DLAP 第38回 文献紹介用ノート　　　橋本幸士 2022.3.10

発表論文　Jonathan Lam, Yi-Zhuang You
"Machine Learning Statistical Gravity from Multi-Region Entanglement Entropy"
arXiv:2110.01115 (hep-th)

元になる論文. Yi-Zhuang You, Zhao Yang, Xiao-Liang Qi
"Machine Learning Spatial Geometry from Entanglement Features"
arXiv: 1709.01223 (cond-mat.dis-nn)

目次

§1. 動機：AdS/CFTとbulk reconstruction

AdS/CFT対応. (Maldacena, hep-th/9711200
Witten, hep-th/9803131
Gubser, Klebanov, Polyakov hep-th/9802109)

CFT側　　　　　　　　重力側.

$$\int \mathcal{D}\phi \; e^{-S_{CFT}[\phi] - \int J(x)\,\mathcal{O}[\phi]} = \int \mathcal{D}g\,\mathcal{D}\Phi \; e^{-S_{grav}[g,\Phi]}$$

$$\left(\mathcal{O}[\phi] \leftrightarrow \Phi \qquad \Phi(r=\infty) \sim J(x). \right)$$

↓ large N　　　　　　　↓ 古典極限

large N Generating Functional $= \int \mathcal{D}\Phi \; e^{-S_{grav}[\tilde{g},\Phi]}$

$\tilde{g}$：古典背景,

or $e^{-S_{grav}[\tilde{g},\tilde{\Phi}]}$

$\tilde{\Phi}$：古典解.

例： $S_{CFT} = S_{\mathcal{N}=4\ SYM}$, $S_{grav} = S_{IIB\ SUGRA}$, 解 $AdS_5 \times S^5$
(量子なら IIB string)

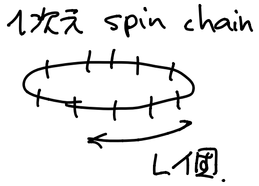
- 通常, AdS/CFTのcheckには, 重力側からstartして. 古典背景時空を解いたりして. それをCFTへの予言とする。
  (例：holographic QCD)
- 一方, CFTやQFTを与えて, 重力側を構築することをbulk reconstructionという

Reconstruction methods

- bulkのEOMを解く (例. holographic renormalization)
- bulkのextremal surfaceからgeometryを解く
  (例. Entanglement entropy, Correlator, Wilson loop, Complexity)
- Machine learning. --- この論文.

加えて. 我々の論文 1802.08313
1809.10536
1903.04951
⋮

§2 この論文の内容と結果



1) L個の qubit 系の ground state を考える.

$$|\psi\rangle = \sum_{s_1} \psi(s_1, \cdots, s_L)\,|s_1, \cdots, s_L\rangle$$

$$H = -\sum_{i=1}^{L} i\, u_i\, \chi_i \chi_{i+1}$$

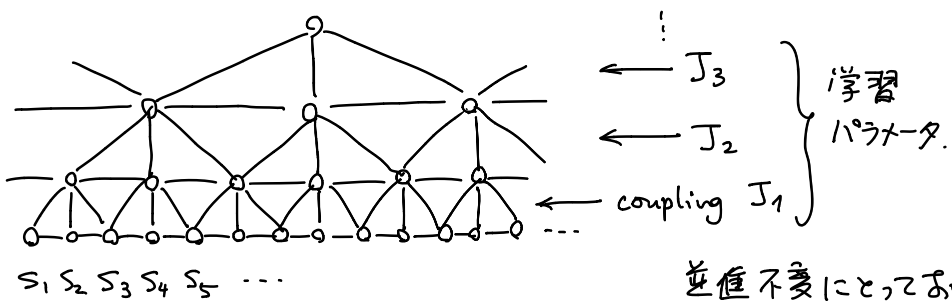
$$u_i = 1 + m(-1)^i$$

χ_i : Majorana fermion

注) Jordan-Wigner 変換を用いて fermion に直した.

2) 領域 A と $\overline{A}$ に分けて, 2nd Renyi entropy $S_E^{(2)}(A)$ の data を集める.

3) その data を再現できるような 深層ボルツマンマシンを学習.



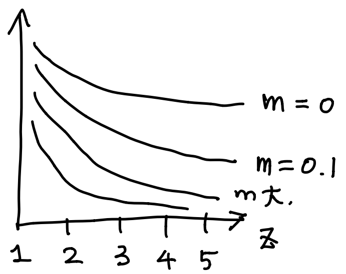
並進不変にとっておく.

① そのまま J_z ($z = 1, 2, 3, \cdots$) を学習 arXiv:1709.01223

② $P[J_z]$ (J の確率分布) を学習 arXiv:2110.01115

○ 結果①

- L=64
- 5層.
- connected A.

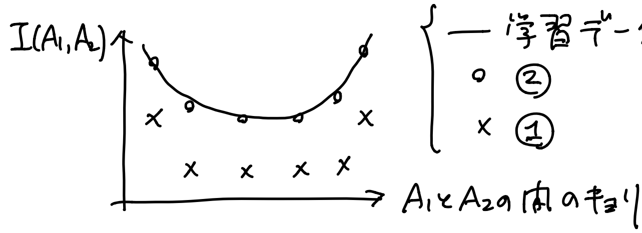


- おおよそAdS的なgeometryが m=0 で出現した.
- 質量mが大きくなると, AdSの内側の時空が順次消えていく.

⇒ AdS/CFTと consistent.

ただし mutual information $I(A_1, A_2) = S(A_1) + S(A_2) - S(A_1 \cup A_2)$ は合わない.

○ 結果② mutual information も学習可能.

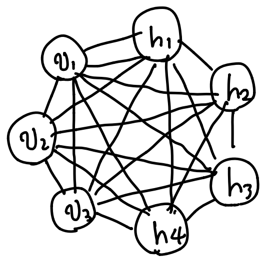


— 学習データ
○ ②
× ①

注) ①と異なり dual Ising model に bulkで移り,
S_E を bulk scalar theory の correlator で評価している.

§3 機械学習とボルツマンマシン

◎ ボルツマンマシン：確率分布を学習するマシン（プログラム）



$$\begin{cases} v_i : \text{visible unit} & v_i = \pm 1 \\ h_i : \text{hidden unit} & h_i = \pm 1 \end{cases}$$

出力 $P_J(v_1, v_2, v_3) = \sum_{\{h_i\}} \exp\left(-\mathcal{E}(v_i, h_j)\right)$

$$\mathcal{E}(v_i, h_j) \equiv \sum_{ij} J_{ij} v_i v_j + \sum_{ij} \tilde{J}_{ij} v_i h_j + \sum \tilde{\tilde{J}}_{ij} h_i h_j$$

$J_{ij}, \tilde{J}_{ij}, \tilde{\tilde{J}}_{ij}$ ：学習パラメータ.

この関数 P_J は確率分布を表す.

入力：(v_1, v_2, v_3)　　出力：$P_J(v_1, v_2, v_3)$ （(v_1, v_2, v_3) が発生する「確率」）

パラメータに依存

学習方法

1) 実験値 (v_1, v_2, v_3, P_{ex}) の set をたくさん用意する

2) $P_{ex}(v_1, v_2, v_3)$ に $P_J(v_1, v_2, v_3)$ に近くなるように J を変更する（「学習」）

誤差関数 $E \equiv D_{KL}(P_{ex} \| P_J)$　　"Relative entropy" "Kullback-Leibler divergence"

$$\equiv \sum_{v_i} \left(P_{ex} \log P_{ex} - P_{ex} \log P_J \right)$$

注) $E \geq 0$ を示すことが出来る.

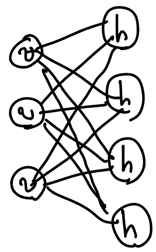
注) $E = 0 \iff P_{ex} = P_J$.

・学習は steepest-descent を用いる.

$$J_{new} = J_{old} - \epsilon \left.\frac{\partial E}{\partial J}\right|_{J_{old}}.$$

ϵ：学習の速さを制御するパラメタ.

・学習をうまくやるためにnetworkを修正する方法がある.

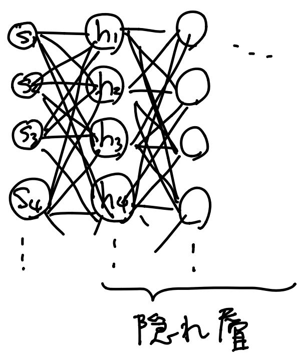


制限ボルツマンマシン.

◎ ボルツマンマシンの物性物理への応用

Carleo, Troyer (2017)

$$\left.\begin{aligned} \psi(s_1, s_2, \cdots s_L) &= P_J(s_1, \cdots, s_L) \\ \text{誤差関数} &= \langle\psi|\hat{H}|\psi\rangle. \end{aligned}\right\}$$ とし J を学習.



・隠れ層の数を多くしていくと基底状態のエネルギーが下がる

・従来のtensor networkよりエネルギーが下がる.

(2次元ハイゼンベルグ強磁性模型)

・物理的解釈（今田 et al 2018）

隠れ層方向を Euclidean time と考え，bond (weight) を Hamiltonian の時間発展と考えると.

$$|\psi\rangle_{\text{ground state}} = \lim_{\tau\to\infty} e^{-\tau H}|\text{any}\rangle$$

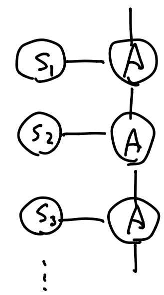
§4 ネットワークと物性物理

1) テンソルネットワーク：

多体 H の Schrödinger 方程式を解くための ansatz

・Matrix product state

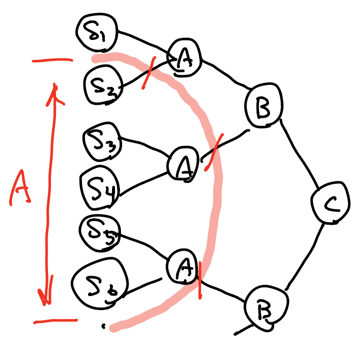
$$\psi(s_1,\cdots,s_L) = \mathrm{Tr}\left(A^{(s_1)}A^{(s_2)}\cdots A^{(s_L)}\right)$$



A: $D\times D$ matrix

D: bond dim.

・Tensor network state



$$\psi(s_1,\cdots,s_L) = C_{ij}\, B_{ikl}\, B_{jmn}\, A_{ks_1s_2}\, A_{ls_3s_4}\, A_{ms_5s_6}\cdots$$

この時、entanglement entropy は 領域 A の長さ γ_A により $S_E \sim \log \gamma_A$ と与えられるが、それは領域の端と端をつなぐ線が tensor net を何回切るかできまる。

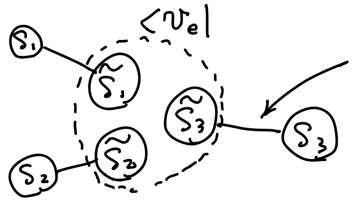
(線一本 ～ EPR pair)

(matrix product state の時は γ_A によらない。)

※ AdS/CFT との関係 (Swingle, 2010)

$A = B = C$ となる時、network は hyperbolic space の離散幾何. = AdS の time slice

・Random tensor network. (RTN)



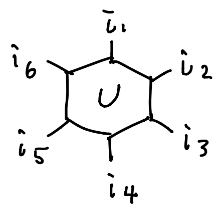
EPR pair : $|s_1\tilde{s}_1\rangle^{EPR} \equiv \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right)$

$$\psi(s_1, s_2, s_3) = \underbrace{\langle v_e|}_{\langle\tilde{s}_1|\otimes\langle\tilde{s}_2|\otimes\langle\tilde{s}_3|\text{ の基底上の状態}}\left(|s_1\tilde{s}_1\rangle^{EPR}\otimes|s_2\tilde{s}_2\rangle^{EPR}\otimes|s_3\tilde{s}_3\rangle^{EPR}\right)$$

。このように tensor network を定める時、
RTN とは $\langle v_e|$ として random な状態をとったものを言う。(それについて和をとることをしてもよい)

。こうとりたい理由：
random tensor はおおよそ完全 tensor の性質を兼ねるから、AdS/CFT の toy model になる。

注) 完全 tensor の定義

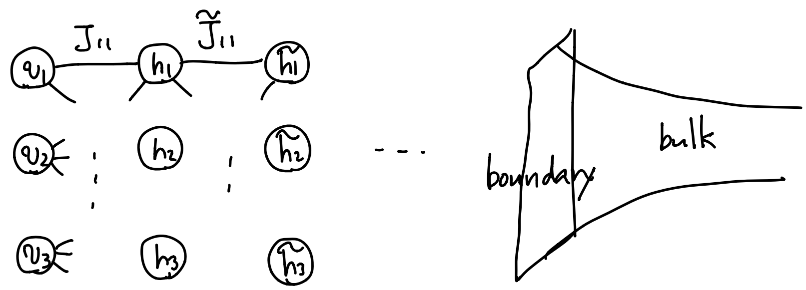


$i_1 \sim i_6$ を3つの入力と3つの出力に任意に分けた時いつも U がユニタリー。

注) HAPPY code (1503.06237)：完全 tensor で bulk の情報と boundary の情報を等価にできる。

。ランダムネスについて和をとる $\cong$ bulk で経路積分する。
$\langle v_e|$ が bulk の場と考えることが出来る。

2) AdS/CFT とボルツマンマシンの関係.



ボルツマンマシン $\longleftrightarrow$ AdS/CFT

$$P_J(v_i) \equiv \sum_{\{h_i\}} e^{-\mathcal{E}(v_i,h_i;J)} \longleftrightarrow Z_{bulk}(\phi(r=\infty)) \equiv \int \mathcal{D}\phi\, e^{-S_{grav}(\phi;g)}$$

背景 g は fix.

J も重みづけすると

$$P(v) \equiv \int dJ\, P_J(v_i)\, P[J] \longleftrightarrow Z_{bulk} \equiv \int \mathcal{D}g\,\mathcal{D}\phi\, e^{-S_{Einstein}-S(\phi;g)}$$

背景 g も積分.

ここで. $\mathcal{E}(v_i, h_i; J)$ と $S_{grav}(\phi; g)$ は

全く同じものとみなせるよう対応づけることもできる.

(拙著 1903.04951)

§5. この論文のモデル構築法

◎ AdS/CFT としての CFT 側は mass m を持つ 1+1 次元の Majorana fermion を lattice に乗せた自由理論を考える.

$$H = -\sum_{i=1}^{L} i\left(1+m(-1)^i\right)\chi_i\chi_{i+1} \qquad \begin{cases}\chi_i : \text{Majorana fermion}\\ m : \text{fermion mass.}\end{cases}$$

$$\{\chi_i, \chi_j\} = \delta_{ij}$$

この理論は $m=0$ が相転移点であり gapless になることが知られている. (2次相転移, 量子臨界点)

($\Rightarrow$ AdS/CFT を考える上でよい模型の候補.)

理由　1次元 spin chain 系 ("横磁場イジング模型")

$$H = -J\sum_k \sigma_y^{(k)}\sigma_y^{(k+1)} - K\sum_k \sigma_z^{(k)}$$

ここで, fermion への map を考える (Jordan-Wigner transf.):

$$\begin{cases}\chi_{2k-1} \equiv \sigma_z^{(1)}\cdots\sigma_z^{(k-1)}\sigma_y^{(k)}\\ \chi_{2k} \equiv \sigma_z^{(1)}\cdots\sigma_z^{(k-1)}\sigma_x^{(k)}\end{cases}$$

するとこれらは Majorana fermion の $\{\chi_i, \chi_j\} = 2\delta_{ij}$ を満たす.

Hamiltonian は ("Kitaev spin chain")

$$H = -K\sum_k i\chi_{2k-1}\chi_{2k} - J\sum_k i\chi_{2k}\chi_{2k+1}$$

この $K = 1+m$, $J = 1-m$ と置いたものが上のもの。

一方この横磁場 Ising 系は $K=J$ が相転移点であった.

$$\begin{cases} J<K : \text{1つの基底状態}\\ J>K : \sigma_y = \pm 1 \text{ の 2つの強磁性基底状態}\end{cases}$$

//

注) スピンレス超伝導体としての解釈もある.

Majorana と考えるより通常の fermion に直した方が分かり易い。

$\chi_{2k-1} \equiv c_k + c_k^\dagger$, $\chi_{2k} \equiv i(c_k - c_k^\dagger)$ とおくと

c_k は通常の fermion で, $\{c_k, c_{k'}^\dagger\} = \delta_{kk'}$ を満たす。

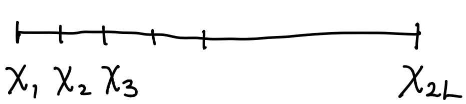
この時.

$$
\begin{aligned}
H = \; & K \sum_k (c_k + c_k^\dagger)(c_k - c_k^\dagger) \\
& + J \sum_k (c_k - c_k^\dagger)(c_{k+1} + c_{k+1}^\dagger) \\
= \; & J \sum_k \left(c_k c_{k+1}^\dagger + c_{k+1} c_k^\dagger \right) && \leftarrow \text{Hopping term.} \\
& + J \sum_k \left(c_k c_{k+1} - c_k^\dagger c_{k+1}^\dagger \right) && \longleftarrow \text{超伝導 pairing term} \\
& + K \sum_k \left(2 c_k^\dagger c_k - 1 \right) && \longleftarrow \text{化学 potential.}
\end{aligned}
$$

つまり. スピンレスの(p波)超伝導体となる.

補足) この Majorana chain は Kitaev 模型と呼ばれ, トポロジカル物質の典型例である。

∵ 一次元 chain に端があるとすると



$K=0$ の時, χ_1 と χ_{2L} は他と decouple ⇒ 端状態あり

実際. $K < J$ なら端状態があり基底状態がそのため縮退することが示される。 $K > J$ なら χ_1 や χ_{2L} は他とよく couple し端状態は存在しない.

⑨ 2nd-Renyi Entropy の計算法

(2nd Renyi をデータとする際に, spin 系だとよい方法がある。
それを用いて. bulk を含めた Boltzmann machine に map してしまう。)

• 定理っぽい事実:

spin 系において 2-nd Renyi entropy $S_E^{(2)} \equiv \mathrm{Tr}_A\left(\left(\mathrm{Tr}_{\bar{A}}\,\rho\right)^2\right)$ が,

τ という「ピン止め spin」を追加で導入した

$$\tilde{H} = -\sum_{\langle ij\rangle} \sigma_i \sigma_j J_{ij} \underbrace{- h \sum_i \tau_i \sigma_i}_{\text{追加項}}$$

$$\tau_i = \begin{cases} +1 & (i \in A) \\ -1 & (i \in \tilde{A}) \end{cases}$$

h は $O(1)$ の定数. (具体的には $h = \frac{1}{2}\log D$ D: 各サイト i での σ_i の自由度の数)

によって

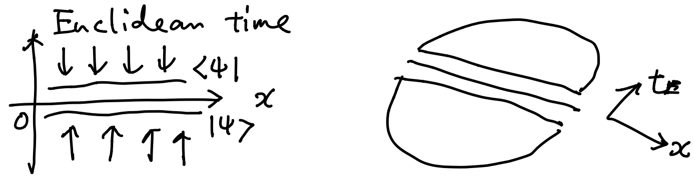
$$e^{-S_E^{(2)}} = F_{(A,\bar{A})} - \underbrace{F_{(\bar{A}=\phi)}}_{\text{全領域をAと考えたもの}}$$

と与えられる.

F: 自由エネルギー

(これは
Random tensor network で表される state について証明される.
1601.01694 Hayden, Nezami, X.L.Qi, Thomas, Walter, Yang
"Holographic duality from Random Tensor Networks")

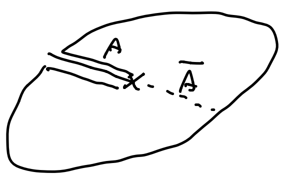
定理の直観的理由の説明. $\rho = |\psi\rangle\langle\psi|$　$|\psi\rangle$: ground state

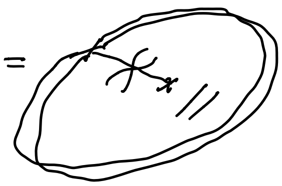


$\left(|\psi\rangle = \lim_{t_E\to\infty} e^{-t_E H} |any\rangle \right.$ で ground state ($E=0$) が選ばれる $\left.\right)$

$\mathrm{Tr}\,\rho = \mathrm{Tr}\,|\psi\rangle\langle\psi| = \langle\psi|\psi\rangle = 1.$

A と $\bar{A}$ に分けて. $\mathrm{Tr}_{\bar{A}}\,\rho$ を考えると

$\rho_A \equiv \mathrm{Tr}_{\bar{A}}\,\rho =$ 

$\therefore \mathrm{Tr}_A\,\rho_A^2 =$ 

Aの領域は上下入れかえてつなげる.

…これは実質的に, Hilbert 空間を double copy して

Aの領域に $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, $\bar{A}$ の領域に $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ を

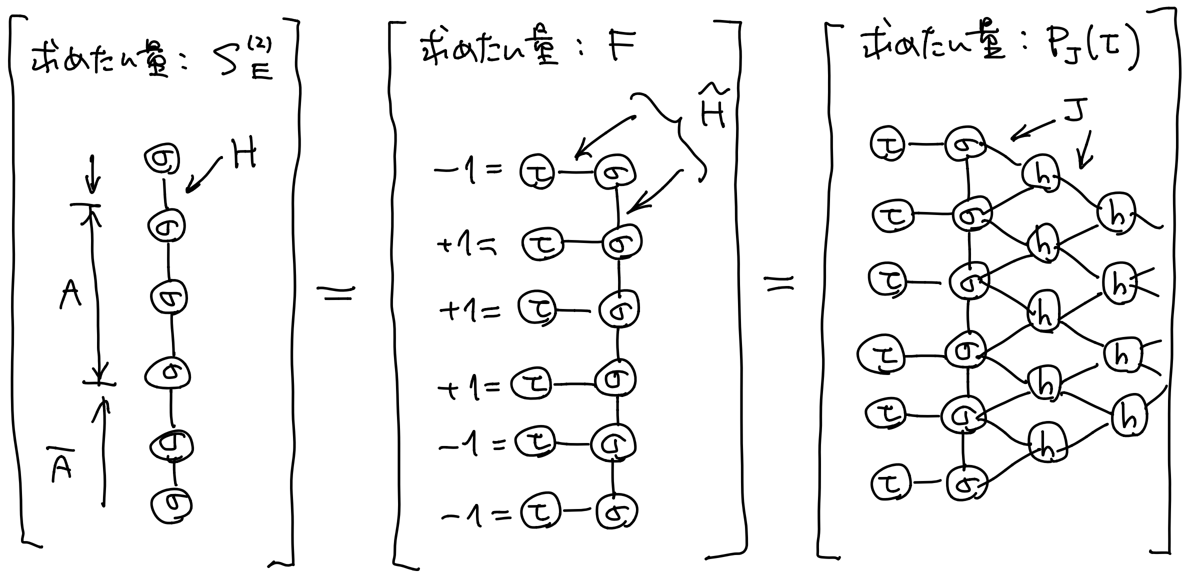
tensor して trace をとることと同じように考えられる

(実際は random tensor states についてより diagramatic に perturbative に証明する)

$\Rightarrow$ A に σ_1, $\bar{A}$ に $\mathbb{1}$.

これは、経路積分の意味では、A と $\bar{A}$ で $+1$ と -1 の spin を挿入することと同じである. //

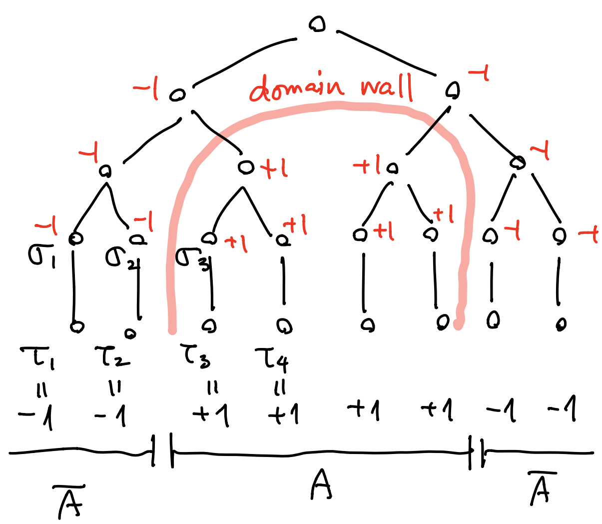
◎ この定理を用いて, free fermion 系の $S_E^{(2)}$ をボルツマンマシンにmapできる.



◎ 定理の意味するところ: Ryu-Takayanagi 公式の直観的理解

Random tensor network を, spin系とみなしたとして

それが強磁性相にあったとする.



τ スピンが tensor net 内部に伝わり, domain を形成する

⇩

Domain wall が出現, Free energyに効く

⇩

2nd Renyi に RT surface が出る.