize \quad & \|\Psi m\|_1 \\ s.t. \quad & \|\mathcal{F}_u m - y\|_2 < \epsilon \end{aligned}$$

m :image
Ψ :wavelet transform
$\mathcal{F}_u$:Fourier transform
y :Experimental data in k-space

L_1 norm: $||x||_1 = \sum_i |x_i|$

L_2 norm: $||x||_2 = \sqrt{\sum_i |x_i|^2}$



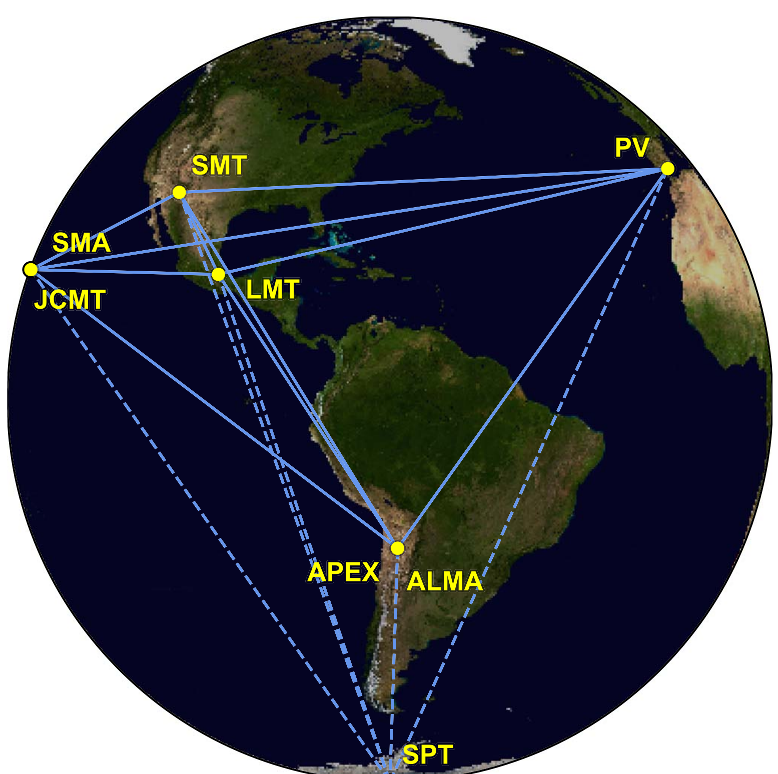
Example of sparse modeling

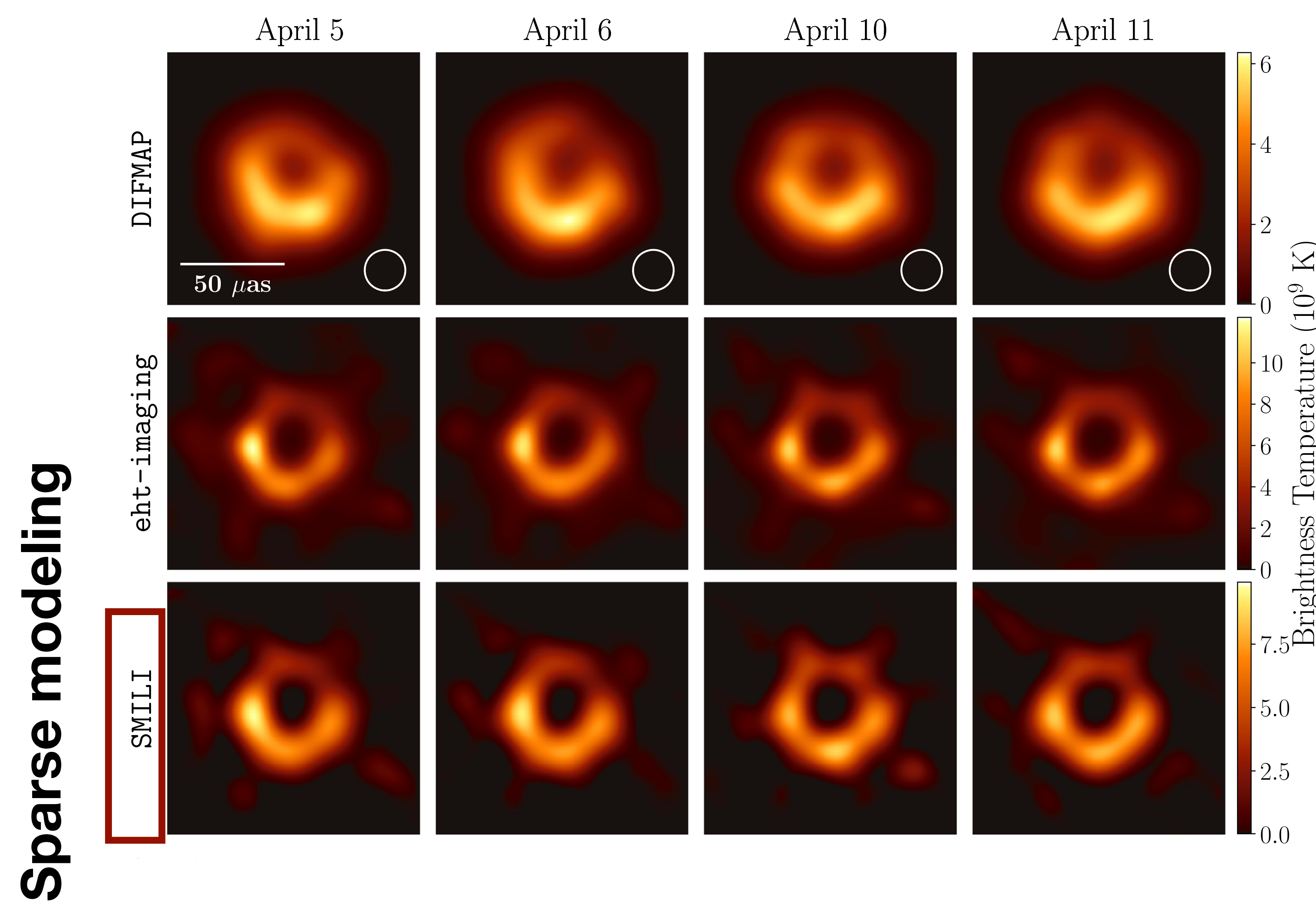
Reconstruction of a black hole image

EHT collaboration, Astrophys. J. Lett. **875**, L1- L6 (2019).

The Event Horizon Telescope (EHT)

Network of **radio telescopes**
(電波望遠鏡)





Why we need information compression?

2. We cannot treat entire data in the present computers.

Available memories in the present computers:

		Double precision real number = 8 Bite
Personal computers:	~10 GB	$\sim 10^{9}$
Supercomputers:	~100 GB / node	$\sim 10^{10}$
Fugaku@RIKEN, **Oakbridge-CX**@UTokyo, **Ohtaka**@ISSP, UTokyo ...	~1 PB (whole system)	$\sim 10^{14}$

Notice: In quantum system, the size of Hilbert space is $O(e^{N})$

Why we need information compression?

2. We can not treat entire data in the present computers.

Try to reduce the "effective" dimension of a (vector) space.

By taking proper basis set,
we can represent a (quantum) state efficiently.

- Low rank approximation (4th lecture)
- Krylov subspace (6th and 9th lectures)
- Tensor network states (10th, 11th 12th lectures)
- ...

Examples of information compression 1

Krylov subspace

linear subspace generated by a square matrix (M) and a vector (v) as

$$\mathcal{K}_n(M, \vec{v}) = \mathrm{span}\left\{\vec{v}, M\vec{v}, M^2\vec{v}, \ldots, M^{n-1}\vec{v}\right\}$$

For quantum many body problems:

$M = \mathcal{H}$:Hamiltonian

$\vec{v} = |\phi\rangle$:wavevector

⇒ Solve the eigenvalue problem within
a restricted space (Krylov subspace)

Lanczos method, Arnoldi method

* In these method, we do not necessarily need explicit matrix.
It is enough to know the result of matrix vector multiplication.

Examples of information compression 2

Compression of an image

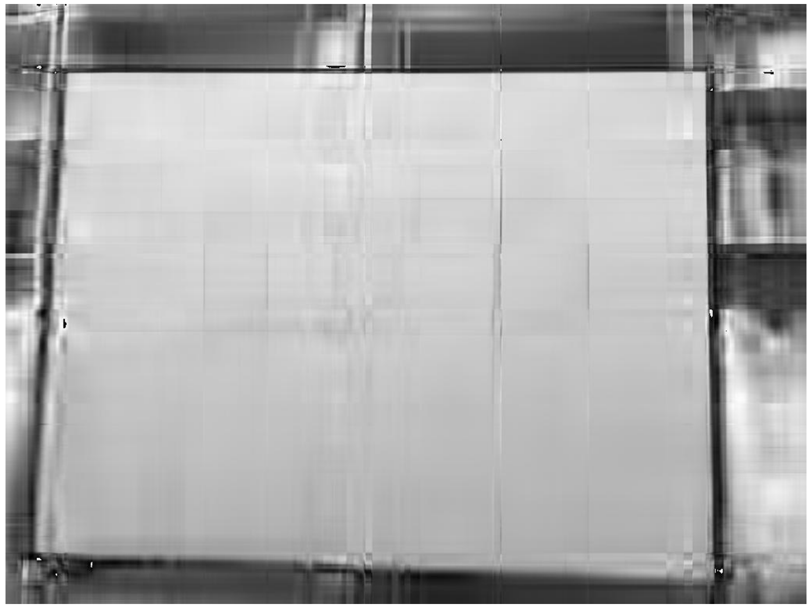
$\boxed{\text{image} = \text{matrix}}$

Original

$\chi = 768$

of "singular values"

Compressed



$\chi = 10$

$\chi = 100$

Examples of information compression 2

Compression of a color image

About 10% compressed

$\boxed{\text{image} = \text{tensor}}$

Original

$\chi = 768$

SVD

HOSVD

Singular value decomposition （特異値分解）

Singular value decomposition (SVD):

For a K × L matrix M,

$$M = U\Lambda V^\dagger$$

$U, V^\dagger$: (half) unitary

Λ : Diagonal

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}$$

$$M_{i,j} = \sum_m U_{i,m}\lambda_m V^\dagger_{m,j}$$

Singular values: $\lambda_m \geq 0$

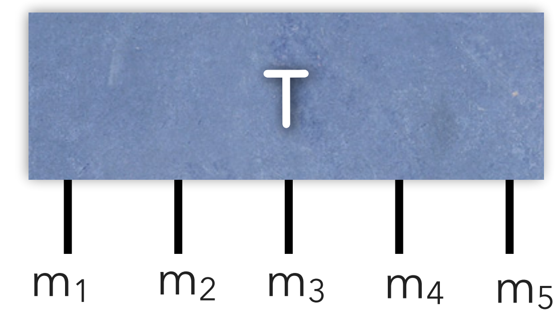
Singular vectors:
$$\sum_i U_{i,m}U^\dagger_{m,j} = \delta_i, j$$
$$\sum_i V_{i,m}V^\dagger_{m,j} = \delta_i, j$$

By taking only several larger singular values,
we can approximate M as **a lower rank matrix**.

(low rank approximation)

Examples of information compression 3

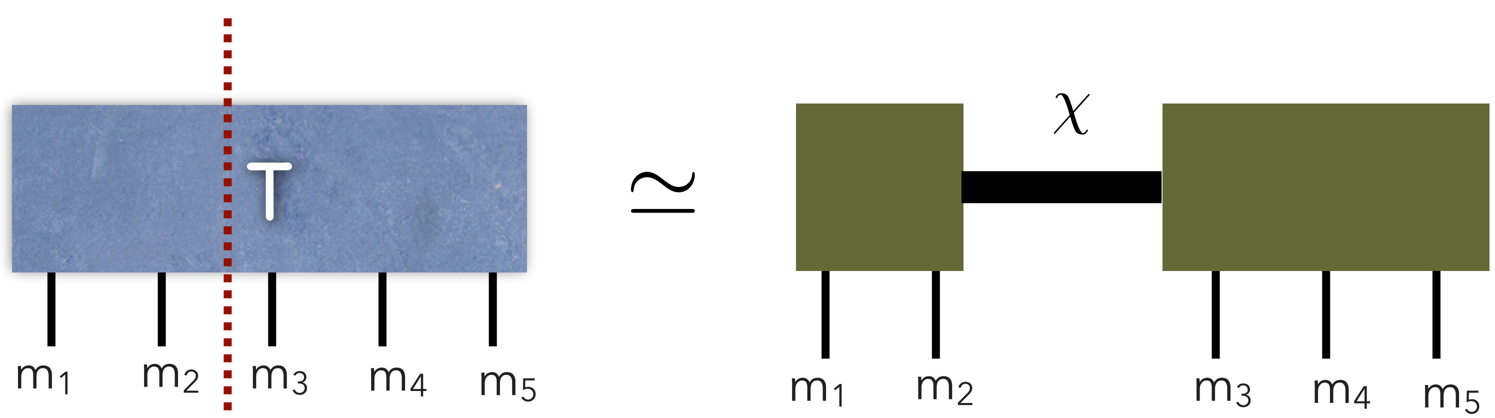
$T_{m_1,m_2,\cdots,m_N}$:N-leg tensor (or Vector)



We can consider it as a matrix by making two groups:

$$T_{\{m_1,m_2,\cdots,m_M\},\{m_{M+1},\cdots,m_N\}}$$

We can perform the low rank approximation of T.



We may consider similar approximation for each block.

Tensor network decomposition

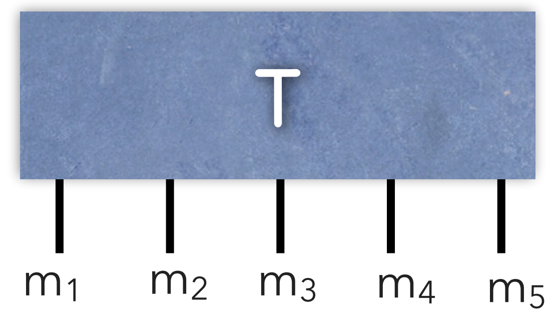
Examples of information compression 3

Wave function: $|\Psi\rangle = \sum_{\{m_i = \uparrow\downarrow\} \text{ or } \{m_i = 0,1\}} T_{m_1,m_2,\cdots,m_N} |m_1, m_2, \cdots, m_N\rangle$

(波動関数)

$T_{m_1,m_2,\cdots,m_N}$ N-rank tensor (or Vector)

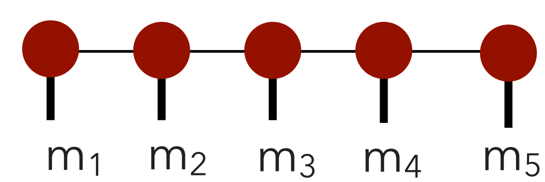
# of Elements＝2^N

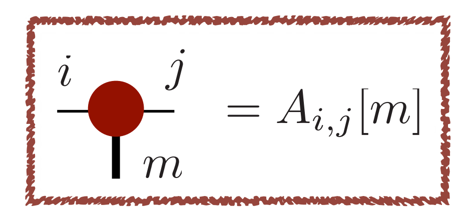


Approximation as a product of "matrices"

Matrix Product States **(行列積状態)**

(Tensor train decomposition)





$$T_{m_1,m_2,\cdots,m_N} \simeq A_1[m_1]A_2[m_2]\cdots A_N[m_N]$$

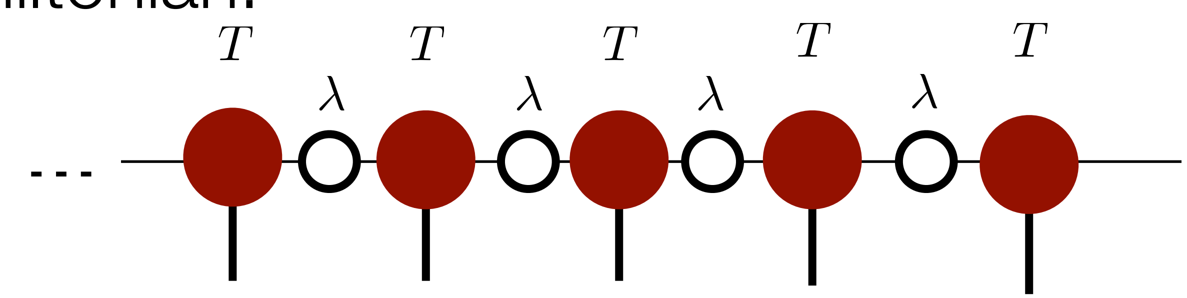
$A[m]$: Matrix for state m

Example of MPS: AKLT state

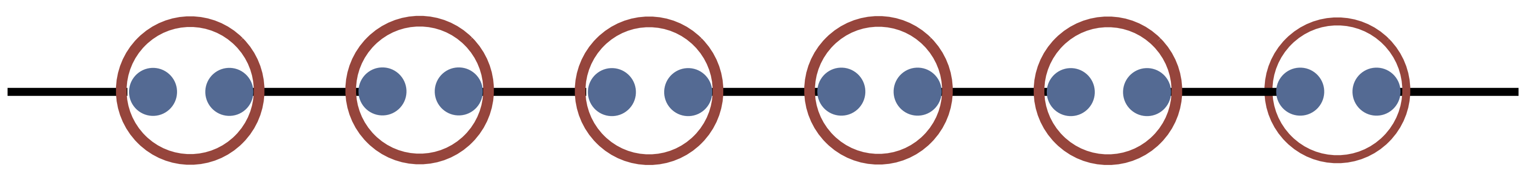
S=1 Affleck-Kennedy-Lieb-Tasaki (AKLT) Hamiltonian:

$$\mathcal{H} = J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j + \frac{J}{3} \sum_{\langle i,j \rangle} \left(\vec{S}_i \cdot \vec{S}_j\right)^2$$

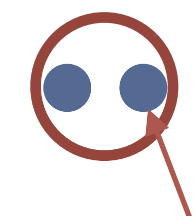
$$(J > 0)$$



The ground state of AKLT model:



S=1 spin:



S=1/2 spin

χ=2 iMPS: (U. Schollwock, Annals. of Physics **326**, 96 (2011))

Spin singlet

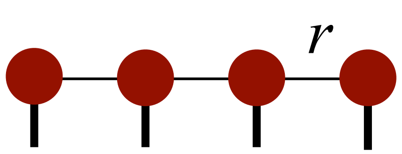
$$T[S_z = 1] = \sqrt{\frac{4}{3}} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

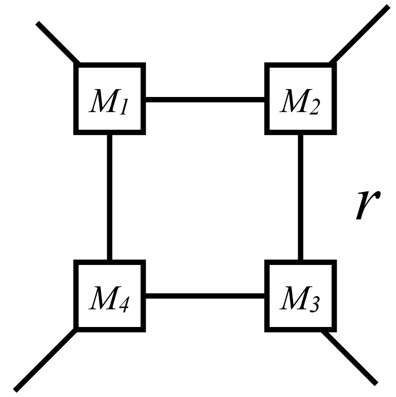
$$T[S_z = 0] = \sqrt{\frac{2}{3}} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \lambda = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$T[S_z = -1] = \sqrt{\frac{4}{3}} \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}$$

Application of MPS to data science

Tensor train (TT) and tensor ring (TR) decompositions for real data.





COIL-100 dataset = $\underline{32 \times 32} \times \underline{3} \times \underline{7200}$ tensor

Pixel color # of pictures

Compression by tensor ring decomposition.

(Q. Zhao, et al arXiv:1606.05535)

error Rank

	ϵ	r_{max}	$\bar{r}$	Acc. (%) ($\rho = 50\%$)	Acc. (%) ($\rho = 10\%$)
CP-ALS	0.20	70	70	97.46	80.03
	0.30	17	17	97.56	83.38
	0.39	5	5	90.40	77.70
	0.47	2	2	45.05	39.10
TT-SVD	0.19	67	47.3	99.05	89.11
	0.28	23	16.3	98.99	88.45
	0.37	8	6.3	96.29	86.02
	0.46	3	2.7	47.78	44.00
TR-SVD	0.19	23	12.0	99.14	89.29
	0.28	10	6.0	99.19	89.89
	0.36	5	3.5	98.51	88.10
	0.43	3	2.3	83.43	73.20

Compressing deep neural network

Z.-F. Gao et al, Phys. Rev. Research **2**, 023300 (2020).

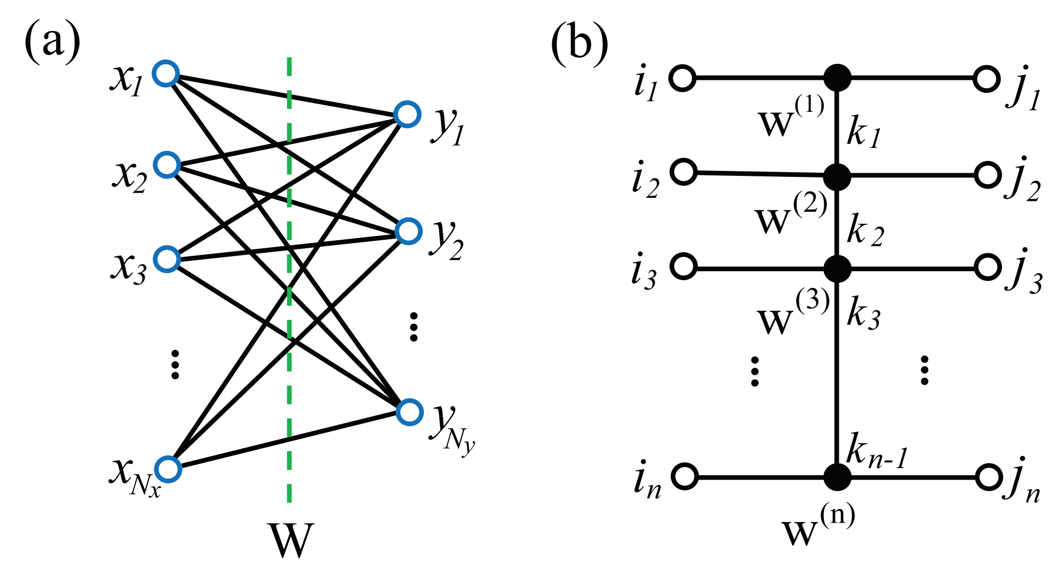
MPO approximation of the weight matrix

x_i: input neuron (pixel)

y_i: output neuron

W_{ij}: weight matrix connecting x and y

MPO approximation of W



Example: application to classification problems

TABLE I. Test accuracy a and compression ratios ρ obtained in the original and MPO representations of LeNet-5 on MNIST and VGG on CIFAR-10.

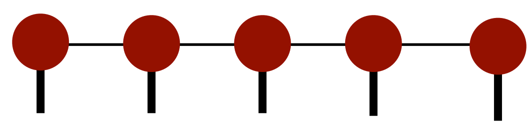
Data set	Network	Original Rep a (%)	MPO-Net a (%)	MPO-Net ρ
MNIST	LeNet-5	99.17 ± 0.04	99.17 ± 0.08	0.05
CIFAR-10	VGG-16	93.13 ± 0.39	93.76 ± 0.16	$\sim$0.0005
	VGG-19	93.36 ± 0.26	93.80 ± 0.09	$\sim$0.0005

a:**accuracy (%)**

ρ:**compression ratio**

Examples of tensor decompositions

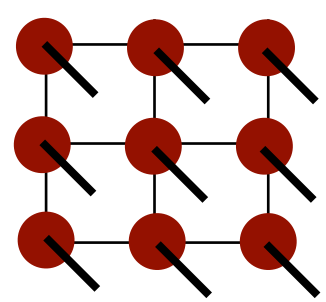
MPS:



Good for 1d gapped systems

(1d correlation in data)

PEPS, TPS:

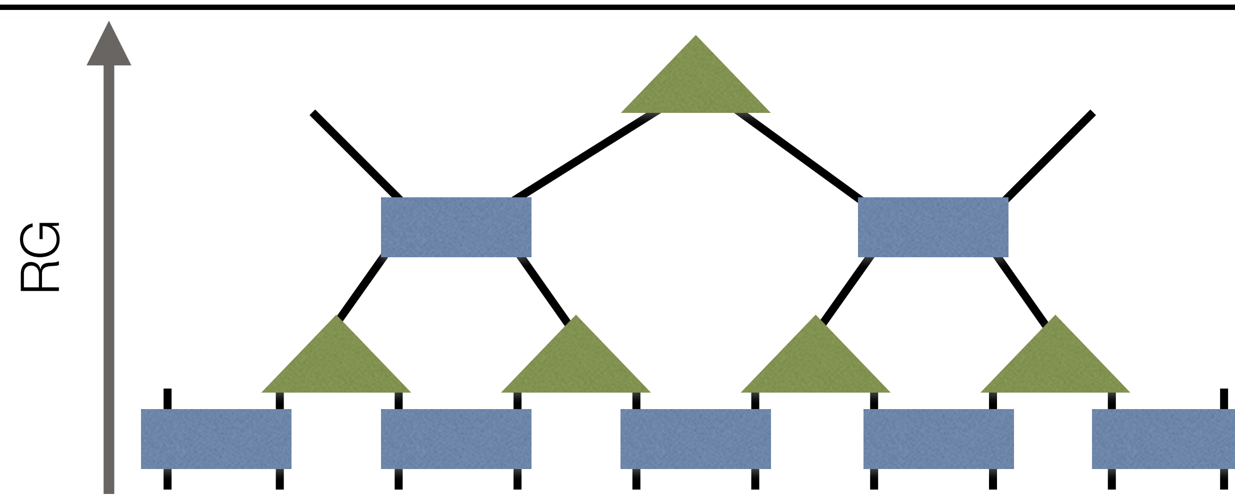


For higher dimensional systems

Extension of MPS

(higher correlation in data)

MERA:

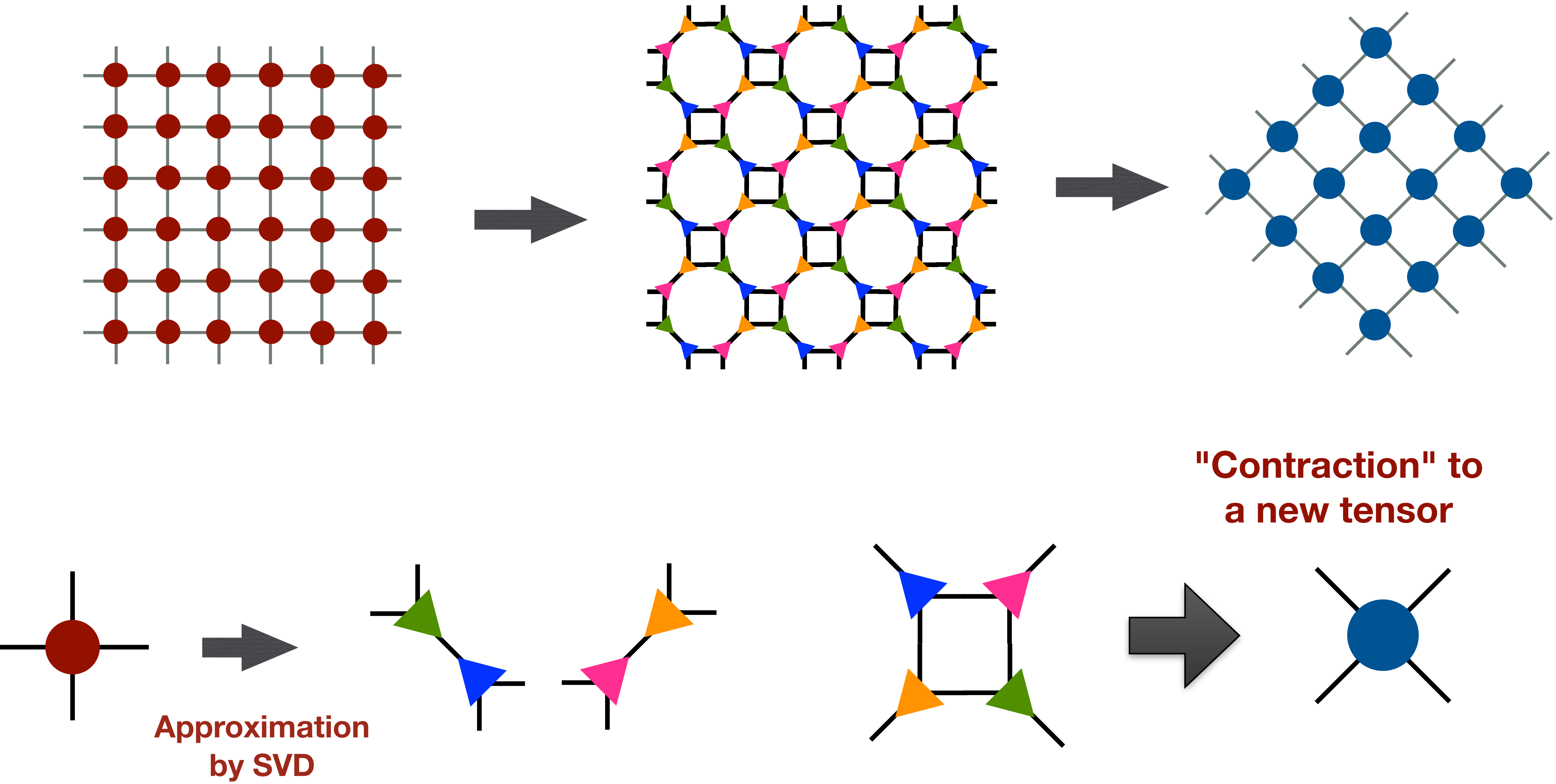


Scale invariant systems

（スケール不変）

Real space renormalization（実空間繰り込み）

Corse graining of a tensor network representing a **scaler.**



Example of calculation

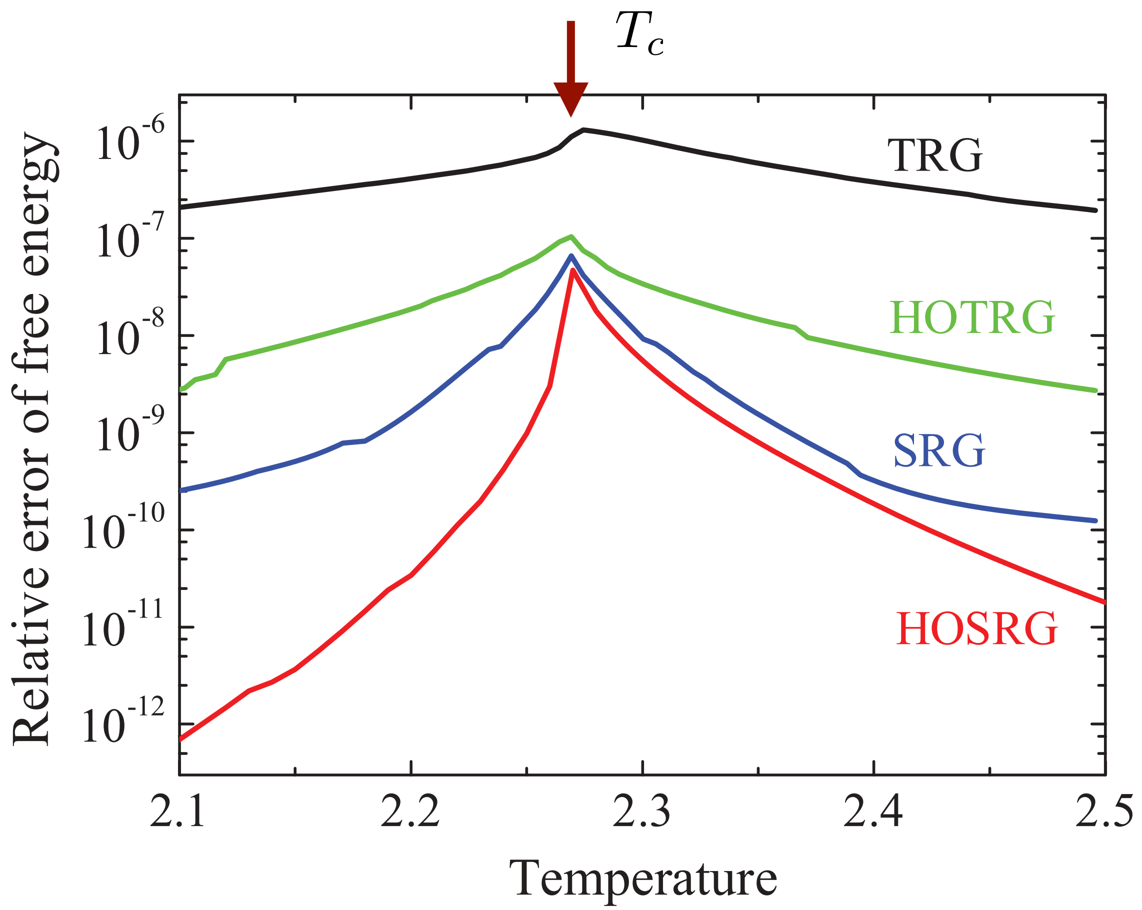
Ising model in infinite size

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

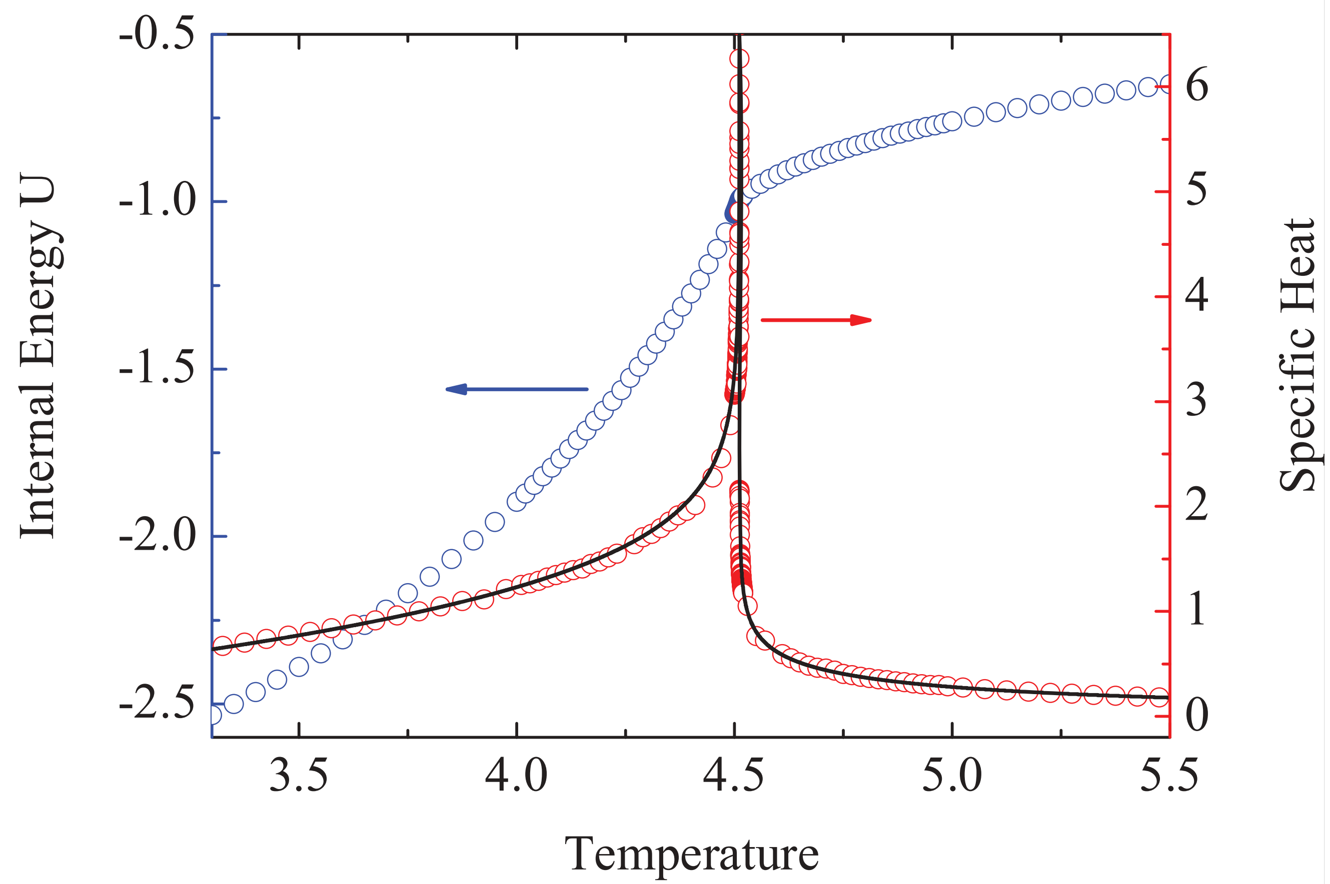
Partition function = tensor network

Z. Y. Xie *et a*l, Phys. Rev. B **86**, 045139 (2012)

Error of free energy for 2D Ising model



Energy and specific heat of 3D Ising model



$$T_c/J = \frac{2}{\ln\left(1+\sqrt{2}\right)} \simeq 2.269$$

Next week

1st: Huge data in modern physics (Today)

2nd: Information compression in modern physics

(+review of linear algebra)

3rd: Review of linear algebra (+ singular value decomposition)

4th: Singular value decomposition and low rank approximation

5th: Basics of sparse modeling

6th: Basics of Krylov subspace methods

7th: Information compression in materials science

8th: Accelerating data analysis: Application of sparse modeling

9th: Data compression: Appication of Krylov subspace method

10th: Entanglement of information and matrix product states

11th: Application of MPS to eigenvalue problems

12th: Tensor network representation

13th: Information compression by tensor network renormalization