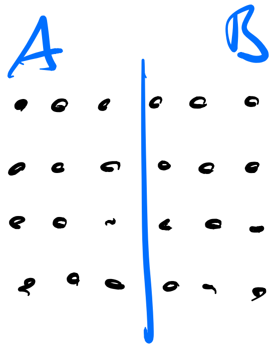
2. Entanglement

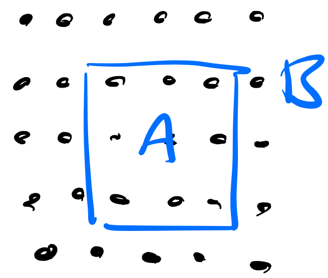
a) Introduction

Consider a system consisting of two parts:

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B .$$

(In a many-body system, this could come from a bipartition:





)

- States $|\psi\rangle$ which can be written as

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$$

(i.e.: $|\psi_A\rangle = \sum a_i |i\rangle$, $|\psi_B\rangle = \sum b_j |j\rangle$:

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle = \sum a_i b_j |i,j\rangle \text{)}.$$

for some $|\psi_A\rangle$, $|\psi_B\rangle$ are called **product states** or **separable states**.

- States $|\psi\rangle$ which are <u>not</u> of this form, i.e. which <u>cannot</u> be written as $|\psi_A\rangle \otimes |\psi_B\rangle$, are called <u>entangled</u>.

That is,

$$(*) \quad |\psi\rangle = \gamma_1 |\psi_{A,1}\rangle \otimes |\psi_{B,1}\rangle + \gamma_2 |\psi_{A,2}\rangle \otimes |\psi_{B,2}\rangle + \ldots$$

with <u>more</u> than one term!

This suggests that entanglement is related to some kind of correlations betw. A & B:

$|\psi_{A,1}\rangle \longleftrightarrow |\psi_{B,1}\rangle$ (weight $|\gamma_1|^2$)

$|\psi_{A,2}\rangle \longleftrightarrow |\psi_{B,2}\rangle$ (weight $|\gamma_2|^2$)

$\vdots$ etc.

How to <u>quantify</u> <u>amount of entanglement</u>?

Intuitively, it should depend on weights $|\gamma_k|^2$

and distinguishability $1-|\langle\psi_{A,k}|\psi_{A,\ell}\rangle|^2$ & $1-|\langle\psi_{B,k}|\psi_{B,\ell}\rangle|^2$.

But naively, this is not even invariant under writing $|\psi\rangle$ in different ways as $(*)$.

Q: How can we measure entanglement in a meaningful way?

b) The singular value decomposition

Theorem (Singular Value Decomposition, SVD):

Any complex $m\times n$-matrix M can be written as

$$M = UDV^\dagger ,$$

with U, V isometries (i.e. $U^\dagger U = V^\dagger V = I$), and

$$D = \begin{pmatrix} s_1 & & & 0 \\ & s_2 & & \\ & & \ddots & \\ 0 & & & s_r \end{pmatrix} \quad ; \quad r \leq m, n.$$

with $s_1 \geq s_2 \geq \dots \geq s_r > 0$ the singular values of Π.

The s_k are the non-zero eigenvalues of $\Pi\Pi^+$ or equivalently of $\Pi^+\Pi$.

(Note: U, V are unique up to rotations in subspaces of degen. s_i. Often, the SVD is stated with U, V unitary and D a $n\times m$-matrix. It is obtained from the form above by padding D with zeros and completing U and V to unitaries by adding columns.)

Proof: Diagonalize $\Pi\Pi^+$:

$$\Pi\Pi^+ = W\Lambda W^+; \quad W \text{ unitary},$$

$$\Lambda = \underbrace{\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots \lambda_r & \\ & & & 0 \ddots 0 \end{pmatrix}}_{n} \Big\}\, n \quad \text{with } \lambda_1 \geq \lambda_2 \geq \dots \lambda_r > 0$$

define $\Pi := \underbrace{\left(\begin{array}{c|c} 1 \ddots 1 & 0 \end{array}\right)}_{n} \Big\}\, r$,

$$\mathcal{U} := W\Pi^\dagger, \quad D := \begin{pmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_r}\end{pmatrix}, \quad \text{and}$$

$$V^\dagger := D^{-1}\Pi W^\dagger M.$$

Then, $\mathcal{U}^\dagger \mathcal{U} = \Pi W^\dagger W \Pi^\dagger = \Pi \cdot I \cdot \Pi^\dagger = I$,

$$V^\dagger V = D^{-1}\underbrace{\Pi \underbrace{W^\dagger M M^\dagger W}_{=\Lambda}\Pi^\dagger}_{=D^2} D^{-1} = I,$$

(i.e. $\mathcal{U}, V$ isometries), and

$$\underbrace{(I-\Pi^\dagger\Pi)}_{=\left(\begin{array}{c|c} \begin{smallmatrix}0&&\\&\ddots&\\&&0\end{smallmatrix} & \\ \hline & \begin{smallmatrix}1&&\\&\ddots&\\&&1\end{smallmatrix}\end{array}\right)} W^\dagger M M^\dagger W (I-\Pi^\dagger\Pi) = (I-\Pi^\dagger\Pi)\Lambda(I-\Pi^\dagger\Pi) = 0,$$

$$\Longrightarrow (I-\Pi^\dagger\Pi)W^\dagger M = 0. \quad \text{Thus,}$$

$$\underline{\underline{\mathcal{U}DV^\dagger}} = (W\Pi^\dagger)D(D^{-1}\Pi W^\dagger M)$$

$$= W\Pi^\dagger\Pi W^\dagger M = WIW^\dagger M = \underline{M}.$$

c) The Schmidt decomposition

Back to bipartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$.

Consider ONB's $|i\rangle_A$, $|j\rangle_B$.

Write
$$|\psi\rangle = \sum c_{ij} |i\rangle_A |j\rangle_B .$$

Use SVD $C = (c_{ij}) = U \cdot D \cdot V^\dagger$,

i.e. $c_{ij} = \sum u_{ik} s_k \overline{v_{jk}}$

$$\Rightarrow |\psi\rangle = \sum_k s_k \underbrace{\Big(\sum_i u_{ik} |i\rangle\Big)_A}_{=: |\psi_A^k\rangle} \underbrace{\Big(\sum_j \overline{v_{jk}} |j\rangle\Big)}_{= |\psi_B^k\rangle}$$

$=: |\psi_A^k\rangle$ ONS as u_{ik} isometry!

$= |\psi_B^k\rangle$ ONS as $\overline{v_{jk}}$ isometry!

$$\Rightarrow \boxed{|\psi\rangle = \sum_{k=1}^{r} s_k |\psi_A^k\rangle \otimes |\psi_B^k\rangle}$$

with $s_k > 0$.

The Schmidt decomposition, with Schmidt coefficients $s_k > 0$, and Schmidt rank r.

Note: • The $\{|\psi_A^k\rangle\}$ and the $\{|\psi_B^k\rangle\}$ each form an orthonormal set!

• The Schmidt decomposition is unique, up to rotations within subspaces with degenerate Schmidt coefficients.

d) Reduced density matrices

Density matrices: typ. introduced to describe states where we have partial knowledge:

Consider $\langle\psi|\Pi|\psi\rangle$, with Π e.g. an observable, or a projection onto a meas. result:

$\langle\psi|\Pi|\psi\rangle = \mathrm{tr}[\Pi \cdot \underbrace{|\psi\rangle\langle\psi|}]$

(Projector onto $|\psi\rangle$

$\mathrm{tr}(X) = \sum X_{ii}$. Basis-independent!

Then, if we have state $|\psi_i\rangle$ w/ probability p_i:

Avg. outcome is

$$\sum p_i \langle\psi_i|\Pi|\psi_i\rangle = \sum p_i \, \mathrm{tr}[\Pi |\psi_i\rangle\langle\psi_i|]$$

$$= \mathrm{tr}[\Pi \cdot \sum p_i |\psi_i\rangle\langle\psi_i|]$$

$$= \mathrm{tr}[\Pi \rho]$$

with $\rho := \sum p_i |\psi_i\rangle\langle\psi_i|$ the density matrix (or density operator).

(Can be used to describe ensemble $\{p_i, |\psi_i\rangle\}$.

Note: This is not uniquely determined by ρ!)

Back to bipartite states. Consider $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$.

How can we describe the expectation value of an operator Ω_A on A? (E.g. measurement)

Operator "ignores" B system. Thus, on any product state $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, we must have

$$\text{"}\langle\psi|\Pi_A|\psi\rangle\text{"} := \langle\psi_A|\Pi_A|\psi_A\rangle$$

$$= \langle\psi_A|\Pi_A|\psi_A\rangle\langle\psi_B|\psi_B\rangle$$

That is, Π_A acts on $|\psi_A\rangle \otimes |\psi_B\rangle$ as

$$|\psi_A\rangle \otimes |\psi_B\rangle \longmapsto (\Pi_A|\psi_A\rangle) \otimes |\psi_B\rangle.$$

This is exactly the definition of the operator $\underline{\underline{\Pi_A \otimes \mathbb{1}_B}}$! – Due to linearity, Π_A must act as $\Pi_A \otimes \mathbb{1}_B$ on <u>all</u> states $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$!

Now let $|\psi\rangle = \sum c_{ij} \, |i\rangle_A \otimes |j\rangle_B$.

Then, $\langle\psi| \Pi_A \otimes \mathbb{1}_B |\psi\rangle =$

$$= \sum c_{ij} \, \overline{c_{i'j'}} \left(\langle i'|_A \otimes \langle j'|_B\right)\left(\Pi_A \otimes \mathbb{1}_B\right)\left(|i\rangle_A \otimes |j\rangle_B\right)$$

$$= \sum c_{ij} \overline{c_{i'j'}} \underbrace{\langle i' | \Pi_A | i \rangle}_{= tr[\Pi_A |i\rangle\langle i'|]} \underbrace{\langle j' | \mathbb{1}_B | j \rangle}_{= \delta_{jj'}}$$

$$= \sum_{ii'j} c_{ij} \overline{c_{i'j}} \, tr[\Pi_A |i\rangle\langle i'|]$$

$$= tr[\Pi_A \rho],$$

with $\rho = \sum_{ii'j} c_{ij} \overline{c_{i'j}} |i\rangle\langle i'|,$

or $\rho_{ii'} = \langle i | \rho | i' \rangle = (CC^\dagger)_{ii'}, \; C = (c_{ij}).$

This can be formalized through the concept of the partial trace: Given ρ_{AB}, the partial trace is

$$\rho_A = tr_B \, \rho_{AB} := \sum_j (\mathbb{1}_A \otimes \langle j |_B) \rho_{AB} (\mathbb{1}_A \otimes | j \rangle_B)$$

$$\equiv \sum_{j} \langle j_B | \rho_{AB} | j \rangle_B$$

$$\equiv \sum_{i,i',j} |i\rangle_A \langle i,j| \rho_{AB} |i',j\rangle \langle i'|_A$$

Again, ρ_A describes anything pertaining to system A.

In particular, for the case $\rho_{AB} = |\psi\rangle\langle\psi|$,

$|\psi\rangle = \sum c_{ij} |i\rangle \otimes |j\rangle$:

$$\rho_A = \sum c_{ij} \overline{c_{i'j'}} \, tr_B \left[\left(|i\rangle_A \langle i'|_A \right) \otimes |j\rangle_B \langle j'|_B \right]$$

$$= \sum c_{ij} \overline{c_{i'j'}} \, |i\rangle_A \langle i'|_A \, \underbrace{tr\left[|j\rangle\langle j'| \right]}_{= \delta_{jj'}}.$$

Finally, consider Schmidt decomposition of $|\psi\rangle$:

$$|\psi\rangle = \sum_{k=1}^{r} s_k \, |\psi_A^k\rangle \otimes |\psi_B^k\rangle.$$

Then,

$$\underline{\rho_A} = \mathrm{tr}_B \left[\sum_{k,l} s_k s_l \, |\psi_A^k\rangle\langle\psi_A^l| \otimes |\psi_B^k\rangle\langle\psi_B^l| \right]$$

$$= \sum_{kl} s_k s_l \, |\psi_A^k\rangle\langle\psi_A^l| \, \underbrace{\mathrm{tr}_B\left[|\psi_B^k\rangle\langle\psi_B^l| \right]}_{}$$

$= \delta_{kl}$ as $|\psi_B^\bullet\rangle$ ONS (cyclicity or trace in $|\psi_B^\bullet\rangle$)

$$\underline{= \sum_k s_k^2 |\psi_A^k\rangle\langle\psi_A^k| .}$$

Similarly, $\rho_B = \mathrm{tr}_A |\psi\rangle\langle\psi| = \sum_k s_k^2 |\psi_B^k\rangle\langle\psi_B^k|$.

$\Rightarrow$ <u>Schmidt coefficients</u> are the <u>non-zero eigenvalues of ρ_A (or ρ_B)</u>.

(In particular: For a pure state $|\psi\rangle = |\psi\rangle_{AB}$, ρ_A and ρ_B have the same non-zero eigvals.)

The <u>Schmidt vectors</u> are the eigenvectors of ρ_A & ρ_B, respectively.

Unless there are degenerate s_k, this uniquely determines the Schmidt decomposition.

e) Quantifying entanglement

Recall: $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$:

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle \Leftrightarrow: |\psi\rangle \text{ product (or separable)}$$

$$|\psi\rangle \neq |\psi_A\rangle \otimes |\psi_B\rangle \Leftrightarrow |\psi\rangle \text{ entangled}$$

– i.e., $|\psi\rangle$ has non-trivial quantum correlations which cannot be created by local operations & classical communication.

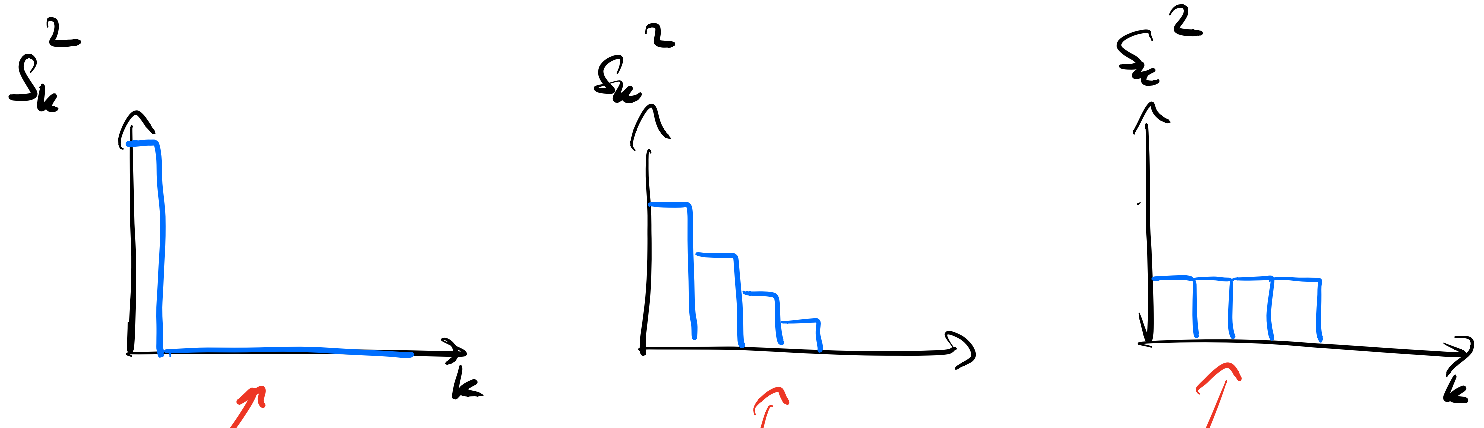
What determines if, and how much, a state is entangled?

Use Schmidt basis:

$$|\psi\rangle = \sum s_k |\psi_A^k\rangle \otimes |\psi_B^k\rangle$$

$|\psi_A^k\rangle$ ONS, $|\psi_B^k\rangle$ ONS: For each k, we have perfect (i.e. orthogonal/distinguishable) correlations between A & B. The amount of correlations should dep. on the distribution of the s_k — if more events can occur with same probability, there are more correlation.

Indeed: The $|\psi_A^k\rangle$ & $|\psi_B^k\rangle$ can be changed with local rotations, and this is all that local rotations can do $\Rightarrow$ all info. about entanglement is in the s_k.



no corrs./ entanglement

some (imperfect) corrs./ entanglement

perfect corrs./ maximal entanglement

(Intuitively:)

Amount of entanglement $\longleftrightarrow$ Amount of disorder in $p_k = s_k^2$.
($\sum p_k = 1$!)

Use entropy as measure of disorder:

$$H(\{p_k\}) := -\sum p_k \log p_k$$

("Shannon entropy")

base e (cond.-mat.)
or base 2 (q. info)

or equivalently the von Neumann entropy

$$\boxed{S(\rho_A) := -\mathrm{tr}\left[\rho_A \log \rho_A\right]}$$

Defined on the eigenvalues, i.e.

$$\rho_A = \sum p_i |\psi_i\rangle\langle\psi_i| \Rightarrow \log \rho_A = \sum \log p_i |\psi_i\rangle\langle\psi_i|$$

$$\Rightarrow S(\rho_A) = -\mathrm{tr}\left[\sum p_i \log p_i \, |\psi_i\rangle\langle\psi_i|\right]$$

$$= -\sum p_i \log p_i$$

In particular, for $|\psi\rangle = \sum s_k |\psi_A^{(k)}\rangle \otimes |\psi_B^{(k)}\rangle$:

$$S(\mathrm{tr}_A |\psi\rangle\langle\psi|) = S(\mathrm{tr}_B |\psi\rangle\langle\psi|) = H(\{s_k^2\}).$$

Define the <u>entropy of entanglement</u>, or <u>entanglement entropy</u>, as

$$\boxed{E(\psi) = S(\mathrm{tr}_A |\psi\rangle\langle\psi|) = S(\mathrm{tr}_B |\psi\rangle\langle\psi|)}$$

Alternatively, we can also use other entropy measures, most importantly the α-<u>Rényi entropies</u>

$$\boxed{S_\alpha(\rho) = \frac{1}{1-\alpha}\log\left(\mathrm{tr}(\rho^\alpha)\right)}$$

or $$H_\alpha(\{p_k\}) = \frac{1}{1-\alpha}\log\left(\sum p_i^\alpha\right).$$

<u>Special cases:</u>

$$S_0(\rho) = \log(\mathrm{rank}(\rho))$$

$$S_\infty(\rho) = -\log(\lambda_{max}(\rho))$$

$$\lim_{\alpha\to 1} S_\alpha(\rho) = S(\rho).$$

Important – though not for this lecture:

$E(\psi) = S(\mathrm{tr}_A|\psi\rangle\langle\psi|)$ has a clear operational interpretation when converting betw. entangled states w/ local operations &

class. communication (LOCC) in the case of many copies:

$$|\psi\rangle^{\otimes N} \underset{\longleftarrow}{\overset{\longrightarrow}{\text{LOCC}}} |\phi\rangle^{\otimes M}$$

can be converted back and forth asympt. reversibly if & only if

$$\underbrace{E(\psi)\cdot N}_{\text{"total ent. in } |\psi\rangle^{\otimes N}\text{"}} = \underbrace{E(\phi)\cdot M}_{\text{"total ent. in } |\phi\rangle^{\otimes M}\text{"}} .$$

f) Entanglement in ground states

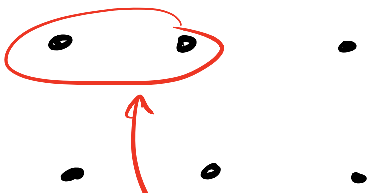
What can we say about entanglement in ground states of local (gapped) Hamiltonians?

What is the role played by entanglement in g.s.?

Example: Heisenberg antiferromagnet (HAFM):

$$H = \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j$$

sum over nearest neighbors

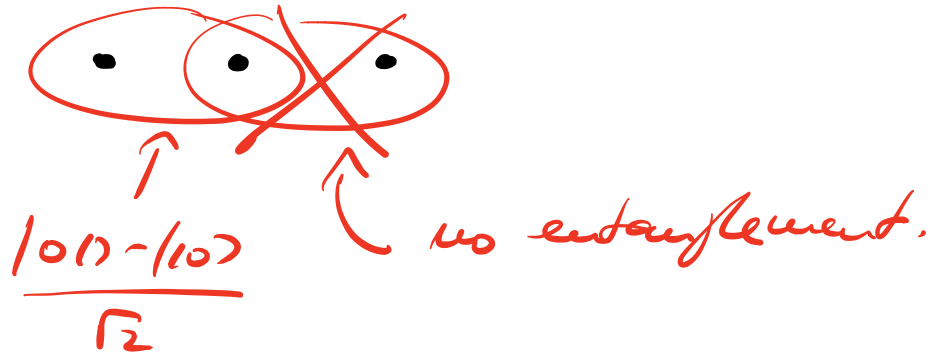


$h = \vec{S}_i \cdot \vec{S}_j \Rightarrow$ ground state $\frac{|01\rangle - |10\rangle}{\sqrt{2}}$

$\Rightarrow$ entanglement betw. adjacent sites

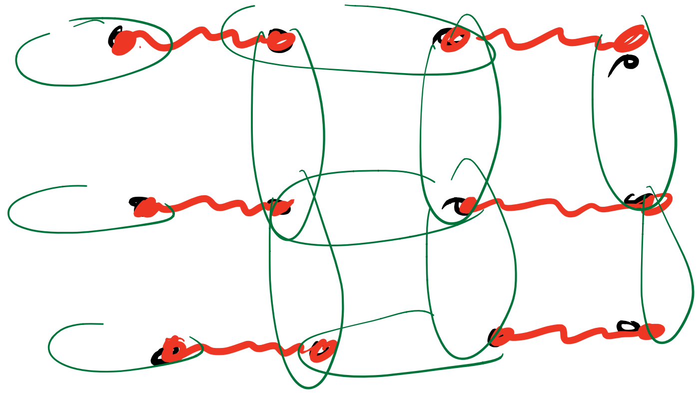
$h = \vec{S}_i \cdot \vec{S}_j$ wants max. entanglement betw. adjacent sites (NN: nearest neighbors)

Problem: We cannot max. ent. a spin with more than one other spin ("monogamy of entanglement")



True ground state has to "balance out" between entangling diff. NN $\Rightarrow$ as a side-effect, this also induces long-range quantum correlations, but only as a "higher order correction".

Can also be (intuitively) understood using pert. theory: Start from one pairing of singlets,



and introduce the other couplings perturbatively.

Each order in pert. theory (which is suppressed exponentially) will increase range of correlations by one site.

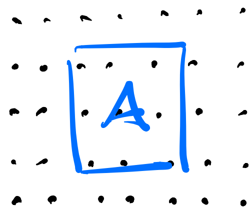
— N.B.: This is all just a very hand-wavy explanation!! —

Consequence: Entanglement in ground states builds up locally!

Formalisation:

g) The entanglement area law

Definition:
We say that a state (or better a family of states defined on increasing lattice sizes N) satisfies an area law (or entanglement area law) if for every "nice" region A, the



147

entanglement $E(A) = S(\mathrm{tr}_B|\psi\rangle\langle\psi|) = S(\rho_A)$

scales as

$$\boxed{S(\rho_A) \leq \alpha\,|\partial A|}$$

where $|\partial A|$ is the length of the boundary of A.

(More precisely, we should state this for a uniform family $|\psi_N\rangle$ of states, e.g. ground states of a transl. invariant $H = \sum^N h$, and for a uniform choice of regions A, e.g. rectangles with a fixed / bounded aspect ratio, circles, ..., and α has to be uniform.)

<u>Key point:</u> <u>Ground states of local Hamiltonians satisfy an area law!</u>

More precisely: Known area results — both "common knowledge" and proven:

spatial dim. / gap of H	1 D	2D and higher
gapped H	$S(\rho_A) \le \alpha \lvert\partial A\rvert$ (in 1D: $\lvert\partial A\rvert = \text{const.}$ $\Longrightarrow S(\rho_A) \le \text{const.}$) proven by Hastings https://arxiv.org/abs/0705.2024 and improved by Arad, Kitaev, Landau, Vazirani https://arxiv.org/abs/1301.1162	$S(\rho_A) \le \alpha \lvert\partial A\rvert$ area law "known" to be true no proof (yet), except for special cases
gapless H	$S(\rho_A) \le \gamma \log \lvert A\rvert$ (still exp. better than the worst case $S(\rho_A) \sim \lvert A\rvert$) "known" — hold for all physically relevant cases. But: (contrived) counterex. known	$S(\rho_A) \le \alpha \lvert\partial A\rvert$ area law still holds for spin systems $S(\rho_A) \le \alpha \lvert\partial A\rvert \log \lvert A\rvert$ for free fermions w/ Fermi surface (= metals) "known" — no proof

<u>Important:</u> Even critical systems generally display only a logarithmic entanglement scaling.

This is in <u>stark contrast</u> to a random (Haar-random) state, for which

$$S(\rho_A) = |A| - c \log |A| \;!$$

$\Rightarrow$ ground states are <u>very special</u> in the space of all states!

(We knew this from parameter counting, but now we know what makes them special: They have (comparatively) very little entanglement!)

So... what is the structure of many-body states with little entanglement?