Canonical Tensor Model through data analysis – Dimensions, topologies, and geometries –

Naoki Sasakura

Yukawa Institute for Theoretical Physics, Kyoto University

Based on a collaboration with Taigen Kawano and Dennis Obster
Phys.Rev. D97 (2018) no.12, 124061, arXiv:1805.04800

Discrete Approaches to the Dynamics of Fields and Space-Time
September 9-12, 2018,
Lecture Theater, TOKYO ELECTRON House of Creativity, Tohoku University

§ Introduction and main results

Canonical Tensor Model (CTM) — A tensor model for gravity

- Hamilton (canonical) formalism with first-class constraints.

  Mimicking the ADM formalism of GR to incorporate general covariance

$$H = n_a \mathscr{H}_a + m_{ab} \mathscr{J}_{ab} \qquad a, b = 1,2,\ldots, N$$

$\mathscr{H}_a$, $\mathscr{J}_{ab}$: "Hamiltonian" & "Momentum" constraints

n_a, m_{ab} : "Shift" & "Lapse"

- Dynamical variables: Q_{abc}, P_{abc}

  A canonical conjugate pair of real symmetric three-index tensors

$$\{Q_{abc}, P_{def}\} = \sum_{\sigma} \delta_{a\sigma_d} \delta_{b\sigma_e} \delta_{c\sigma_f} \qquad \sigma : \text{Permutations of indices}$$

The constraints (unique under some reasonable assumptions)

$$\mathcal{H}_a = \frac{1}{2}\left(P_{abc}P_{bde}Q_{cde} - \lambda\, Q_{abb}\right)$$

$\lambda = 0, \pm 1$

"Cosmological" constant

$$\mathcal{J}_{ab} = \frac{1}{4}\left(Q_{acd}P_{bcd} - Q_{bcd}P_{acd}\right)$$

: SO(N) generators

The constraint algebra — First class

$$\left\{H(n_1), H(n_2)\right\} = J\left([\tilde{n}_1, \tilde{n}_2] + 2\lambda\, n_1 \wedge n_2\right)$$

$$\{J(m), H(n)\} = H(mn)$$

$$\left\{J(m_1), J(m_2)\right\} = J\left([m_1, m_2]\right)$$

$$H(n) \equiv n_a \mathcal{H}_a \qquad J(m) \equiv m_{ab}\mathcal{J}_{ab}$$

$$(n_1 \wedge n_2)_{ab} \equiv n_{1a}n_{2b} - n_{1b}n_{2a}$$

$$\tilde{n}_{bc} \equiv n_a P_{abc}$$

Non-linearity exists

Not a genuine Lie algebra

Similar structure as ADM

depending on $g^{\mu\nu}$.

Classical correspondences to GR

- N=1 agrees with mini-superspace approximation of GR. λ plays as the cosmological constant.

Y.Sato & NS

- Taking the (formal) continuum limit

$$P_{abc} \rightarrow P_{xyz} \qquad a, b, c \in \mathbb{N} \xrightarrow{N \to \infty} x, y, x \in \mathbb{R}^D, \; x \sim y \sim z$$

Constraint algebra → Constraint algebra of ADM formalism

Y.Sato & NS

EOM → Equivalent to a GR system in Hamilton-Jacobi formalism

H.Chen, Y.Sato & NS

$$S = \int d^D x \; [2R - \frac{1}{2}(\nabla \phi)^2 - e^{-\alpha\phi} + \text{higher spins/derivatives}]$$

$$\alpha = \sqrt{\frac{6-D}{8(D-1)}}$$

CTM is straightforward to quantize

$$[\hat{Q}_{abc}, \hat{P}_{def}] = i \sum_{\sigma} \delta_{a\sigma_d}\delta_{b\sigma_e}\delta_{c\sigma_f}$$

- First-class constraints being kept consistent (no anomaly).

- Exact physical wave functions (states) can be obtained.

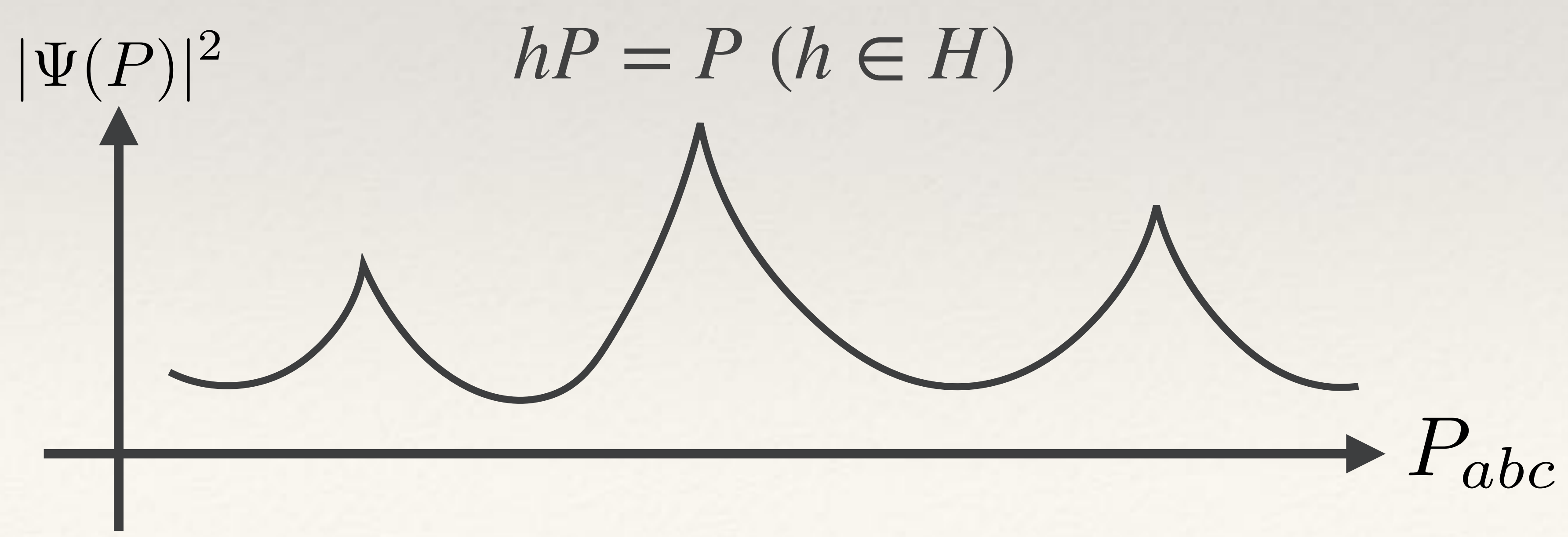
G.Narain, Y.Sato & NS

$$\hat{\mathscr{H}}_a|\Psi\rangle = \hat{\mathscr{J}}_{ab}|\Psi\rangle = 0 \rightarrow \text{Generalized Airy functions}$$

- The exact wave functions have large peaks at Lie-group symmetric configurations for $\lambda > 0$. — Emergence of Lie-group symmetries

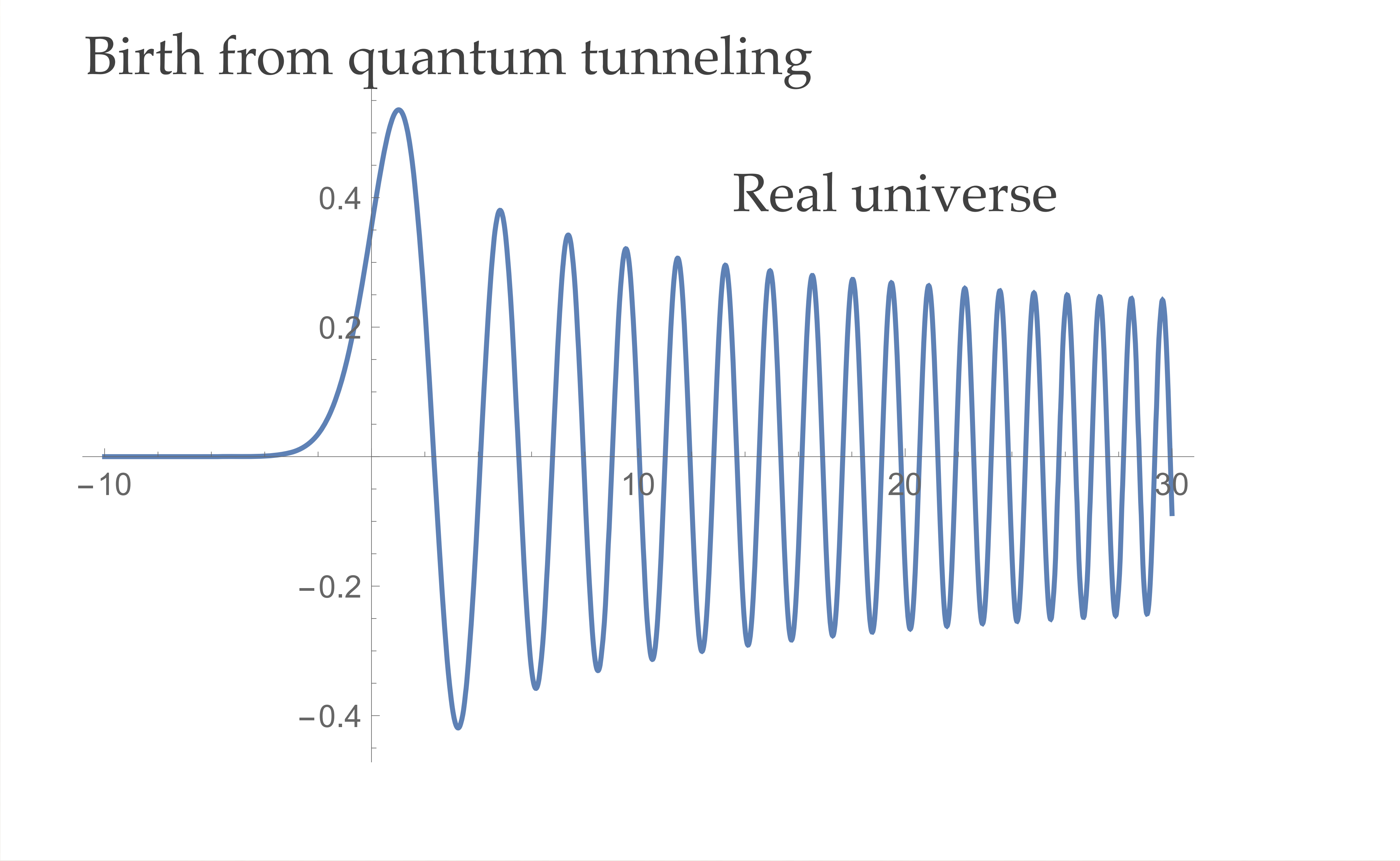
Encouraging toward spacetime emergence

D.Obster & NS



The Airy function is often used as a wave function of the universe.

AiryAi(-x)



The question we answer in this talk

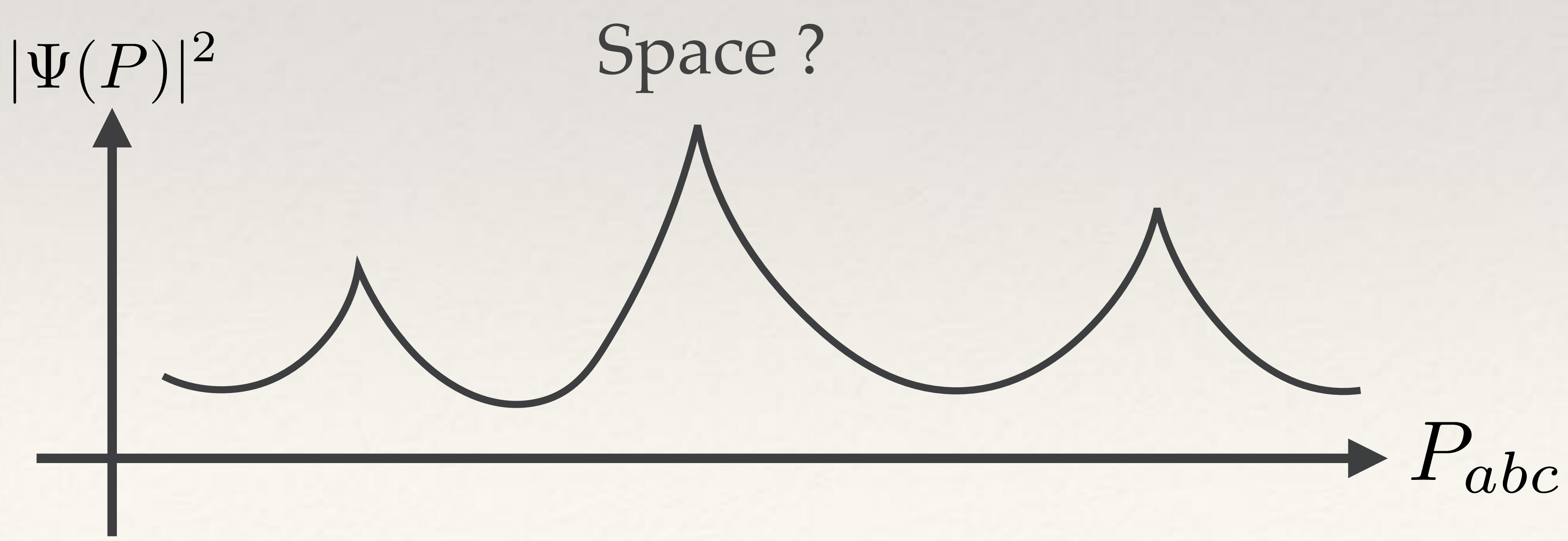
Continuum limit was a formal one in the previous works.

$$P_{abc} \rightarrow P_{xyz} \qquad a, b, c \in \mathbb{N} \xrightarrow{N \rightarrow \infty} x, y, x \in \mathbb{R}^D,\ x \sim y \sim z$$

Namely, the existence (emergence) of space was an input.

Rather we want to obtain emergent spaces as peaks of the wave functions.

Can we interpret a value of a tensor as a space ? P_{abc} = space ?



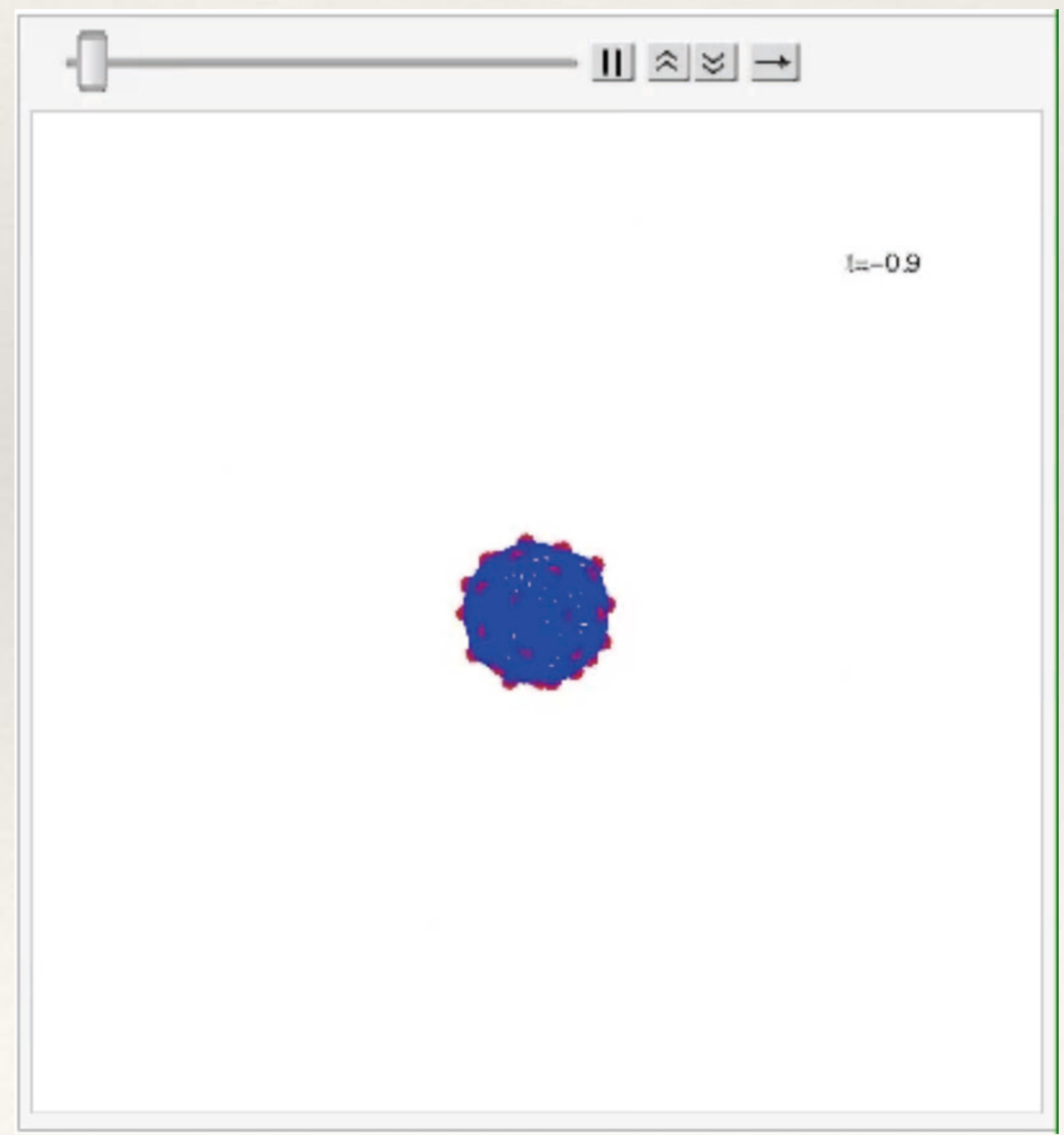
The answer

Using tensor-rank decomposition & persistent homology

— important mathematical techniques in data analysis

$P_{abc} \leftrightarrow$ a space   For finite N (but large enough)

Eg. Time-evolution of two-sphere



In classical CTM, this provides

$P_{abc}(t) \leftrightarrow$ a spacetime

Classical solutions

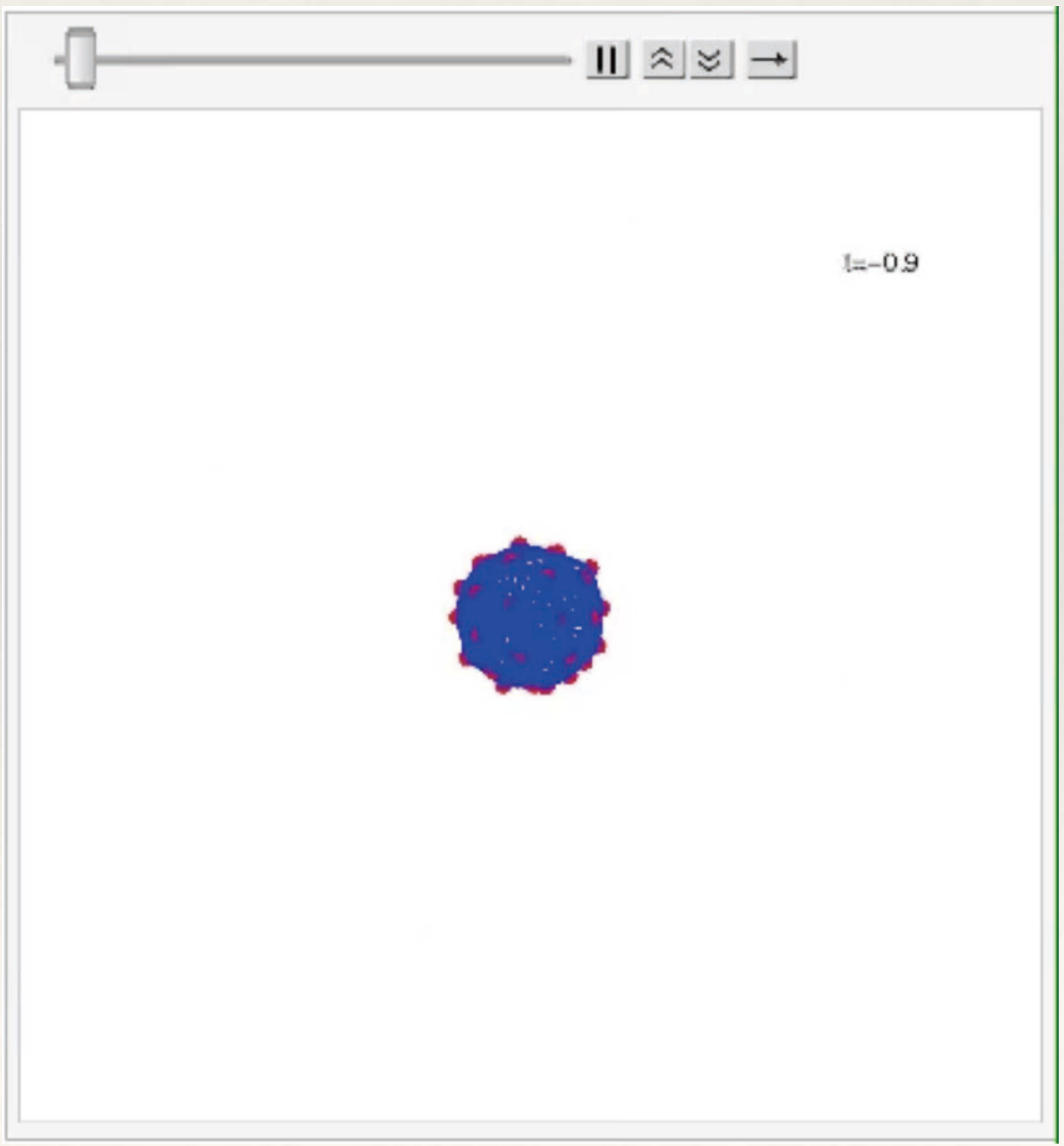
The answer

Using tensor-rank decomposition & persistent homology

— important mathematical techniques in data analysis

$P_{abc} \leftrightarrow$ a space For finite N (but large enough)

Eg. Time-evolution of two-sphere



In classical CTM, this provides

$$P_{abc}(t) \leftrightarrow \text{a spacetime}$$

↑

Classical solutions

Contents

§ Tensor-rank decomposition and the notions of point & space

§ The notion of neighborhoods

§ Global topology — Persistent homology presented by Taigen

§ Distances between points

§ Time evolutions of spaces in canonical tensor model

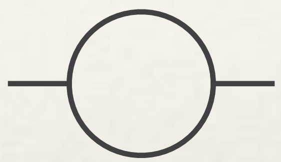
§ Summary and future prospects

§ Tensor-rank decomposition and the notions of point & space

The problem of formulating points from tensors is hard to grasp. What is the question ?
I had been wondering for several years, ending up with difficulties.

One would try the following, but they do not work :

(i) Diagonalize the matrix $M_{ab} \equiv P_{acd}P_{bcd}$



Corresponding to singular value decomposition (SVD) for tensors.
This ends up with zero modes as the most important contribution.
Non-local modes are more important. Hence one cannot find points from such a procedure, since points are local.

(ii) Diagonalize P_{abc} as much as possible: $P_{abc} = e_a \delta_{ab} \delta_{bc} + \cdots$

- No standard procedure. Non-linear problem hard to solve.
- Turns out to be essentially wrong. The procedure implicitly expects the number of points to be $\leq N$, but it is rather $O(N^2)$ as we will see.

One day I read the following chapter of a book by chance :

§9 r-term Approximation,
in “Tensor Spaces and Numerical Tensor Calculus” by W. Hackbusch.

In fact, I bought the book several years ago, but did not notice this particular chapter important for me for several years, because this book is written by a computer scientist, has about 500 pages, and contain various topics.

I wrote a simple Mathematica program to obtain the approximation for tensors, and I realize that this approximation gives the notion of points !

Now the answer

A point : the simplest tensor e.g. $P_{111} \neq 0$ Others $= 0$

Taking into account the kinematical symmetry $O(N)$ of CTM,

A point: $P_{abc} = v_a v_b v_c$ v_a : a real N-dim vector

Any points are equivalent except their sizes — Tensor analogue of equivalence principle.

A space : a collection of points

$$P_{abc} = \sum_{r=1}^{R} v_a^r v_b^r v_c^r$$

R: the rank of a tensor P for minimum R.

This is called tensor-rank decomposition of a tensor.

This decomposition of a tensor into a number of vectors is known as the tensor-rank decomposition and has a long and multi-discipline history. Seems to have been “discovered” a number of times.

1927 F. L. Hitchcock

1970 J. D. Carroll and J.-J. Chang

1970 R. Harshman

⋮

Called with different names in various fields such as MathPhys, Chemistry, Pshycometrics, Mathematics,…

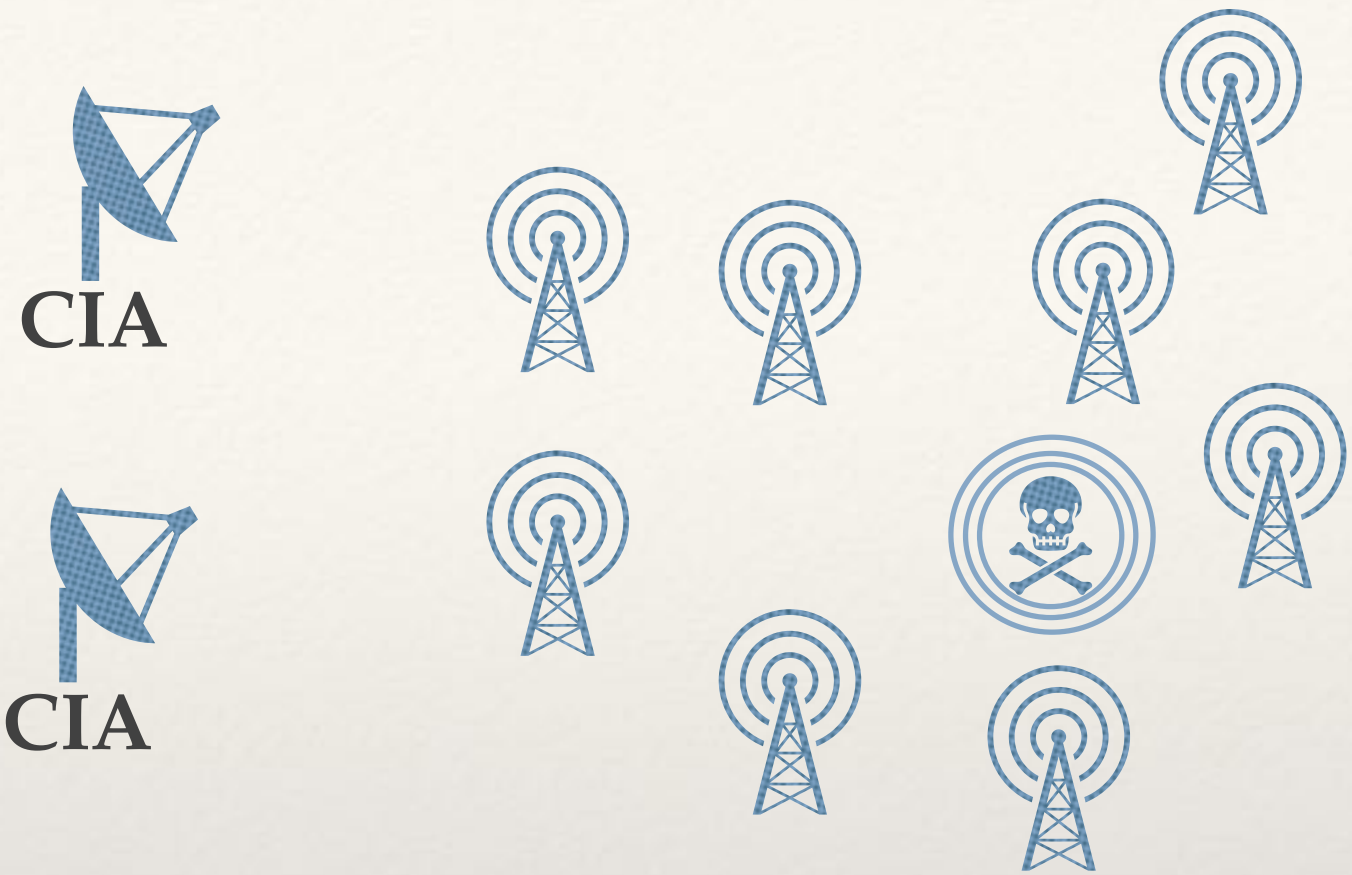
Polyadic.., Generalized Eckart-Young, Parafac, Multi-mordal factor, CP, r-representation, Rank-r decomposition, Rank-one decomposition, Waring’s problem, R-th secant variety of Serge/Veronese variety, Tensor-rank decomposition, etc.

Customarily called CP-decomposition in data analysis/computer science

We use tensor-rank decomposition as a neutral mathematical word.

A typical usage of the decomposition: Blind source separation (BSS)

Where is the terrorist ?



Observed data

Unknown sources

$$y_a(t) = v_a{}^i x_i(t) \qquad P_{abc} \equiv \langle y_a(t) y_b(t) y_c(t) \rangle_t = \sum_i v_a^i v_b^i v_c^i$$

Unknown coefficients

Tensor rank decomposition

Other usages:

Complexity of arithmetics, Fluorescence spectroscopy, Nervous system for muscle, Evaluate intelligence, Geophysics, Analysis of MRI data, Data-mining,…

Recommended references

Easy introductory review:

P. Comon, “Tensors: a Brief Introduction,” IEEE Signal Processing Magazine 31 no. 3, (May, 2014) 44–53. https://hal.archives-ouvertes.fr/hal-00923279.

More comprehensive reviews:

P. Comon, X. Luciani, and A. L. F. de Almeida, “Tensor decompositions, alternating least squares and other tales,” Journal of Chemometrics 23 no. 7-8, 393–405. https://onlinelibrary.wiley.com/doi/pdf/10.1002/cem.1236.

P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, “Symmetric tensors and symmetric tensor rank,” SIAM Journal on Matrix Analysis and Applications 30 no. 3, (2008) 1254–1279, https://doi.org/10.1137/060661569.

Mathematically rigid comprehensive book:

Landsberg, J. M., Tensors: Geometry and Applications. American Mathematical Society, Providence, 2012.

The properties of the tensor-rank decomposition $P_{abc} = \sum_{r=1}^{R} v_a^r v_b^r v_c^r$

Essentially different from the matrix case $M_{ab} = \sum_{r=1}^{R} v_a^r v_b^r$

In the matrix case,

- 😃 No essential differences between real and complex cases
- 😃 $R=N$ with probability one for randomly given M_{ab}.
- 😃 There is a systematic procedure: Eigenvalues/vectors.
- 😥 The decomposition is not unique : $v_a^r \to L^r{}_{r'} v_a^{r'}$, $L \in O(N)$

In the tensor case, on the other hand,

- 😃 The decomposition is essentially unique (No degeneracies like above). The basic reason for the ability of BSS.
- 😫 The formulas for general ranks (ranks with finite probabilities) are only known for limited cases, such as the complex symmetric case. Unknown for real cases in general situations (only in special cases).
- 😱 A standard systematic procedure for the decomposition is unknown.

Alexander-Hirschowitz theorem (1995)

The general rank (rank with probability one) of a symmetric complex tensor with w indices is given by

$$R_g(w,N) = \left\lceil \frac{1}{N}\binom{N+w-1}{w} \right\rceil$$

with the following exceptions :

$$R_g(2,N) = N$$

$$R_g(w,N) = \left\lceil \frac{1}{N}\binom{N+w-1}{w} \right\rceil - 1 \quad \text{for } (w,N) = (3,5),(4,3),(4,4),(4,5)$$

The actual practical procedure — Numerical optimization

1. Suppose P_{abc} is given.
2. Assume a value of R.
3. Then solve

$$\min_{v_a^r} \left| P_{abc} - \sum_{r=1}^{R} v_a^r v_b^r v_c^r \right|^2$$

4. If the remaining error is too large, change R to larger values and repeat.

R cannot be made arbitrary large, because the minimization problem becomes ill-defined.
($v_a^r v_b^r v_c^r$ and $v_a^{r'} v_b^{r'} v_c^{r'}$ tend to cancel with each other).

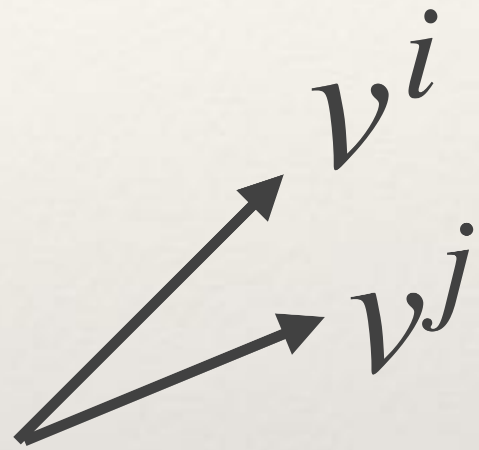
We made a Mathematica and a c++ program for the optimization.
(Some Matlab programs by others are also available on the net.)

§ The notion of neighborhood

From the tensor-rank decomposition

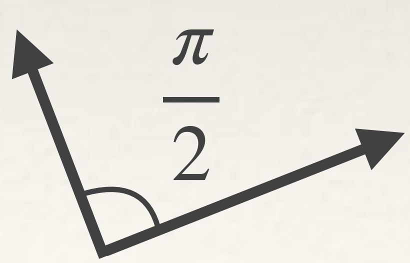
A point $i \quad = v_a^i \;\in\; \mathbb{R}^N$

It is natural to define neighboring points by



$$\mathcal{N}(i, c) = \{ v^j \mid v^i \cdot v^j > c \}$$

The value c determines the range of a neighborhood.
Large c, smaller neighborhood, and vice versa.



Uncorrelated independent points, if $v^i \cdot v^j = 0$

§ Fuzzy spaces

It is not true that any tensor P_{abc} corresponds to a continuous space through the tensor-rank decomposition. Generally it is just a collection of points, not corresponding to any continuous spaces.

However, there is a systematic procedure to construct a tensor describing a continuous space with any dimension, topology, and geometry. This will be presented by Taigen.

Rather than describing the general procedure (which is in our paper), let us restrict ourselves to spheres with any dimensions in this talk.

Fuzzy two-sphere S²

Spherical harmonics

Essentially,

$$P_{(l_1m_1)(l_2m_2)(l_3m_3)} \sim \int_{S^2} d\omega\ Y_{l_1m_1}(\omega)Y_{l_2m_2}(\omega)Y_{l_3m_3}(\omega)$$

Cut-off for the modes

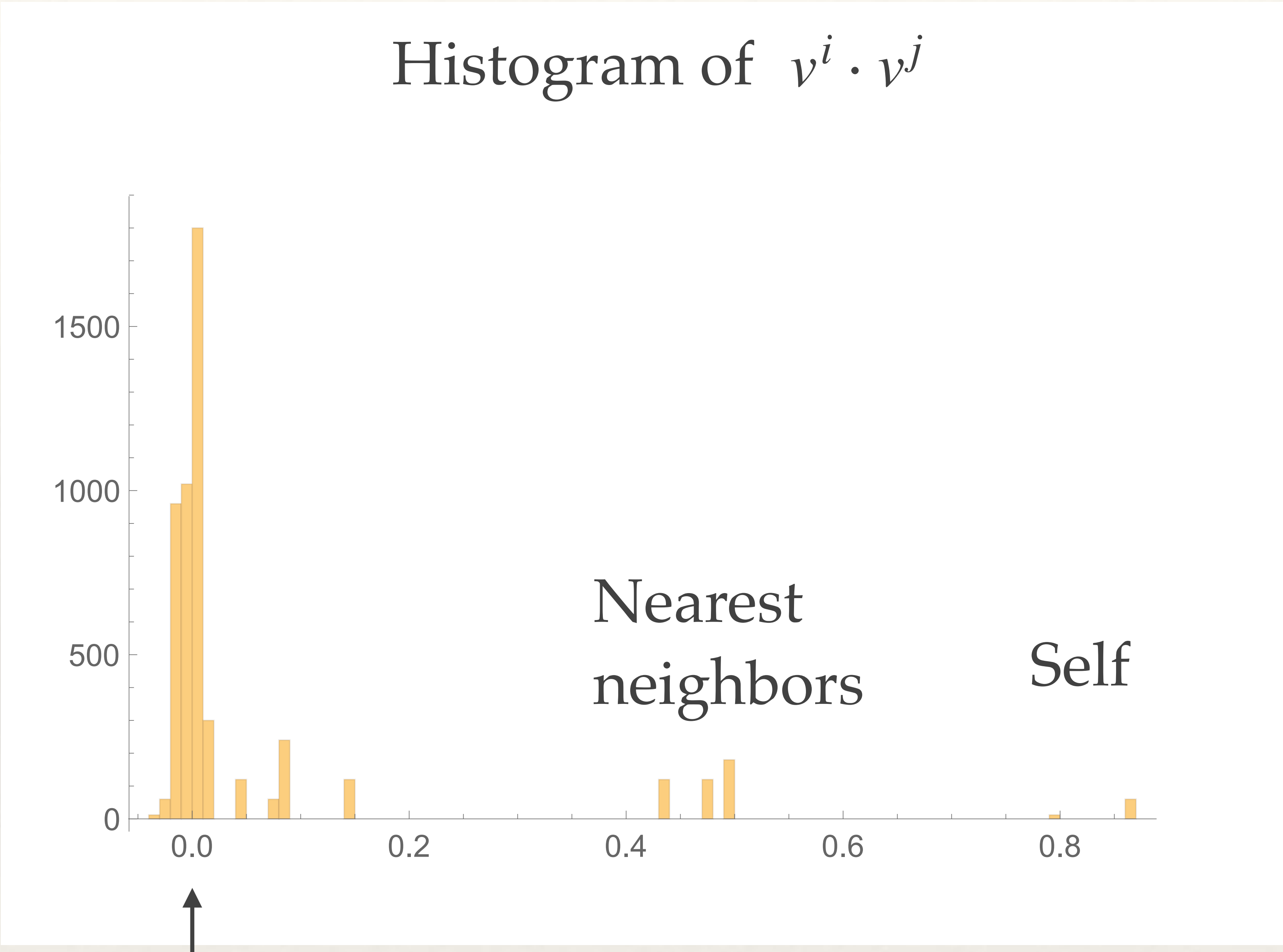
More precisely, $a = (l, m) \quad (|m| \leq l \leq L)$

$$P_{(l_1m_1)(l_2m_2)(l_3m_3)} = \int_{S^2} d\omega\ \tilde{Y}_{l_1m_1}(\omega)\tilde{Y}_{l_2m_2}(\omega)\tilde{Y}_{l_3m_3}(\omega)$$

$$\tilde{Y}_{lm} = \begin{cases} \frac{1}{\sqrt{2}}\left(Y_{lm} + Y^*_{lm}\right) e^{-l^2/L^2} & (m > 0) \\ Y_{l0}\, e^{-l^2/L^2} & (m = 0) \\ \frac{1}{\sqrt{2}i}\left(Y_{lm} - Y^*_{lm}\right) e^{-l^2/L^2} & (m > 0) \end{cases}$$

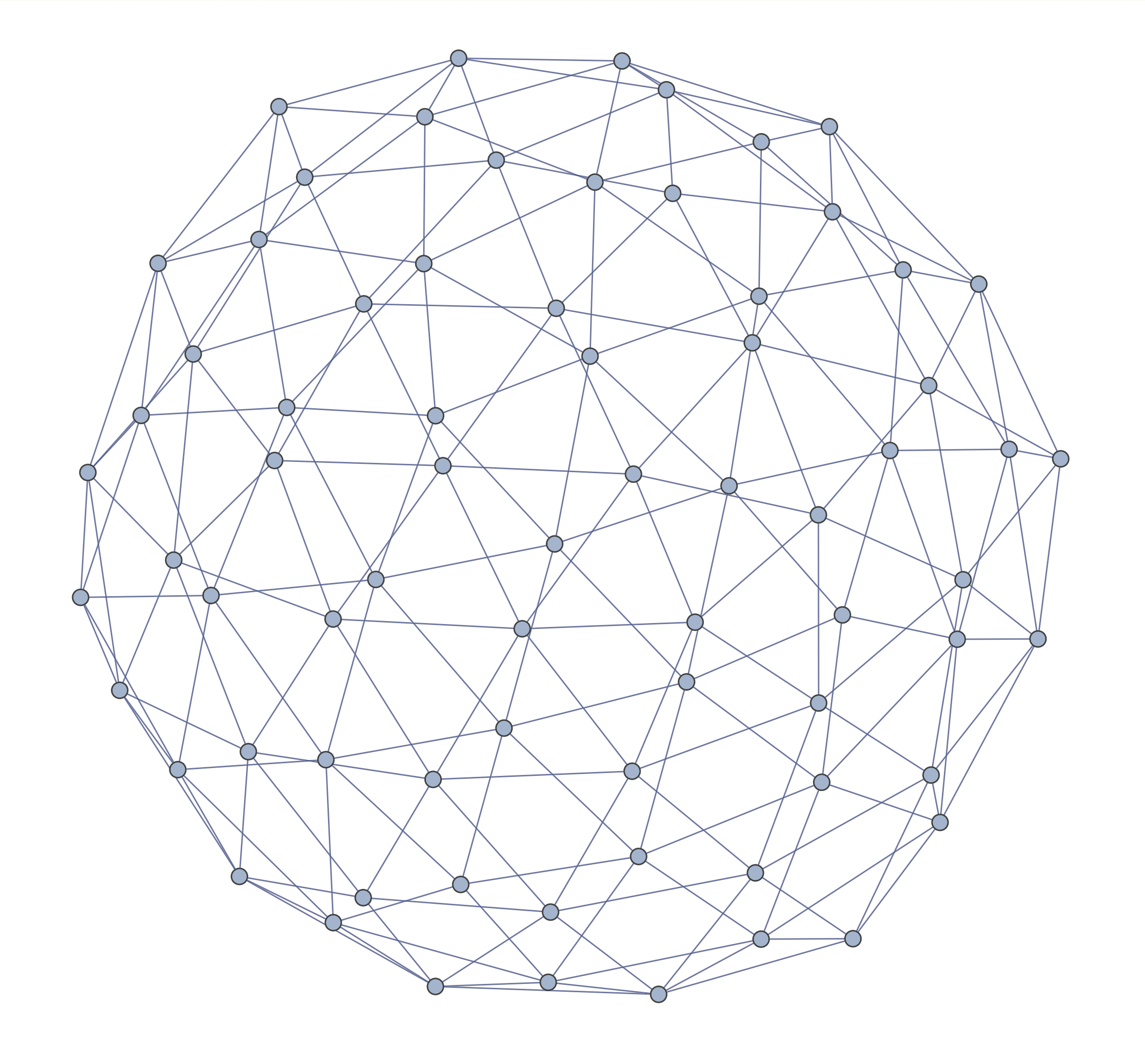
Regulator to smooth the cutoff

Fuzzy 2-sphere L=5 R=72



Most of them are mutually independent.

Eg. L=5 (N=36), R=72 Edges drawn if $v^i \cdot v^j > 0.2$



The construction can straightforwardly be extended to other cases.

- Any dimensional fuzzy spheres

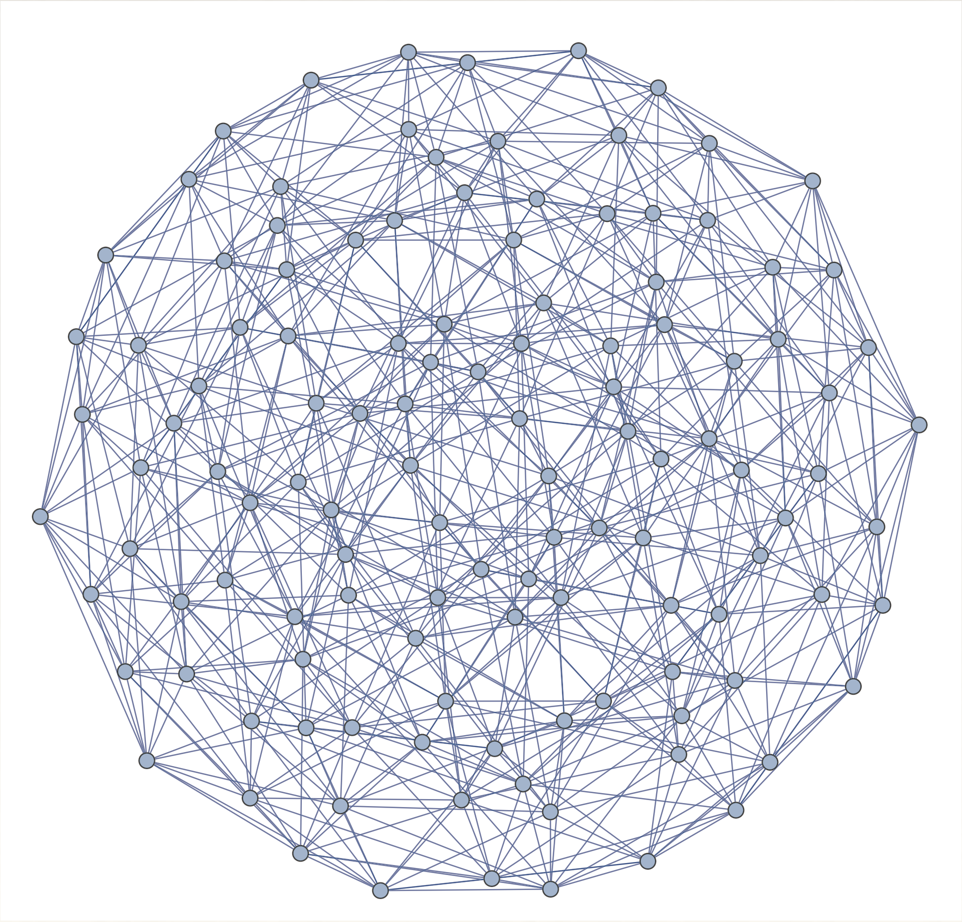
n-dimensional spherical harmonics $Y_{l_1 l_2 \ldots l_n}(\theta_1, \theta_2, \ldots, \theta_n)$

$$|l_1| \le l_2 \le \cdots \le l_n \le L$$

$$P_{(l_1 \ldots l_n)(l'_1 \ldots l'_n)(l''_1 \ldots l''_n)} = \int_{S^n} d\omega \, \tilde{Y}_{l_1 \ldots l_n}(\omega) \, \tilde{Y}_{l'_1 \ldots l'_n}(\omega) \, \tilde{Y}_{l''_1 \ldots l''_n}(\omega)$$

Eg. Fuzzy three-sphere

N=55, R=120



Really ?

How can we be sure ?

Concerning this question, one day I read an article by chance, and was deeply impressed :

物理学会誌　2017年9月号

ランダムの中に見る秩序—パーシステントホモロジーとその応用—

平岡裕章，西浦廉

(Order in Disorder: Persistent Homology and Its Applications
by Yasuaki Hiraoka and Yasumasa Nishiura
In BUTSURI, September volume, 2017)

Persistent homology can be used to recognize global topologies of point sets with mutual distances.

This will be presented by Taigen.

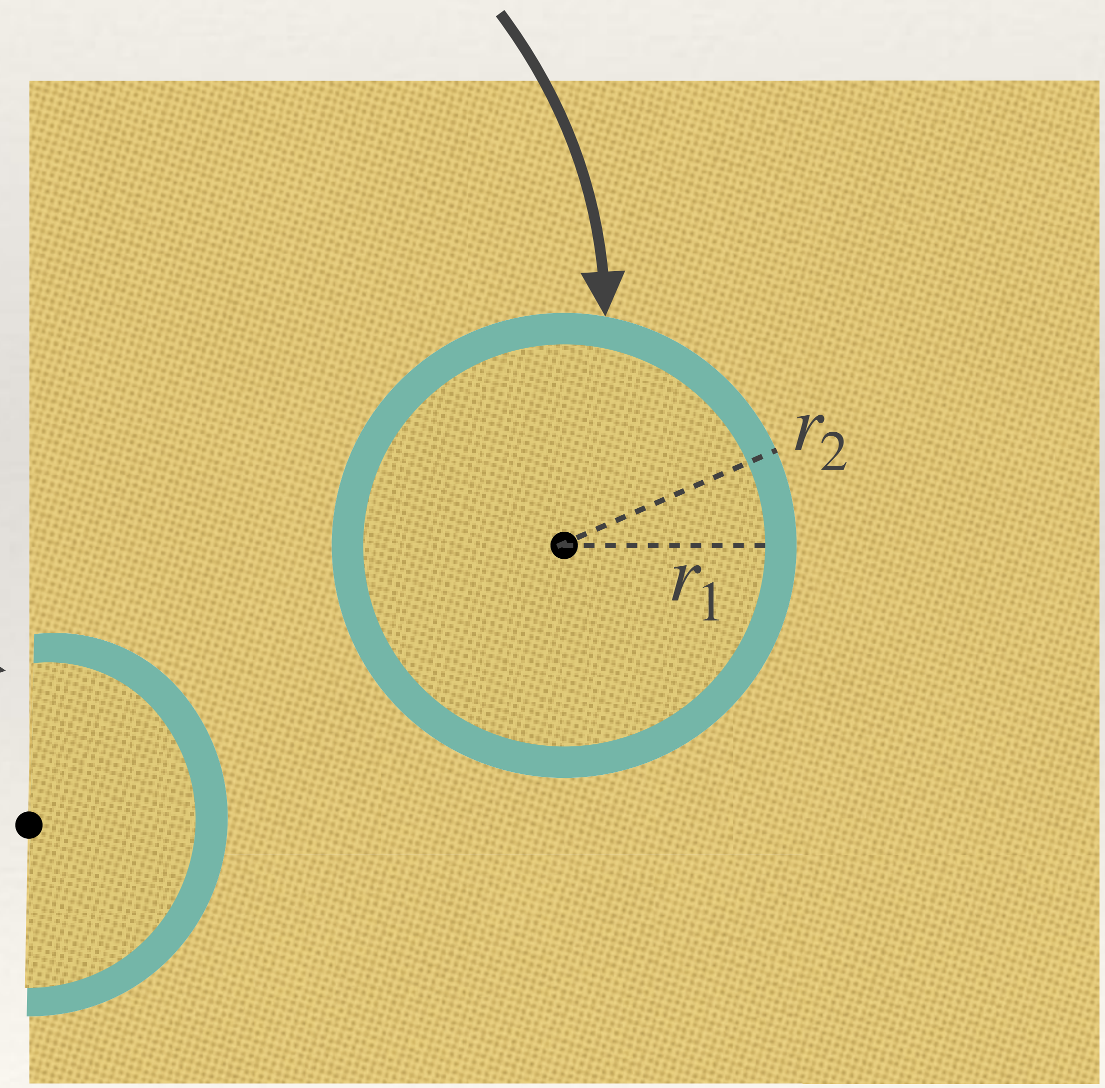
A comment:

Persistent homology can also be used to

- determine the topological dimension of data points
- distinguish inner and boundary data points

The shell region $r_1 < r < r_2$ of an inner point should be S^{D-1}, if the base space has D dimensions.

The shell region of a data point on a boundary should have the topology of a disk.



Recommended references :

Quick practical reference :

JAVAPLEX TUTORIAL

H. Adams and A. Tausz

http://www.math.colostate.edu/%7Eadams/research/javaplex_tutorial.pdf

Mathematical review by the founder :

TOPOLOGY AND DATA

G. Carlsson

Bulletin (New Series) of the American Mathematical Society, Vol 46, Number 2, April 2009, Pages 255–308

By using persistent homology, the realization of the three-sphere can be checked.

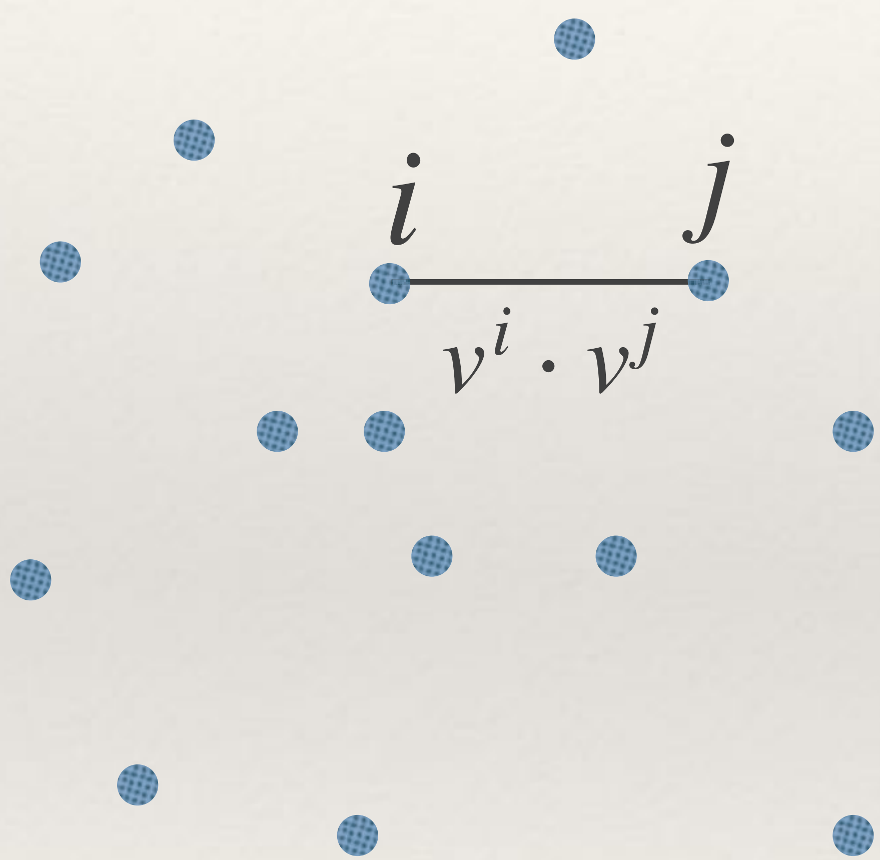
The construction and checking can straightforwardly be generalized to

n-Torus, Strip, Mobius strip, Klein bottle,…

This will be presented by Taigen.

§Distances between points

More detailed distances between points (not just neighborhoods) can be determined by considering virtual diffusion process over points. This is also commonly used in data analysis.



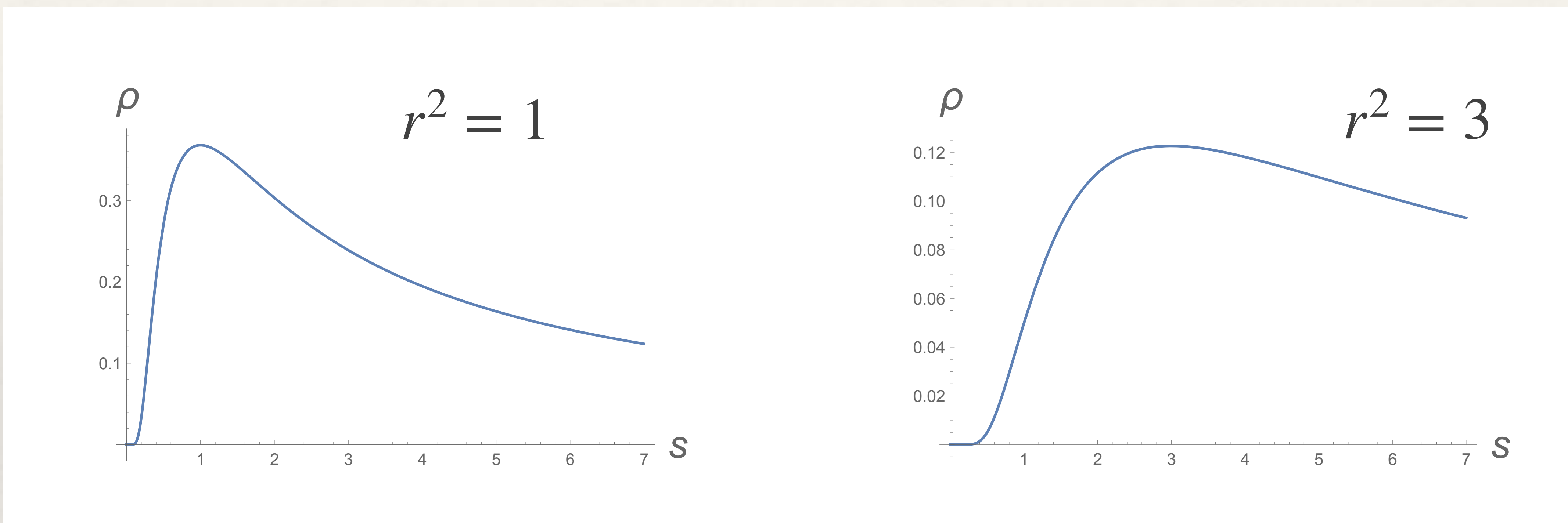
$$\frac{d}{ds}\rho_i(s; i_0) = K_{ij}\rho_j(s; i_0)$$

$$\rho_i(0; i_0) = \delta_{ii_0}$$

K_{ij} should be taken larger for larger $v^i \cdot v^j$.

$$r = (s_{peak})^2$$

$$\rho(s, r) \sim s^{-D/2} \exp(-r^2/s)$$



The canonical tensor model (CTM) motivates the following form of K:

$$K_{ij} = \beta(i)^{-1}\beta(j)^{-1} w^i \cdot w^j - \delta_{ij} \sum_k \beta(k)^{-2}$$

$$P_{abc} = \sum_i \frac{1}{\beta(i)^2} w_a^i w_b^i w_c^i$$

$$\sum_j w^i \cdot w^j \beta(j)^{-2} = 1$$

Solving for w^i and $\beta(i)$

The complication from β is because CTM contains a scalar field.

The derivation is skipped in this talk. See our paper, if you are interested.

§ Time evolutions of spaces in CTM

We have argued P_{abc} ~ A space with geometric structure

EOM of CTM

$$\frac{d}{dt}P_{abc} = \{P_{abc}, H\} = -n_d P_{dae} P_{ebc} - n_d P_{dbe} P_{eca} - n_d P_{dce} P_{eab} + \text{shift}$$

CTM defines $P_{abc}(t) \to$ a spacetime with geometry

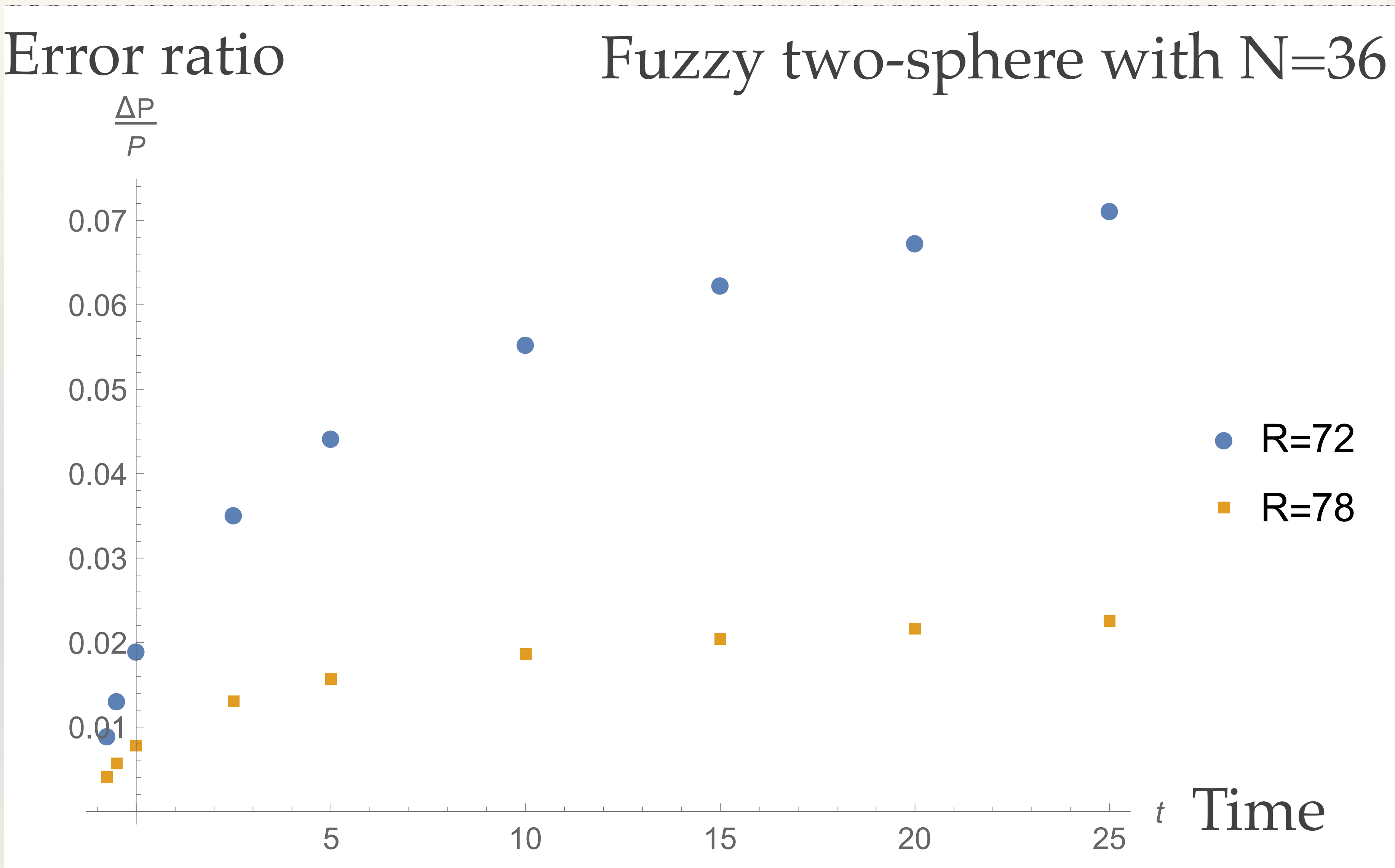
The claims we will show are

(i) Time evolution increases the number of points forming a space.

(ii) Time evolutions of homogeneous spheres agree with the EOM of the GR system previously obtained.

(i) Time evolution increases the number of points.

As time evolves, more points (larger rank R) are necessary to suppress the error of the tensor-rank decomposition. This roughly means that the complexity of the space evolves to necessitate more points to describe it.



$$\Delta P^2 \equiv \left(P_{abc} - \sum_{i=1}^{R} v_a^i v_b^i v_c^i\right)^2$$

(ii) Time evolutions of the homogeneous spheres agrees with the EOM of the GR system previously obtained.

The EOM of the GR system tells

Ignored for homogenous spaces

$$\frac{d}{dt}\beta = -9n\beta^2 + \text{derivative \& higher terms}$$

$$\frac{d}{dt}\beta^{\mu\nu} = -15n\beta\beta^{\mu\nu} + 2n\beta^{\mu\mu'}\beta^{\nu\nu'}R_{\mu'\nu'} + \text{derivative \& higher terms}$$

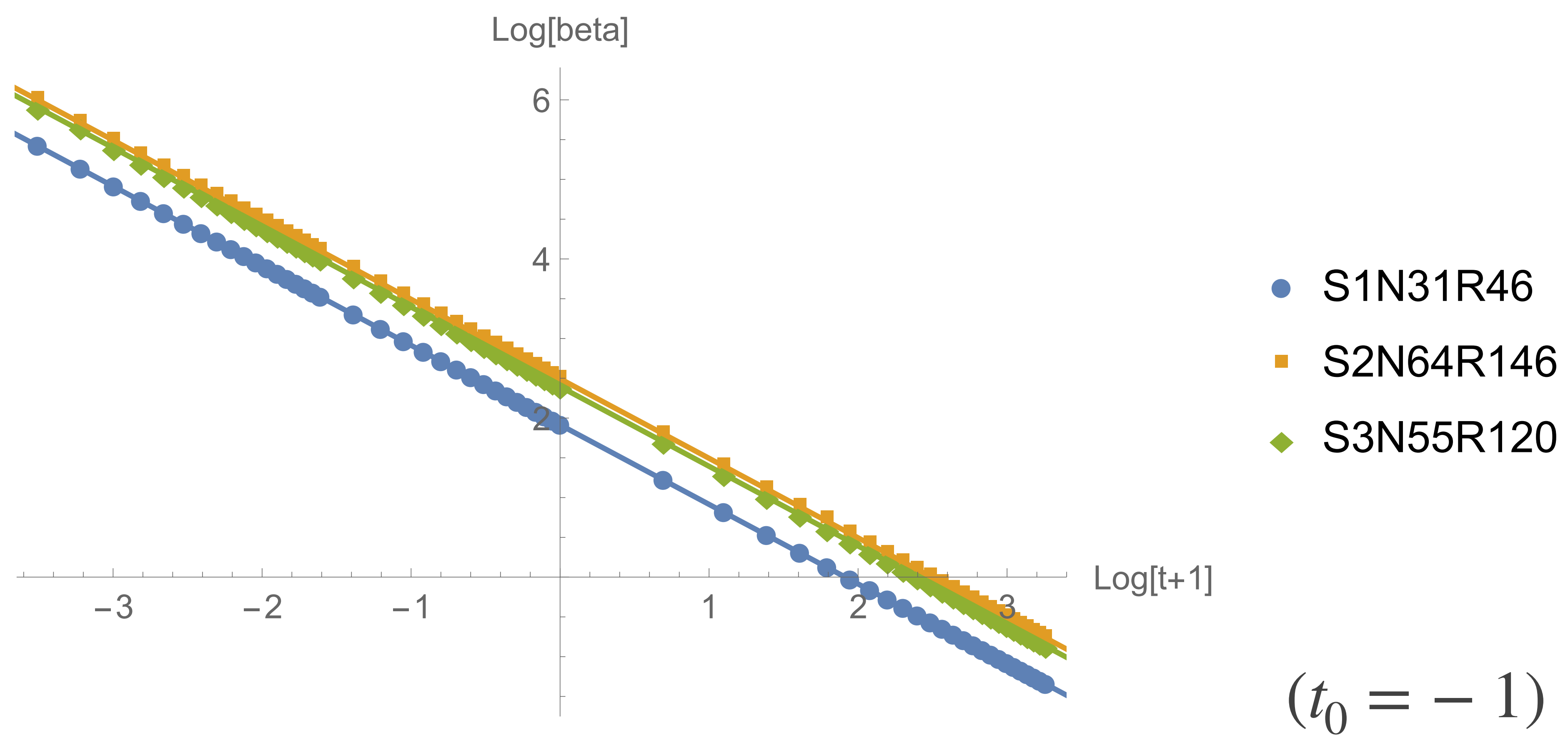
Constant curvature for homogenous spaces

$$\beta(t) = \frac{1}{t-t_0}$$

$$\beta^{\mu\nu}(t) \propto (t-t_0)^{-\frac{5}{3}}\left(1 + c_R(t-t_0)^{-\frac{2}{3}}\right)^{-1}$$

We compared these for the homogeneous fuzzy $S^{1,2,3}$ through the procedure using the virtual diffusion process.

$\langle \beta(i,t) \rangle_i$ compared with $\beta(t)$

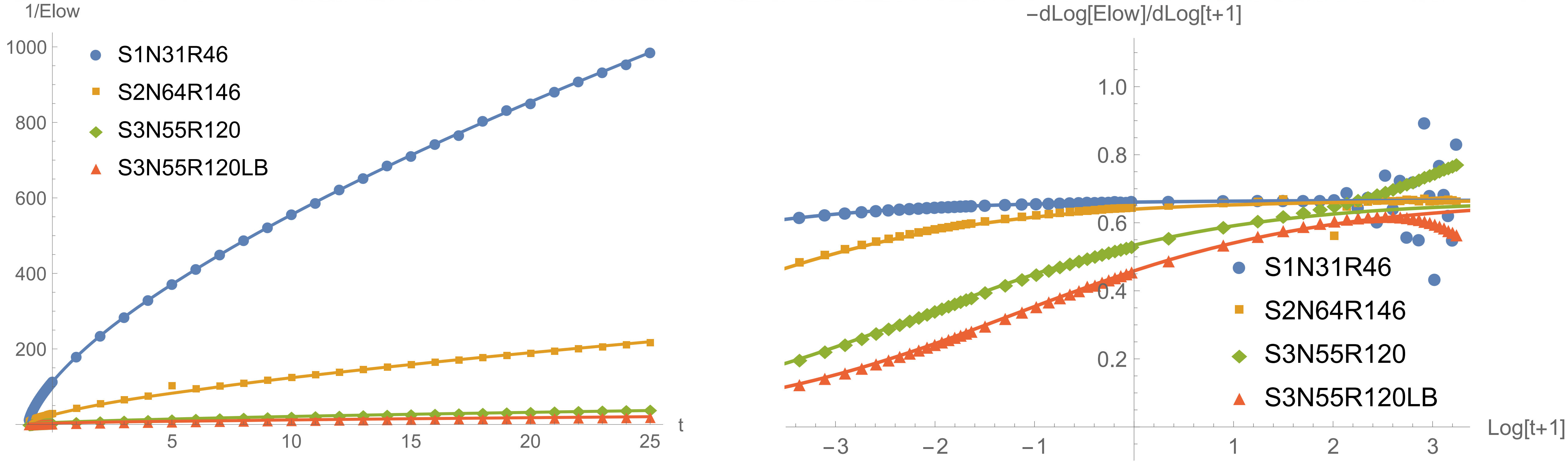


$(t_0 = -1)$

Perfect matches

The inverse of the lowest eigenvalue of $-K_{ij}(t)$ compared with $\beta/\beta^{\mu\nu}(t)$

~ The square sizes of the spaces



Quite a good match except for large time regions, where the sizes of the spaces become too large in comparison with the numbers of points.

(i.e. The distances between points become macroscopic, and continuum description becomes invalid.)

§ Summary and future prospects

Tensor rank-decomposition and persistent homology, well known important techniques in data analysis, and virtual diffusion process provide the spacetime interpretation of the canonical tensor model (CTM).

The initial investigation of the EOM of CTM for the homogenous fuzzy spheres $S^{1,2,3}$ agreed completely with the GR system previously obtained in a formal continuum limit of CTM.

We may expect interdisciplinary researches in the future.

Canonical tensor model

Classical/Quantum gravity — Data analysis

Spacetime — Data

A microscopic description of spacetime