

Data Compression Limit for a class of non-i.i.d quantum information sources

Nilanjana Datta
University of Cambridge

in collaboration with : Yuri Suhov

Ref: [Quantum Information Processing 1(4)
257-281, Aug '02]

Issue : Compression of information from a particular class of quantum info sources

Question : What are the minimal resources needed to store the output of such a source ?

Quantum Information source :

- a set of q.m. states $\{|\psi_j\rangle\}$
(i.e., vectors from a given Hilbert space \mathcal{H}_n)
- a set of corresponding probabilities $\{p_j\}$

$|\psi_j\rangle$: signal/output of the source
(state of a q.m. system)

[message \equiv signal \equiv data \equiv output of the source]

p_j : probabilities with which the signals are produced

Density matrix formalism

$$\rho = \sum_j p_j |\psi_j\rangle \langle \psi_j|$$

Quantum Information Source:

$$\rho = \sum_{i=1}^n p_i |\psi_i\rangle\langle\psi_i| \quad ; \quad |\psi_i\rangle \in \mathcal{H} \quad ; \quad \dim \mathcal{H} = N$$

$$\langle\psi_i|\psi_j\rangle = \delta_{ij} \quad \text{orthormal vectors}$$

To compress data from such a source one

encodes: a signal state $|\psi_j\rangle \in \mathcal{H}$
 $\dim \mathcal{H} = N$

using a vector $|x_i\rangle \in \tilde{\mathcal{H}}$
 $\dim \tilde{\mathcal{H}} = d < N$

(lower-dimensional H.S.)

⑧ What are the minimal resources needed to compress data from such a source?

resource: dimension (d) of the H.S. of the $(\tilde{\mathcal{H}})$ encodings

⑧: What is the maximum possible compression of H.S. dimensions per signal state such that the signal states can be reconstructed with an arbitrarily low probability of error from the compressed version?

Find: the limiting rate of data compression

Consider: $\left\{ \begin{array}{l} \text{signals } |\psi_j\rangle \in \mathcal{H}_n; \dim \mathcal{H}_n = 2^n \\ \mathcal{H}_n: \text{Hilbert space of } n \text{ qubits} \\ \text{source } \rho_n = \sum_j p_j^{(n)} |\psi_j\rangle \langle \psi_j| \text{ acts on } \mathcal{H}_n \end{array} \right.$

Compression scheme of rate R : ($R < 1$)

$$C: |\psi_j\rangle \in \mathcal{H}_n \longmapsto |x_i\rangle \in \tilde{\mathcal{H}}$$
$$\dim \mathcal{H}_n = 2^n \qquad \dim \tilde{\mathcal{H}} = d = 2^{nR}$$

Rate of compression

$$R := \frac{\log_2 (\dim \tilde{\mathcal{H}})}{\log_2 (\dim \mathcal{H}_n)} = \frac{\log_2 d}{n}$$

Regard: compressed space $\tilde{\mathcal{H}}$: nR qubits

R : # of qubits needed to encode each qubit of information

$$\text{limiting rate } R_\infty := \lim_{n \rightarrow \infty} \frac{\log_2 d}{n}$$

Key to data compression:

some signal states have a higher probability of occurrence than others

R_∞ : found by incorporating the redundancy resulting from an uneven prob. distribution of signals

Use the notion of a "typical subspace"

Typical subspace [Schumacher '95]

$\mathcal{M} \subset \mathcal{H}_n$ [where \mathcal{H}_n : original Hilbert space
s.t. $\forall \epsilon > 0$ and n large enough:
 $\dim \mathcal{H}_n = 2^n$]

$$\text{Prob}(|\psi_j\rangle \in \mathcal{M}) > 1 - \epsilon$$

signal

$$\therefore \text{Prob}(|\psi_j\rangle \notin \mathcal{M}) \leq \epsilon$$

\mathcal{M} : a subspace of states which have a high probability of occurrence

If $|\psi_j\rangle \in \mathcal{M}$: typical state

$|\psi_j\rangle \notin \mathcal{M}$: atypical state

If a typical subspace $\mathcal{M} \subset \mathcal{H}_n$ exists:

where $\dim \mathcal{H}_n = 2^n$; $\dim \mathcal{M} = d < 2^n$

• then it suffices to encode states

$|\psi_j\rangle \in \mathcal{M}$ only

• ignore the atypical states \because they rarely occur

• **encode** each typical state $|\psi_j\rangle \xrightarrow{1-1} |x_i\rangle \in \tilde{\mathcal{H}}$

where $\dim \tilde{\mathcal{H}} = \dim \mathcal{M} = d$

compressed states $|x_i\rangle$ can be decompressed unambiguously $\rightarrow |\psi_j\rangle$

$R_\infty \leftrightarrow h$ **von Neumann entropy rate** of source (ρ_n)
von Neumann entropy

$$S(\rho_n) = -\text{tr} \rho_n \log_2 \rho_n \\ = -\text{tr} \sum_j p_j^{(n)} \log_2 p_j^{(n)}$$

$$h := \lim_{n \rightarrow \infty} \frac{1}{n} S(\rho_n)$$

(if limit \exists)

An i.i.d quantum info source : Density matrix ρ

• acts on a tensor pdkt H.S. $\mathcal{H}_n = \mathcal{H}^{\otimes n}$

• is given by $\rho \equiv \rho_n = \pi^{\otimes n}$

Here :

\mathcal{H} : a fixed H.S. representing an elementary quantum subsystem

$\pi = \sum q_i |\phi_i\rangle\langle\phi_i|$: density matrix acting on \mathcal{H}

e.g.:

\mathcal{H} : single qubit space ; $\dim \mathcal{H} = 2$

$|\phi_i\rangle \in \{|0\rangle, |1\rangle\}$

Spectral decomposition of ρ_n

$$\rho_n = \sum_j p_j^{(n)} |\psi_j\rangle\langle\psi_j|$$

eigenstates:

$$|\psi_j\rangle = |\phi_{j_1}\rangle \otimes |\phi_{j_2}\rangle \otimes \dots \otimes |\phi_{j_n}\rangle$$

eigenvalues:

$$p_j^{(n)} = q_{j_1} \cdot q_{j_2} \cdot \dots \cdot q_{j_n}$$

von Neumann entropy of source $S(\rho_n) = nS(\pi)$

$$S(\pi) = - \sum_{j=1}^2 q_j \log_2 q_j$$

von Neumann entropy of a qubit

Schumacher [PRA 51, 2738, '95]

The minimum # of qubits necessary to represent reliably each qubit of info produced by an i.i.d quantum info source $\rho_n = \pi^{\otimes n}$ is asymptotically given by $S(\pi)$

Note: $S(\rho_n) = n S(\pi) \quad \therefore S(\pi) = h$ vN entropy rate

\exists a reliable compression scheme of rate R only when $R > S(\pi)$

Extensions of Schumacher's theorem:

- Josza et al : to mixed state ensembles
 - Petz et al
 - Bjelakovic et al
 - Kaltchenko et al
- } to some classes of non-product quantum sources

Our work concerns : data compression for a class of quantum info sources modelled by a system of **strongly interacting spins**

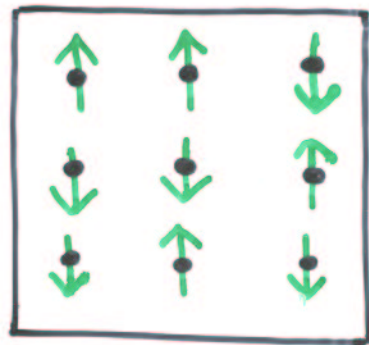
Non-i.i.d quantum information source modelled by a quantum spin system

Basic set-up:

- a quantum-mechanical system defined on a lattice \mathbb{Z}^d
- a spin- $1/2$ particle attached to each site $x \in \mathbb{Z}^d$

$|\uparrow\rangle$: up-spin state

$|\downarrow\rangle$: down-spin state



To each $x \in \mathbb{Z}^d$

$j_x = +1$ if $|\uparrow\rangle$ at x

$= -1$ if $|\downarrow\rangle$ at x

A configuration ω_x on $X \subset \mathbb{Z}^d$:

an assignment $\omega_x = \{j_x \mid x \in X\}$

To each site $x \in \mathbb{Z}^d \leftrightarrow$ H.S. \mathcal{H}_x

$$\mathcal{H}_x \simeq \mathbb{C}^2$$

Hilbert space of a finite lattice $\Lambda \subset \mathbb{Z}^d$:

$$\mathcal{H}_\Lambda = \bigotimes_{x \in \Lambda} \mathcal{H}_x \simeq (\mathbb{C}^2)^{\otimes |\Lambda|}$$

$2^{|\Lambda|}$ -dim space ; $|\Lambda|$: # of sites in Λ

(d) 7/2

a configuration ω on Λ } $\xleftrightarrow[\text{corrs.}]{1-1}$ { a basis vector
in $\mathcal{H}_\Lambda : |\omega\rangle$

Physics : of the quantum spin system on Λ
governed by a Hamiltonian H_Λ

H_Λ : specifies the interactions between the
spin- $1/2$ particles

Quantum spin system \equiv a system of interacting qubits

Consider a class of quantum spin systems:

$$H_\Lambda = H_{0\Lambda} + \epsilon V_\Lambda$$

leading part

quantum perturbation

ϵ : perturbation parameter

$$[H_{0\Lambda}, V_\Lambda] \neq 0$$

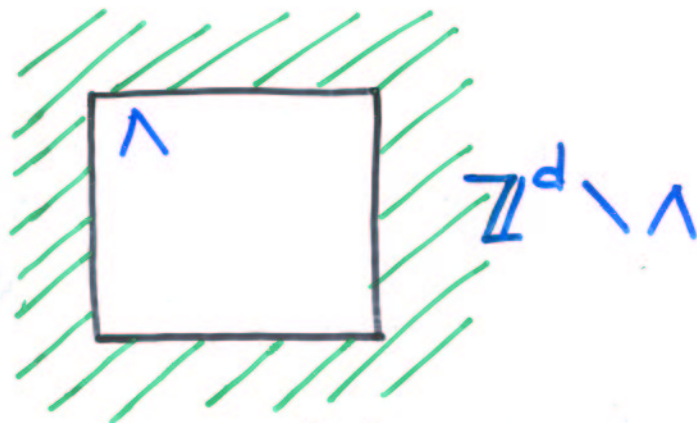
finite range (r) interactions :

e.g.

$$H_\Lambda = - \sum_{i \in \Lambda} \sigma_i^z \sigma_{i+1}^z + \epsilon \sum_{i \in \Lambda} [\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y]$$

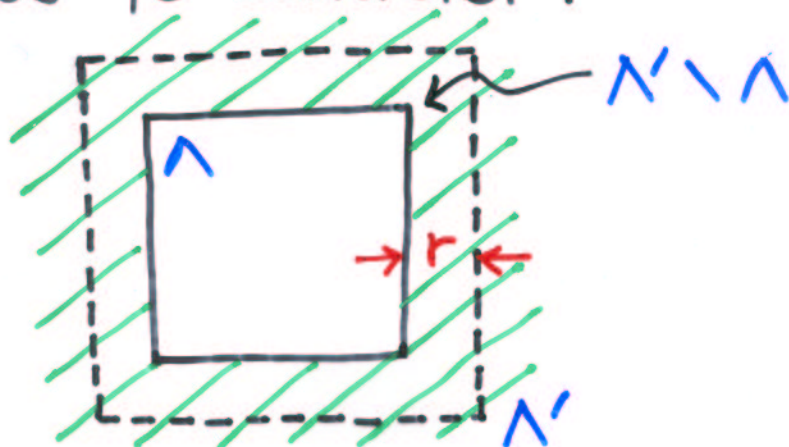
$i, i+1$: pair of n.n. sites $\therefore r = 1$

Λ : system ; $\mathbb{Z}^d \setminus \Lambda$: environment



Fix boundary conditions:

Since interactions are of finite range r it suffices to consider:



- $\Lambda' \supset \Lambda$; Hilbert spaces $\mathfrak{H}_{\Lambda'}$, \mathfrak{H}_{Λ}
- fix config in $\Lambda' \setminus \Lambda$ to be ω [= b.c.]
- Hamiltonian : $H_{\Lambda} \rightarrow H_{\Lambda}^{\omega} := P_{\Lambda'}^{\omega} H_{\Lambda} P_{\Lambda'}^{\omega}$
- $P_{\Lambda'}^{\omega}$ { orthogonal projection onto the subspace of $\mathfrak{H}_{\Lambda'}$ spanned by states corrs. to fixed config ω on $\Lambda' \setminus \Lambda$

Non-i.i.d quantum information source

- Quantum spin system on a finite lattice $\Lambda \subset \mathbb{Z}^d$
- subject to boundary condition ω
- at finite (but low) temperature : $\beta = \frac{1}{k_B T}$
(Boltzmann constant)

Hamiltonian H_Λ^ω :

system of interacting spins (interacting with each other & with environment)
(through b.c.)

Quantum info source : defined by the density matrix ρ_Λ^ω of such a system

$\left. \begin{array}{l} \exists \text{ interactions} \\ \text{between spins} \end{array} \right\} \Rightarrow \rho_\Lambda^\omega \neq \left\{ \begin{array}{l} \text{tensor pdct. of} \\ \text{density matrices} \\ \text{of individual spins} \end{array} \right.$

\therefore Quantum Information source is
non - i.i.d

Note:

- We consider interactions to be strong enough for there to be phase transitions
- Each phase in the phase diagram is labeled by a boundary condition
- By fixing a boundary condition we select a phase

→ prove data compression for it

- To compute the data compression limit R_∞ consider the thermodynamic limit : $\Lambda \uparrow \mathbb{Z}^d$.

Density matrix

$$\rho_{\Lambda}^{\omega} = \frac{e^{-\beta H_{\Lambda}^{\omega}}}{\text{tr } e^{-\beta H_{\Lambda}^{\omega}}} \dots \textcircled{a}$$

Gibbs state
thermal equil. state

Note: **Any** density matrix can be written as **ⓐ**
for some self-adjoint operator H_{Λ}^{ω}
with discrete spectrum s.t.
 $e^{-\beta H_{\Lambda}^{\omega}}$ is trace class

However, we consider H_{Λ}^{ω} to be of form
discussed

Eigenvectors of ρ_{Λ}^{ω} : signals/messages of the
source

Spectral decomposition :

$$\rho_{\Lambda}^{\omega} = \sum_j p_j |\psi_j\rangle \langle \psi_j| \dots \textcircled{b}$$

$$p_j \geq 0 \quad \sum_j p_j = 1 \quad (\because \rho^{\dagger} = \rho, \rho \geq 0, \text{Tr } \rho = 1)$$

p_j : obey same rule as a
(eigenvalues of ρ_{Λ}^{ω}) prob. distr. ρ_{Λ}^{ω}

p_j : prob. of system being in state $|\psi_j\rangle$

Aim: Find data compression limit R_∞ for a quantum info source:

$$\rho_\Lambda^\omega = \frac{e^{-\beta H_\Lambda^\omega}}{\text{tr } e^{-\beta H_\Lambda^\omega}} \quad ; \quad \rho_\Lambda^\omega = \sum_j p_j |\psi_j\rangle \langle \psi_j|$$

Result $R_\infty = h$: von Neumann entropy rate

$$h = \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} S(\rho_\Lambda^\omega) \dots \textcircled{1}$$

$$S(\rho_\Lambda^\omega) = -\text{tr } \rho_\Lambda^\omega \log_2 \rho_\Lambda^\omega \quad ; \quad \text{von Neumann entropy}$$

Thm: (N.D., R. Fernandez, J. Fröhlich)

For β large enough (low temperature) and ϵ small enough (small quantum perturbation) the limit on RHS $\textcircled{1}$ exists.

$\therefore h$ well-defined!

Define a random variable K_Λ^ω : which takes values p_j with prob. p_j

$$\text{Prob}(K_\Lambda^\omega = p_j) = p_j$$

Why?

Ans: • R_∞ related to asymptotical properties of K_Λ^ω

• $E\left(-\frac{1}{|\Lambda|} \log_2 K_\Lambda^\omega\right) = h \dots \textcircled{2}$

• Use properties of K_Λ^ω to prove existence of a **typical subspace**

(= key ingredient for data compression)

Thm: For the class of systems considered for β large enough & ϵ small enough

$\forall \delta > 0$

$$\lim_{\Lambda \rightarrow \mathbb{Z}^d} \text{Prob}\left(\left|-\frac{1}{|\Lambda|} \log_2 K_\Lambda^\omega - h\right| \leq \delta\right) = 1$$

[Weak LLNs due to $\textcircled{2}$]

Alternatively,

$\forall \gamma > 0$ & Λ large enough :

$$\text{Prob}\left(2^{-|\Lambda|(h+\delta)} \leq K_\Lambda^\omega \leq 2^{-|\Lambda|(h-\delta)}\right) \geq 1-\gamma$$

$\forall \gamma > 0$ and Λ large enough:

$$\text{Prob} \left(2^{-\Lambda(h+\delta)} \leq K_{\Lambda}^{\omega} \leq 2^{-\Lambda(h-\delta)} \right) \geq 1-\gamma$$

Since K_{Λ}^{ω} takes values p_j with prob. p_j :

$$2^{-\Lambda(h+\delta)} \leq p_j \leq 2^{-\Lambda(h-\delta)} \dots \textcircled{A}$$

with a prob. of at least $(1-\gamma)$.

Typical states : $\left\{ \begin{array}{l} \text{eigenstates } |\psi_j\rangle \text{ of } \rho_{\Lambda}^{\omega} \\ \text{whose eigenvalue } p_j \text{ satisfies} \end{array} \right. \textcircled{A}$
(δ -typical states)

Typical subspace :

$$\mathcal{M}_{\delta} = \text{span} \left\{ |\psi_j\rangle \mid p_j \text{ satisfies } \textcircled{A} \right\}$$

$$\sum_j p_j = 1 \implies \sum_{j:\textcircled{A}} p_j \leq 1$$

sum of probs. of δ -typical states

Let $|\mathcal{M}_{\delta}| := \dim \mathcal{M}_{\delta}$

$$|\mathcal{M}_{\delta}| \cdot 2^{-\Lambda(h+\delta)} \leq \sum_{j:\textcircled{A}} p_j \leq 1$$

$$\implies |\mathcal{M}_{\delta}| \leq 2^{\Lambda(h+\delta)}$$

Using

$$|\mathcal{M}_\delta| \leq 2^{|\Lambda|(h+\delta)}$$

we prove:

(i) if $R > h$ then there exists a reliable compression scheme of rate R

(ii) a compression scheme of rate $R \leq h$ is not reliable

(i) & (ii) together implies that

$$\text{Data compression limit : } R_\infty = h$$

Proof of (i): $R > h$;

Choose $\delta > 0$ such that $h + \delta \leq R$

$$|\mathcal{M}_\delta| \leq 2^{|\Lambda|(h+\delta)} \leq 2^{|\Lambda|R}$$

\Rightarrow one can uniquely identify a δ -typical state using $\lceil |\Lambda|R \rceil$ qubits

Data compression:

Typical state $|\psi_j\rangle \longrightarrow |\underline{x}\rangle \in \tilde{\mathcal{H}}$

- $\tilde{\mathcal{H}}$: Hilbert space of $[\Lambda|R]$ qubits

- $\underline{x} = (x_1, x_2, \dots, x_{[\Lambda|R]}) \in \{0, 1\}^{[\Lambda|R]}$

$$|\underline{x}\rangle = |x_1\rangle \otimes |x_2\rangle \otimes \dots \otimes |x_{[\Lambda|R]}\rangle$$

$$|x_i\rangle \in \{|0\rangle, |1\rangle\}$$

$|\underline{x}\rangle$: quasiclassical state of $[\Lambda|R]$ qubits

- a $2^{|\Lambda|}$ -dim vector

$$|\psi_j\rangle \xrightarrow{1-1} |\underline{x}\rangle$$

- a $2^{[\Lambda|R]}$ -dim vector

- info contained in $|\Lambda|$ qubits $\} \longrightarrow \{ [\Lambda|R]$ qubits

- $|\underline{x}\rangle$ can be decompressed unambiguously

- As $\Lambda \nearrow \mathbb{Z}^d$ this scheme succeeds with probability 1

(ii) A compression scheme of rate $R < h$ is not reliable

Proof of (ii): This follows from the lemma:

Lemma: Let \mathcal{S}_Λ be any set of eigenstates $\{|\psi_j\rangle\}$ such that

$$|\mathcal{S}_\Lambda| = 2^{\lfloor \Lambda R \rfloor}$$

where $R < h$ is fixed.

Then for any $\gamma > 0$ and sufficiently large Λ

$$\sum_{j \in \mathcal{S}_\Lambda} p_j \leq \gamma \quad \dots \textcircled{B}$$

Proof: LHS of \textcircled{B} = $\left\{ \begin{array}{l} \text{Prob. that an eigenstate} \\ \text{of } \rho_\Lambda^\omega \text{ belongs to } \mathcal{S}_\Lambda \end{array} \right\}$

$$\sum_{j \in \mathcal{S}_\Lambda} p_j = \sum_{\substack{j \in \mathcal{S}_\Lambda \\ \text{(typical)}}} p_j + \sum_{\substack{j \in \mathcal{S}_\Lambda \\ \text{(atypical)}}} p_j \quad \dots \textcircled{C}$$

$$\text{Prob} \left(2^{-\Lambda(h+\delta)} \leq K_\Lambda^\omega \leq 2^{-\Lambda(h-\delta)} \right) \geq 1 - \gamma \quad \dots \textcircled{D}$$

(Prob. of a "state being typical")

$$\therefore \sum_{\substack{j \in \mathcal{S}_\Lambda \\ \text{(atypical)}}} p_j = \left\{ \begin{array}{l} \text{Prob. of a state} \\ \text{being atypical} \end{array} \right\} < \gamma \quad \dots \textcircled{E}$$

What about $\sum_{j \in \mathcal{S}_n} p_j$?
 (typical)

We have $R < h$:

Choose $\gamma > 0$ s.t. $\bullet R < h - \gamma$
 $\bullet 0 < \delta < \gamma/2$

Then

$$\sum_{j \in \mathcal{S}_n} p_j \leq 2^{[1 \wedge R]} \cdot 2^{-|A|(h-\delta)} \leq 2^{-|A|\gamma/2} \xrightarrow{\uparrow \uparrow \uparrow} 0 \quad \text{ⓕ}$$

(typical)

[since \bullet there are at most $|\mathcal{S}_n| \equiv 2^{[1 \wedge R]}$ typical states in \mathcal{S}_n
 \bullet for each typical state, the eigenvalue p_j is bounded by $2^{-|A|(h-\delta)}$ (from ⓓ)]

\therefore from ⓐ, ⓔ & ⓕ we have

$$\sum_{j \in \mathcal{S}_n} p_j \leq \gamma \quad \dots \text{ⓑ}$$



Data compression limit & limiting fidelity of the compression scheme for general decompositions (i.e., not necessarily orthogonal) of the density matrix ρ_Λ^ω :

Consider any representation of ρ_Λ^ω :

$$\rho_\Lambda^\omega = \sum_i p_i |\phi_i\rangle \langle \phi_i|$$

$|\phi_i\rangle \in \mathcal{H}_\Lambda$

- arbitrary vectors of unit norm
- not necessarily orthogonal

$$p_i \geq 0 \quad ; \quad \sum_i p_i = 1$$

To apply the data compression scheme:
consider an orthogonal projection:

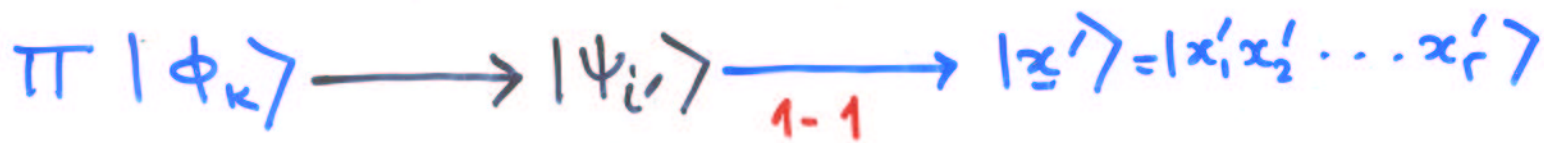
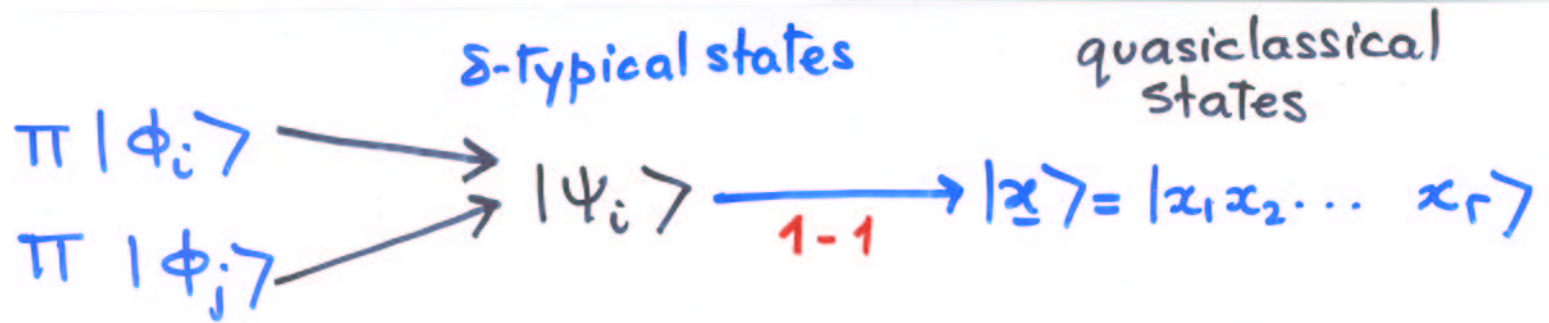
$$\Pi: \mathcal{H}_\Lambda \rightarrow \mathcal{C}$$

\mathcal{C} : subspace of \mathcal{H}_Λ s.t. the vectors

$\Pi|\phi_i\rangle$
 either collinear
 or orthogonal
 for different i

let $\mathcal{C} = \mathcal{M}_\delta$: subspace of δ -typical states

$$\therefore r := \dim \mathcal{C} = |\mathcal{M}_\delta| \leq 2^{[n\lambda H]}$$



- Each vector $\pi |\phi_k\rangle \mapsto \delta$ -typical state (not 1-1)
- Each δ -typical state \mapsto quasiclassical state

$| \underline{x} \rangle$; $\underline{x} \equiv$ binary string of length r
 $r = \dim \mathcal{C} \equiv \dim \mathcal{M}_\delta$

\therefore To each $\pi |\phi_k\rangle \longleftrightarrow$ quasiclassical state

Compression - Decompression Scheme

Encoding $\pi |\phi_k\rangle \xrightarrow{\text{(not 1-1)}} | \psi_j \rangle$

Compression $\mathcal{C} : | \psi_j \rangle \mapsto | \underline{x} \rangle$; $\underline{x} \in \{0,1\}^r$

Prescription for Decompression:
 (i.e., for decoding a state $| \psi_j \rangle$)

$\mathcal{D} : | \psi_j \rangle \mapsto | \phi_k \rangle$, where
 $\langle \phi_k | \psi_j \rangle = \max_i \langle \phi_i | \psi_j \rangle$

Fidelity

$$F_\Lambda := \sum_i p_i \langle \phi_i | \Pi | \phi_i \rangle$$

We prove that :

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} F_\Lambda = 1.$$

Conclusions

- class of non-i.i.d quantum information sources modelled by Gibbs state of a quantum spin system with strong interactions
- **Data compression limit** :
 $R_\infty = h$: von Neumann entropy rate of the source
- Key to data compression :

existence of a typical subspace of low dimension .