INTERNSHIP REPORT

Approximating Gibbs states with MPOs

Joël Charles-Rebuffé

Supervisor: Ángela Capel Cuevas

Junior Professor at the University of Tübingen

Contents

Part I
Introduction

All notations used in this report are standard quantum theory notations, but we define some of them and make several linear algebra recalls in the Appendix. My full work, including the report and its appendix, and my MATLAB codes can be found at https://github.com/JoeID/2023InternshipReport.

1 Motivation

Gibbs states are a special class of quantum states that aim to represent the interactions in a set of particles. A more formal definition will be given in the next section. They are of great interest in Quantum Computer Science, as they frequently arise in quantum machine learning [1] as well as in algorithms such as quantum SDP solvers [10], which is an extension of Linear Programming.

A major problem with quantum computing is the important noise that quantum states have to face in real quantum computers. While quantifying this noise in real quantum computers in not as straigthforward as it is for clasical computers because of the huge diversity in architectures and types of noise, for example a survey [11] has found a real-world implementation of an OR gate to fail in as much as 10-20% of the number of runs in some IBM 5-qubits quantum computers. This can prevent the good preparation of a Gibbs state.

The idea we will explore in this report is a quantum state learning algorithm, which can be formalized as follows. The algorithm receives a set of independent noised copies of an unknown Gibbs state, and uses quantum measurements to reconstruct an approximation of the original state, the goal being to minimize the distance between the original state and its approximation. We use further assumptions, presented in the next section.

2 Problem definition

Consider a chain of N particles $I \coloneqq \{1 \dots N\}$ (modeled as qubits for instance). For $J \subseteq I$, let us denote by $\phi(J)$ the $2^{|I|} \times 2^{|I|}$ self-adjoint matrix acting non-trivially on $|J|$ qubits representing the interactions of the particles in J. let us suppose that, for all $J \subseteq I$:

1. the interactions have finite range, that is: $\exists r \geq 1$ s.t. $\max(J) - \min(J) > r \implies \phi(J) = 0$. This property is also known as $r-$*locality*
2. the interactions are *translationally invariant*, that is: $\forall i \in \mathbb{Z}, \phi(\tau_i(J)) = \phi(J)$ where τ_i is the map that adds i to every element of J (defined only if we can do so while staying in I)
3. the strength of the interactions is uniformally bounded: $\exists \Gamma > 0$ s.t. $\forall J \subseteq I, \|\phi(J)\|_\infty \leq \Gamma$

Note that 3. is a consequence of 1. and 2. in our case where the number of particles is finite.

We then denote by

$$H_I \coloneqq \sum_{J \subseteq I \text{ s.t. } \max(J) - \min(J) \leq r} \phi(J)$$

the $2^N \times 2^N$ Hamiltonian matrix representing the interactions.

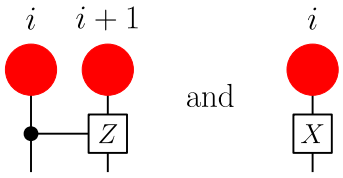

Figure 1: CZ and X quantum gates

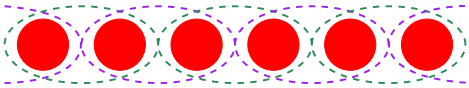

Figure 2: 2-local interactions in a chain of particles

For instance, with $r = 2$ and $\forall i, \phi(\{i, i+1\}) = CZ$ and $\phi(\{i\}) = X$, we get

$$H_I = \sum_{i=1}^{N-1} \mathbb{1}_{2^{i-1}} \otimes CZ \otimes \mathbb{1}_{2^{N-i-1}} + \sum_{i=1}^{N} \mathbb{1}_{2^{i-1}} \otimes X \otimes \mathbb{1}_{2^{N-i}}$$

which represents the situation of Figures 1 and 2.

Then, for any inverse temperature $\beta > 0$, we define

$$\rho_I \coloneqq \frac{e^{-\beta H_I}}{\mathrm{Tr}\left(e^{-\beta H_I}\right)}$$

the Gibbs state that represents the state of the system. We will simply write ρ when it is clear what system we are considering. The results presented in this report hold for every inverse temperature $\beta > 0$. Thus, for simplicity, we fix the temperature and omit β in the expression of the Gibbs states.

Let us suppose that the chain of particles I is divided into n subintervals of size m each:

$$I = A_1 \dots A_n \text{ with } |A_i| = m$$

For ρ a Gibbs state representing the interactions in I, we will suppose that we have access to copies of ρ that are guaranteed to be correct on small subsystems of the form A_i or $A_i A_{i+1}$. That is, we are given $\widetilde{\rho_1}, \dots, \widetilde{\rho_{n-1}}, \widetilde{\rho_{1:2}}, \dots, \widetilde{\rho_{n-1:n}}$ s.t.

$$\mathrm{tr}_{I \setminus A_i}\left(\widetilde{\rho_i}\right) = \mathrm{tr}_{I \setminus A_i}(\rho) \text{ and } \mathrm{tr}_{I \setminus A_i A_{i+1}}\left(\widetilde{\rho_{i:i+1}}\right) = \mathrm{tr}_{I \setminus A_i A_{i+1}}(\rho)$$

Given these copies, we want to approximate ρ with the best possible bound, that is compute a $\widetilde{\rho}$ s.t.

$$\|\rho - \widetilde{\rho}\|_1 \leq f(n, m)$$

Intuitively, $f(n, m)$ should be:

- decreasing in m because a bigger chunk size means that the correct parts of the ρ copies are bigger, thus improving the approximation
- increasing in n because more chunks means a chunk size that is smaller compared to the total system size

In the end, for $n \in \mathbb{N}$ and $\epsilon > 0$ fixed, we want to know what is the minimum chunk size m s.t. $f(n, m) \leq \epsilon$.

3 State of the art

Let ρ be our original Gibbs state and $\widetilde{\rho}$ denote the approximation of it. N is the number of particles.

There are several previous works in the literature that have addressed the problem of approximating Gibbs states by matrix product operators (MPOs) or projected entangled pair operators (PEPOs), their analogues in higher dimensions, like [6, 7, 8]. Notably, in [8], Kuwahara, Alhambra and Anshu present an algorithm to approximate Gibbs states by MPOs.

Theorem 1. *Kuwahara's algorithm*
If the temperature T is s.t $T = \Omega\left(\frac{1}{log(N)}\right)$, this algorithm gives an approximation $\widetilde{\rho}$ s.t.

$$\|\rho - \widetilde{\rho}\|_1 \leq \epsilon$$

as long as $m = \Omega\left(\frac{N}{\epsilon}\right)$.

The approach is temperature dependant and does not have the same purpose as ours. Indeed, their goal is to reconstruct an easy to compute, classically speaking, approximation of the original Gibbs state using truncated power series in place of the exponents and then controlling them with a technique called cluster expansions. The base idea behind their approach is however very similar to the one behind our first approach (section 1), even if the mathematicals tools are very different, so it is of great interest to us.

In [9], the authors showed several new examples of approximations by MPOs. In particular, they provide an algorithm to derive an MPO to approximate a thermal state in 1D with block dimension $\exp\left(\tilde{\Omega}\left(\beta^{\frac{2}{3}} + \left(\beta log\left(\frac{N}{\epsilon}\right)\right)^{\frac{1}{2}}\right)\right)$ with $\tilde{\Omega}(N) = \Omega(N \log(N))$. In the current project, we construct an MPO approximation of a thermal state that improves on such a dimension, as it scales at most with N and ϵ as $\Omega\left(\frac{log\left(\frac{n}{\epsilon}\right)}{log\left(log\left(\frac{n}{\epsilon}\right)\right)}\right)$.

Part II
Mathematical toolbox

In this part we present basic mathematical definitions, notation, lemmas as well as a short introduction to the Tensor Network notation. More details and in-depth explanations can be found in Álvaro M. Alhambra[13].

1 Definitions and notation

We use the Schatten norms, defined as: for a given operator (matrix) M with singular values $\{\lambda_i\}_i$ and $p \in [1, +\infty)$

$$\|M\|_p := \left(\sum_i |\lambda_i|^p\right)^{\frac{1}{p}}$$

And $\|M\|_\infty := \lim_{p \to +\infty} \|M\|_p$.

The closeness of two quantum states σ, ρ is typically measured with the *trace distance* $\frac{1}{2}\|\sigma - \rho\|_1$. This is because if we are given two different quantum states $\sigma = \sum_j q_j |j\rangle \langle j|, \rho = \sum_j p_j |j\rangle \langle j|$, the probability of getting a different result (i.e differentiating them) when measuring both is

$$\frac{1}{2}\sum_j |q_j - p_j| = \frac{1}{2}\|\sigma - \rho\|_1$$

If σ and ρ are not expressed in the same basis, the remark still applies, see [12].

Proposition 1. *Hölder's inequality:* $\forall p, q_1, q_2 \in [1, +\infty]$ *s.t.* $\frac{1}{p} = \frac{1}{q_1} + \frac{1}{q_2}$, $\forall M, N$ *multipliable matrices:*

$$\|MN\|_p \leq \|M\|_{q_1} \|N\|_{q_2}$$

Corollary 1. *A useful application is:*

$$\|MN\|_1 \leq \|M\|_1 \|N\|_\infty$$

Let I be a spin chain of N particles with r-local interactions.

- If I is divided into three subintervals $I = ABC$ (with A or C possibly empty), the marginal state ρ_B is defined as:
$$\rho_B \coloneqq \mathrm{tr}_A\left(\mathrm{tr}_C\left(\rho^{ABC}\right)\right)$$
- In order not to mix up the notations for I and the identity matrix, we will designate by $\mathbb{1}_A$ the $2^{|A|}$-dimensional identity matrix for any subinterval $A \subseteq I$. We will drop the $_A$ when it can be clearly deduced from context

When considering ρ_B or ρ^B in the system J, we will extend these matrix with $\mathbb{1}$ so that they match the $2^{|J|} \times 2^{|J|}$ dimension of ρ^J. Formally:

- Considered in system ABC, ρ_B designates in fact $\mathbb{1}_A \otimes \mathrm{tr}_A\left(\mathrm{tr}_C\left(\rho^{ABC}\right)\right) \otimes \mathbb{1}_C$
- Considered in system ABC, ρ^B designates in fact $\frac{1}{Z}\left(\mathbb{1}_A \otimes e^{-H_B} \otimes \mathbb{1}_C\right)$ where Z is s.t. $\mathrm{Tr}\left(\rho^B\right) = 1$

Note that we do not need to normalize ρ_B because ρ^{ABC} is already a mixed state, so we simply "fill the gaps" in A and C with $\mathbb{1}_A$ and $\mathbb{1}_C$.

In order not to use too heavy notations, we will simply write ρ_B and ρ^B when the context makes it clear what overall system we are considering.

When considering I as a juxtaposition of intervals (i.e. $I = A_1 \dots A_n$), we will use the following notation:

$$\rho_i \coloneqq \rho_{A_i} \text{ and } \rho_{i:j} \coloneqq \rho_{A_i \dots A_j}$$

According to our assumptions, our quantum learning algorithm will have access, as an input, to the marginals every ρ_i and $\rho_{i:i+1}$.

2 Useful properties and definitions

From now on, for the sake of conciseness $\|\cdot\|$ will mean $\|\cdot\|_\infty$. If not provided, the proofs of these properties can be found in the Appendix. Let $I = [\![1, N]\!]$ be a chain of particles.

2.1 Technical facts

Lemma 1. *States that do not live in the same Hilbert space commute: if A, B are separate subintervals of I, $\rho_A \rho_B = \rho_B \rho_A$.*

Remark 1. This can be extended to separate subsets: $\forall J, J' \subseteq I$ s.t. $J \cap J' = \emptyset$,

$$\rho_J \rho_{J'} = \rho_{J'} \rho_J$$

Lemma 2. *For $J \subseteq I$, the applications $\mathrm{tr}_J\left(\bullet \rho^J\right)$ and $\mathrm{tr}_J\left(\rho^J \bullet\right)$ are contractive, i.e.*

$$\begin{aligned} \text{For } M \text{ a } 2^{|I|} \times 2^{|I|} \text{ matrix}: \left\|\mathrm{tr}_J\left(M\rho^J\right)\right\| &\leq \|M\| \\ \left\|\mathrm{tr}_J\left(\rho^J M\right)\right\| &\leq \|M\| \end{aligned}$$

Lemma 3. *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be the Hilbert spaces associated with disjoint subsystems A and B of I, respectively. Then:*

$$\forall Z_{AB} \in \mathcal{H}_I, \forall Y_A \in \mathcal{H}_A, \quad \mathrm{tr}_B\left(Z_{AB}\left(Y_A \otimes \mathbb{1}_B\right)\right) = \mathrm{tr}_B\left(Z_{AB}\right) Y_A$$

Corollary 2. *If I consists of at least A and B separate subsystems:*

$$\mathrm{tr}_B\left(\rho^I e^{-H_A}\right) = \mathrm{tr}_B\left(\rho^I\right) e^{-H_A}$$

2.2 Matrix closeness

Definition 1. A matrix A is said to be *ϵ-close to B* if $\|A - B\| \leq \epsilon$. We will denote this property by

$$A \approx B\,[\epsilon]$$

Proposition 2. *If $A \approx \mathbb{1}\,[\epsilon_1]$ and $B \approx \mathbb{1}\,[\epsilon_2]$ then $AB \approx \mathbb{1}\,[(1+\epsilon_1)(1+\epsilon_2) - 1]$*

Corollary 3. *By induction, if $A_1, \ldots, A_n$ are matrices s.t. $\forall i$, $A_i \approx \mathbb{1}\,[\epsilon_i]$ then:*

$$A_1 \ldots A_n \approx \mathbb{1}\left[\prod_{i=1}^{n}(1+\epsilon_i) - 1\right]$$

Proof.

$$\begin{aligned} \|AB - \mathbb{1}\| &= \|(A-\mathbb{1})(B-\mathbb{1}) + (A-\mathbb{1}) + (B-\mathbb{1})\| \\ &\leq \|A-\mathbb{1}\|\,\|B-\mathbb{1}\| + \|A-\mathbb{1}\| + \|B-\mathbb{1}\| \text{ by submultiplicity of } \|\cdot\| \\ &\leq \epsilon_1\epsilon_2 + \epsilon_1 + \epsilon_2 \\ &= (1+\epsilon_1)(1+\epsilon_2) - 1 \end{aligned}$$

□

Proposition 3. *If A, B are hermitian matrices,*

$$AB \approx \mathbb{1}\,[\epsilon] \Leftrightarrow BA \approx \mathbb{1}\,[\epsilon]$$

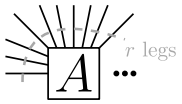

Figure 3: r-rank tensor A

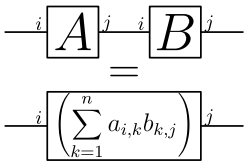

Figure 4: *Tensor contraction* rule : AB product

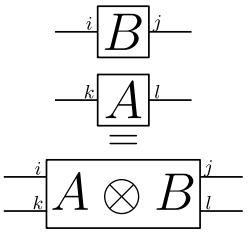

Figure 5: *Tensor product* rule : AB tensor product using TN notation

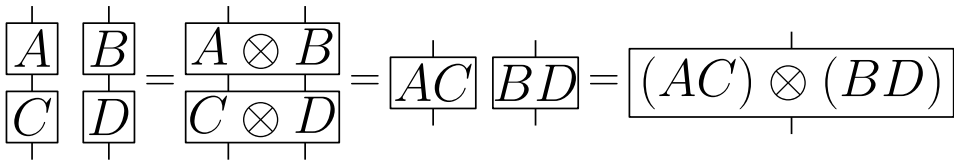

Figure 6: graphical illustration of $(AB) \otimes (CD) = (A \otimes C)(B \otimes D)$

Proof. We only need to prove one way. Let us suppose that $\|AB - \mathbb{1}\| \leq \epsilon$. Then:

$$\begin{aligned} \|BA - \mathbb{1}\| &= \left\|(BA - \mathbb{1})^\dagger\right\| \text{ by property of Schatten norms} \\ &= \left\|A^\dagger B^\dagger - \mathbb{1}\right\| = \|AB - \mathbb{1}\| \leq \epsilon \end{aligned}$$

□

2.3 Tensor network notation and MPOs

Definition 2. A Tensor Network (TN) is a convenient way to represent tensor and product operations. Formally, a tensor A belonging to some Hilbert space $\mathcal{H}_I = \bigotimes_{i=1}^{r} \mathcal{H}_{A_i}$ (i.e a r-rank tensor) is represented by a node with r legs, denoting the r coordinates we need to use to *navigate* A: see Figure 3.
Note that a 0-rank tensor is a scalar, a 1-rank tensor is a vector and a 2-rank tensor is a matrix.

This notation is combined with the following conventions:

- *Tensor contraction :* When the legs of two different nodes get connected, the resulting merged node denotes the tensor where the index corresponding to the legs is summed over. For instance, if $A = [a_{i,j}]_{i,j=1\ldots n}$ and $B = [b_{i,j}]_{i,j=1\ldots n}$, the node obtained by connecting the j index of A with the i index of B is the matrix product AB: see Figure 4.
- *Tensor product :* Two nodes situated next to each other without being connected denote their tensor product: see Figure 5

We can aply these merging rules in any order we want thanks to the property $(AB) \otimes (CD) = (A \otimes C)(B \otimes D)$: see Figure 6

Definition 3. A Matrix Product Operator (MPO) is a TN of the form depicted in Figure 7

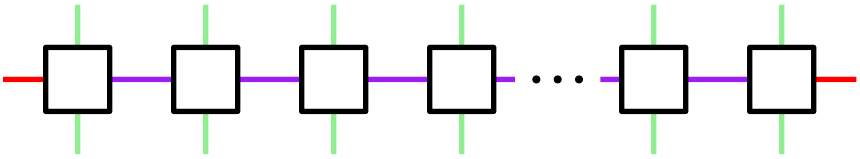

Figure 7: Matrix Product Operator (MPO)

MPOs are a tool to represent some matrix and tensor products in a simple and concise way, which the usual sum and product notation does not allow. More precisely, MPOs are particularly good at compressing tensor contractions of local terms (i.e. that only act non-trivially on some qubits), and that is what we are going to use them for.

A fundamental 1969 article from Araki [2] provides us with some results, that were reformulated as corollaries suited for our needs in the Corollary 3.4 in [3], as follows.

Theorem 2. *Araki's corollary results*
Consider a particle chain. For A, B subintervals of I, let $E_{A,B} \coloneqq e^{-H_{AB}} e^{H_A} e^{H_B}$. There exists a constant $\Delta > 1$ depending on Γ, β, r s.t.

1. If the interval I includes two subintervals AB: $I = I_1 ABI_2$, then
$$\|E_{A,B}\|, \left\|E_{A,B}^{-1}\right\| \leq \Delta$$

2. If $I = I_1 ABI_2$, then
$$\|E_{I_1 A,BI_2} - E_{A,B}\|, \left\|E_{I_1 A,BI_2}^{-1} - E_{A,B}^{-1}\right\| \leq \frac{\Delta^l}{\left(\lfloor \frac{l}{r} \rfloor + 1\right)!} \text{ with } l = \min\left(|A|, |B|\right)$$
that is, $E_{I_1 A,BI_2}$ is superexponentially close to $E_{A,B}$ in terms of the size of A and B. This does not depend on $I_1$and I_2. Something that is worth noting is that we can have I_1 or I_2 empty.

3. If $I = I_1 ABI_2$, then
$$\left\|\mathrm{tr}_B\left(E_{A,B}\rho^B\right)\right\|, \left\|\mathrm{tr}_B\left(E_{A,B}\rho^B\right)^{-1}\right\| \leq \Delta$$

Part III
Main results

1 Linear-shaped approach

This approach was inspired by a lemma in one of my supervisor's published papers [3]. As a consequence, the lemma 4 and its annexed proof come directly from that paper. The idea of the linear-shaped approach as well as a first proof of the theorem 3 also come from my supervisor. All the rest, including the matrix closeness notation introduced in 2.2 and the subsequent proof

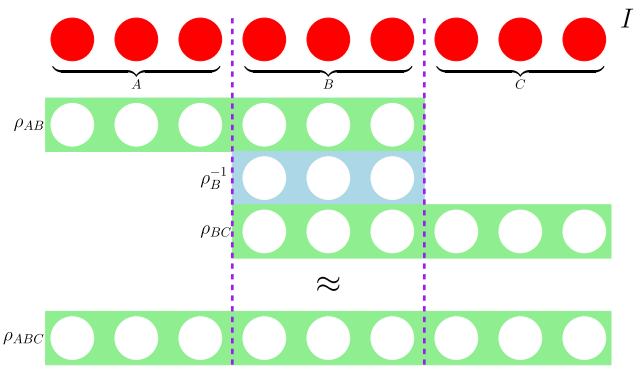

Figure 8: ρ_{ABC} is close to the product $\rho_{AB}\rho_B^{-1}\rho_{BC}$

of theorem 3 that relies heavily on it, constitute my contribution.

1.1 Base lemma

Lemma 4. *If I is a spin chain divided into subintervals $I = ABC$ with $|B| = m$ then there exists a superexponentially decaying function $\epsilon(m)$ s.t.*

$$\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} \approx \mathbb{1}\left[\epsilon(m)\right]$$

Proof. The full proof is quite complex and can be found in Appendix. It consists of translating $\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} - \mathbb{1}$ into a product whose terms can be bounded using Araki's results. □

Intuitively, this means that we can decompose a state over three subsystems in terms of a product of states over one or two subsystems, while losing only a small amount of precision: see Figure 8.

Remark 2. $\epsilon(m)$ can be written using the convenient form $\epsilon(m) := K\frac{G^m}{\lfloor\frac{m}{2r}\rfloor!}$ where K, G are constants depending on Γ, r, β.

Corollary 4. *We switched to the ∞-norm because it's stronger, and we can therefore deduce a result with the 1-norm :*

$$\begin{aligned}\left\|\rho_{AB}\rho_B^{-1}\rho_{BC} - \rho_{ABC}\right\|_1 &\leq \left\|\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} - \mathbb{1}\right\|_\infty \left\|\rho_{ABC}\right\|_1 \quad \textit{with Hölder's inequality} \\ &= \left\|\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} - \mathbb{1}\right\|_\infty \leq \epsilon(m)\end{aligned}$$

We can extend the preceding lemma to the case where I is bigger than simply ABC :

Lemma 5. *If I is a spin chain divided into subintervals $I = ABCD$ with $|B| = m$ then there exists a superexponentially decaying function $\epsilon'(m)$ s.t.*

$$\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} \approx \mathbb{1}\left[\epsilon'(m)\right]$$

Proof. in Appendix □

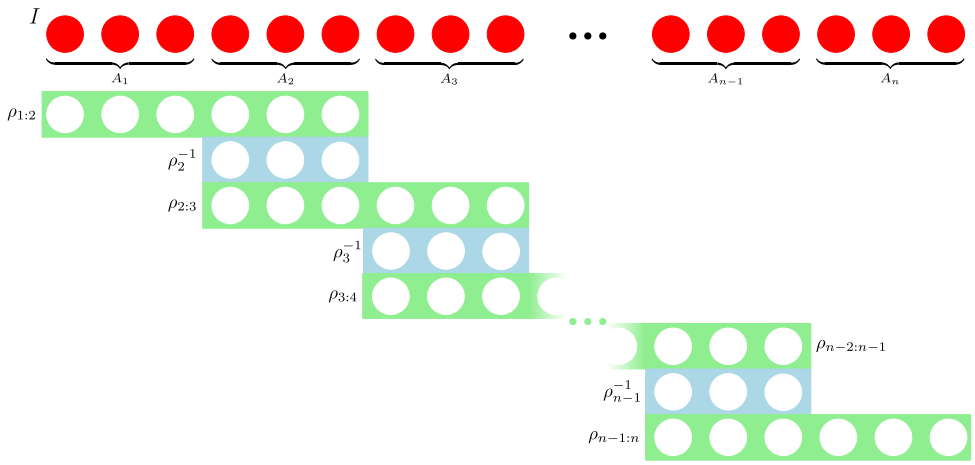

Figure 9: Linear-shaped construction: ρ is close to $\rho_{1:2}\rho_2^{-1}\rho_{2:3}\dots\rho_{n-1}^{-1}\rho_{n-1:n}$

Similarly, we can prove the following extension :

Lemma 6. *If I is a spin chain divided into subintervals $I = I_1ABCI_2$ with $|B| = m$ then there exists a superexponentially decaying function $\epsilon(m)$ s.t.*

$$\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} \approx \mathbb{1}\left[\epsilon(m)\right]$$

1.2 Theorem

An idea naturally comes next: if $\rho = \rho_{1:n}$, this lemma states that $\rho \approx \rho_{1:n-1}\rho_{n-1}^{-1}\rho_{n-1:n}$. Applying it again to $\rho_{1:n-1}$, then $\rho_{1:n-2}$, etc gives a MPO approximation for the whole state: see Figure 9. Formally speaking:

Theorem 3. *If I is a spin chain divided into subintervals $I = [\![1, n]\!]$ of size m each, then, under the assumption that $n \geq 2$,*

$$\rho_{1:2}\left(\prod_{i=2}^{n-1}\rho_i^{-1}\rho_{i:i+1}\right)\rho_{1:n}^{-1} \approx \mathbb{1}\left[(1+\epsilon(m))^{n-2}-1\right] \tag{1}$$

Proof.

$$\rho_{1:2}\left(\prod_{i=2}^{n-1}\rho_i^{-1}\rho_{i:i+1}\right)\rho_{1:n}^{-1} = \prod_{i=2}^{n-1}\rho_{1:i}\rho_i^{-1}\rho_{i:i+1}\rho_{1:i+1}^{-1}$$

For $1 < i < n$, the base lemma applied to I with $A \coloneqq A_1\dots A_{i-1}$, $B \coloneqq A_i$ and $C \coloneqq A_{i+1}$ gives:

$$\rho_{1:i}\rho_i^{-1}\rho_{i:i+1}\rho_{1:i+1}^{-1} \approx \mathbb{1}\left[\epsilon(m)\right]$$

By the corollary 3:

$$\prod_{i=2}^{n-1}\rho_{1:i}\rho_i^{-1}\rho_{i:i+1}\rho_{1:i+1}^{-1} \approx \mathbb{1}\left[(1+\epsilon(m))^{n-2}-1\right]$$

□

Let's now examine the

Proposition 4. *Let $\epsilon_0 > 0$. Denote by $\delta(n,m) \coloneqq (1+\epsilon(m))^{n-2} - 1$ the bound we found. Then, a sufficient condition for $\delta(n,m) \leq \epsilon_0$ is*

$$m = \Omega\left(\frac{\log\left(\frac{n}{\epsilon}\right)}{\log\left(\log\left(\frac{n}{\epsilon}\right)\right)}\right)$$

Proof. Sketch of the proof : we use Stirling's inequality to bound

$$\begin{aligned}\epsilon(m) &= K\frac{G^m}{\lfloor\frac{m}{2r}\rfloor!} \\ &\leq K\left(\frac{Ge^{\frac{1}{2r}}}{m^{\frac{1}{2r}}}\right)^m\end{aligned}$$

and the rest is completely analogous to the proof of Corollary 3.3 of [5]. □

2 Divide and conquer approach

2.1 Presentation

In order to approximate $\rho = \rho_{1:n}$, our approach is "divide and conquer". Let us suppose that $n = 2^k$: we will first recursively approximate the "two halves" of ρ: $\rho_{1:2^{k-1}}$ and $\rho_{2^{k-1}+1:n}$. Then merge these approximations to get ρ. Let us define formally our recursive approach:

Definition 4. For $\rho_{i:j}$ a partial trace of ρ, let us denote by $\widetilde{\rho_{i:j}}$ our approximation of it.

Definition 5. Our "divide and conquer" approximation is defined as follows:

$$\begin{cases} \forall i, & \widetilde{\rho_i} \coloneqq \rho_i \text{ and } \widetilde{\rho_{i:i+1}} \coloneqq \rho_{i:i+1} \\ \forall \text{ interval } [i,j] \text{ of size } 2^{k+1} & \widetilde{\rho_{i:j}} \coloneqq \underbrace{\widetilde{\rho_{i:i+2^k}}}_{\text{left approx.}} \underbrace{\widetilde{\rho_{i+2^k+1:j}}}_{\text{right approx.}} \underbrace{\rho_{i+2^k}^{-1}\rho_{i+2^k+1}^{-1}\rho_{i+2^k:i+2^k+1}}_{\text{to "glue" the approx. together}} \end{cases}$$

Graphically, one can visualize this approach: considering the spin chain $[\![1,24]\!]$:

2.2 Base lemma

Lemma 7. *If I is a spin chain divided into subintervals $I = ABCD$ with $|B|, |C| = m$ and $|A|, |D| \geq m$, then there exists a superexponentially small function $\epsilon(m)$ s.t.*

$$\rho_{AB}\rho_{CD}\rho_B^{-1}\rho_C^{-1}\rho_{BC}\rho_{ABCD}^{-1} \approx \mathbb{1}[\epsilon(m)]$$

Intuitively, this means we can decompose a state over four subsystems in terms of states over one or two subsystems: see Figure 11.

Proof. The proof is completely analogous to the one of lemma 4, albeit more complicated because we're dealing with four subsystems instead of three. A detailed sketch of it can be found in the VIII. □

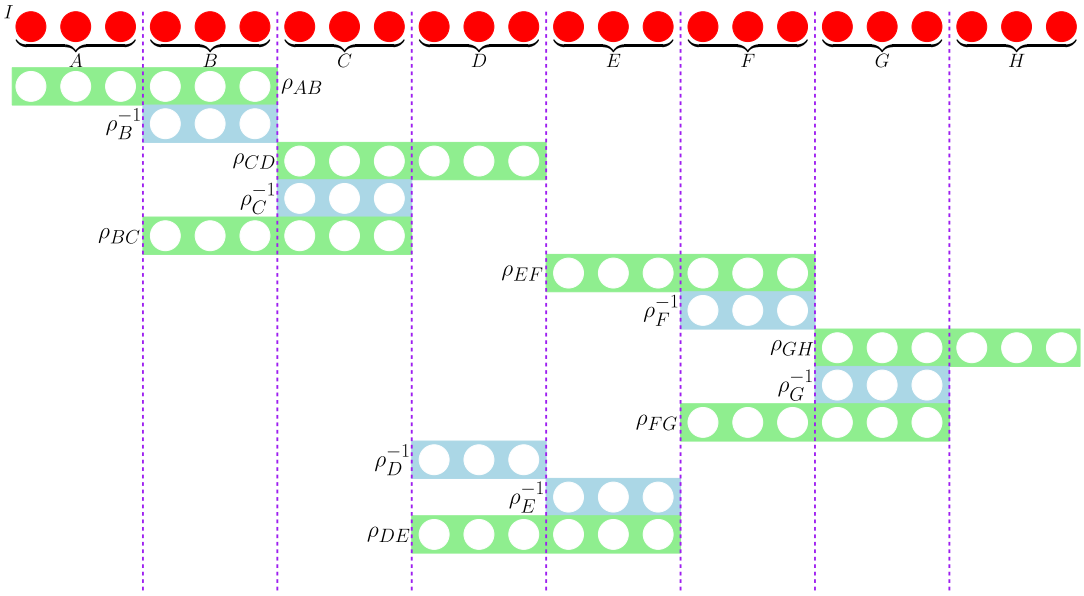

Figure 10: "divide and conquer" approach with $n = 8, m = 3$

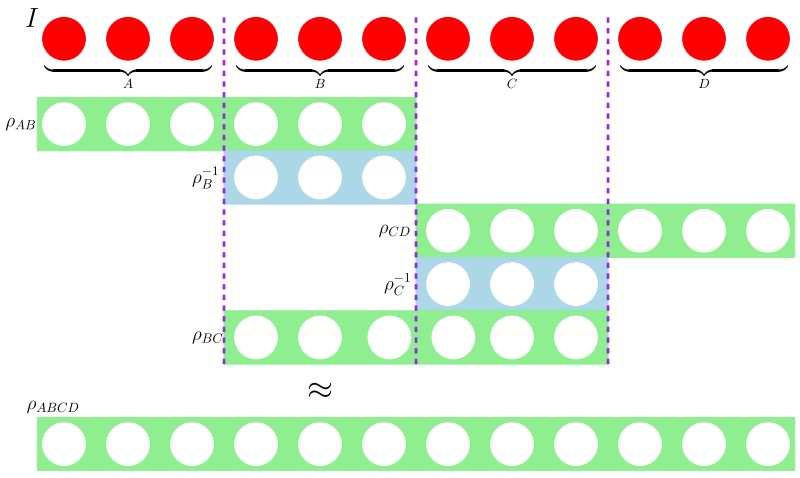

Figure 11: ρ_{ABCD} is close to the product $\rho_{AB}\rho_{CD}\rho_B^{-1}\rho_C^{-1}\rho_{BC}$

2.3 Theorem

Then we want to find an expression for $\delta(k,m)$ s.t. the following theorem holds:

Theorem 4. *If $I = [\![i,j]\!]$ is a spin chain divided into $|I| = n = 2^k$ subintervals of size m each, then*

$$\widetilde{\rho_{i:j}}\rho_{i:j}^{-1} \approx \mathbb{1}\left[\delta(k,m)\right]$$

For $n = 1, 2$, we have $\widetilde{\rho_{i:j}} = \rho_{i:j}$ so we can define $\delta(0,m), \delta(1,m) := 0$. let us find an expression for $\delta(k,m)$ in terms of $\delta(k-1,m)$:

let us suppose that $I := [i,j]$ is a spin chain divided into $|I| = n = 2^k$ subintervals and let $p, q := \frac{i+j-1}{2}, \frac{i+j+1}{2}$ be the two middle elements of $\mathbb{1}$. let us suppose (recursion hypothesis) that:

$$\widetilde{\rho_{i:p}}\rho_{i:p}^{-1} \approx \mathbb{1}\left[\delta(k-1,m)\right], \widetilde{\rho_{q:j}}\rho_{q:j}^{-1} \approx \mathbb{1}\left[\delta(k-1,m)\right]$$

Then:

$$\begin{aligned}
\widetilde{\rho_{i:j}}\rho_{i:j}^{-1} &= \widetilde{\rho_{i:p}}\rho_p^{-1}\widetilde{\rho_{q:j}}\rho_q^{-1}\rho_{p:q}\rho_{i:j}^{-1} \text{ by definition} \\
&= \widetilde{\rho_{i:p}}\rho_{i:p}^{-1}\rho_{i:p}\rho_p^{-1}\widetilde{\rho_{q:j}}\rho_{q:j}^{-1}\rho_{q:j}\rho_q^{-1}\rho_{p:q}\rho_{i:j}^{-1} \\
&= \widetilde{\rho_{i:p}}\rho_{i:p}^{-1}\widetilde{\rho_{q:j}}\rho_{q:j}^{-1}\rho_{i:p}\rho_p^{-1}\rho_{q:j}\rho_q^{-1}\rho_{p:q}\rho_{i:j}^{-1} \text{ as } \rho_{q:j}, \widetilde{\rho_{q:j}} \text{ live in } \mathcal{H}_{q:j} \text{ and } \rho_{i:p}, \widetilde{\rho_{i:p}} \text{ in } \mathcal{H}_{i:p}
\end{aligned}$$

It follows that

$$\begin{aligned}
\widetilde{\rho_{i:j}}\rho_{i:j}^{-1} &= \underbrace{\widetilde{\rho_{i:p}}\rho_{i:p}^{-1}}_{\approx\mathbb{1}[\delta(k-1,m)]} \underbrace{\widetilde{\rho_{q:j}}\rho_{q:j}^{-1}}_{\approx\mathbb{1}[\delta(k-1,m)]} \underbrace{\rho_{i:p}\rho_p^{-1}\rho_{q:j}\rho_q^{-1}\rho_{p:q}\rho_{i:j}^{-1}}_{\approx\mathbb{1}[\epsilon(m)] \text{ by the lemma}} \\
&\approx \mathbb{1}\left[\left(1+\delta(k-1,m)\right)^2\left(1+\epsilon(m)\right)-1\right]
\end{aligned}$$

.

Then, $\left(1+\delta(k-1,m)\right)^2\left(1+\epsilon(m)\right)-1 = \left(1+\epsilon(m)\right)\delta(k-1,m)^2 + 2\left(1+\epsilon\right)\delta(k-1,m) + \epsilon(m)$. We can subsequently set $\delta(k,m) := \left(1+\epsilon(m)\right)\delta(k-1,m)^2 + 2\left(1+\epsilon(m)\right)\delta(k-1,m) + \epsilon(m)$. So, let us study, for $\epsilon > 0$, $u_k := \delta(k,m)$ defined by:

$$\begin{cases} u_0, u_1 & = 0 \\ \text{for } k \geq 2, u_k & = (1+\epsilon)\,u_{k-1}^2 + 2(1+\epsilon)\,u_{k-1} + \epsilon \end{cases} \tag{2}$$

u_k is the approximation error for $n = 2^k$ intervals. To study $(u_k)_{k\geq 0}$, we will prove the following:

Claim 1. Let α be the only positive root of $P(X) := X^2 - 2X - \epsilon = 0$: $\alpha := 1 + \sqrt{1+\epsilon}$. Then:

$$\forall k \geq 2, u_k = \left(1+\frac{\epsilon}{\alpha}\right)^{2^k-2} - 1 \tag{3}$$

Proof. First, notice that $2 \leq \alpha \leq 2+\epsilon$. Then, by definition of α,

$$\begin{aligned}
\left(1+\frac{\epsilon}{\alpha}\right)^2 &= 1 + \frac{2\epsilon}{\alpha} + \frac{\epsilon^2}{\alpha^2} \\
&= 1 + \epsilon\left[\frac{2\alpha+\epsilon}{\alpha^2}\right] \\
&= 1+\epsilon
\end{aligned}$$

Then, the recurrence formula defining (u_k) can be rewritten as follows:

$$\begin{aligned}\forall k \geq 2, u_k + 1 &= (1+\epsilon)\, u_{k-1}^2 + 2\,(1+\epsilon)\, u_{k-1} + (1+\epsilon) \\ &= \left(1+\frac{\epsilon}{\alpha}\right)^2 u_{k-1}^2 + 2\left(1+\frac{\epsilon}{\alpha}\right)^2 u_{k-1} + \left(1+\frac{\epsilon}{\alpha}\right)^2 \\ &= \left(1+\frac{\epsilon}{\alpha}\right)^2 \left(u_{k-1}^2 + 2u_{k-1} + 1\right) \\ &= \left(1+\frac{\epsilon}{\alpha}\right)^2 (u_{k-1}+1)^2\end{aligned}$$

It follows that:

$$\forall k \geq 2, u_k + 1 = \left(1+\frac{\epsilon}{\alpha}\right)^{2^k-2} \left(\underbrace{u_1}_{=0} + 1\right)$$

$$\text{and } u_k = \left(1+\frac{\epsilon}{\alpha}\right)^{2^k-2} - 1 \text{ with } 2 \leq \alpha \leq 2+\epsilon$$

□

Subsequently:

$$\delta(n,m) = \left(1+\frac{\epsilon(m)}{\alpha}\right)^{n-2} - 1 \tag{4}$$

That is almost exactly the same bound as the linear-shaped approach: see Eq. 1. The only difference between the two is that we now have $\frac{\epsilon(m)}{\alpha}$ instead of $\epsilon(m)$, but keep in mind the the $\epsilon(m)$'s of both inequations come from different base lemmas, they are the same up to a constant factor. The relationship between n and m is therefore unchanged in order of magnitude.

But, when paying attention to the recurrence formula for (u_k), we can make more easily than before make the following observation:

As long as $u_k \leq 1$, then $u_{k+1} \leq 3\,(1+\epsilon)\, u_k + \epsilon$, thus $u_k = O\left((3+3\epsilon)^k\right)$

As soon as $u_k \geq 1$, then $u_k \geq (1+\epsilon)\, u_k^2$ and $u_k = \Omega\left((1+\epsilon)^{2^k}\right)$

In order words, the error stays exponential in k (so polynomial in $n = 2^k$) until it gets bigger than 1, at which point it becomes exponential in n. We can see this clear cut when plotting $log\,(u_k)$ in terms of k: see plot 14 on page 18.

3 Every other approach

For $n \geq 2$, let $S_n := \left(\bigcup_{i=2}^{n-1} \left\{\rho_i^{-1}\right\}\right) \bigcup \left(\bigcup_{i=1}^{n-1} \left\{\rho_{i:i+1}\right\}\right)$ the set of 1- and 2-consecutive-subsystems marginals of ρ. Note that $|S_n| = 2n-3$. Let also A_n be the set of products containing exactly once every element of S_n . We can consider A_n as the "set of candidates for approximating ρ". We explored two elements of A_n:

- $\widetilde{\rho}_{lin}$ the linear-shaped approximation of ρ defined in Part III section 1.2
- $\widetilde{\rho}_{dac}$ the divide-and-conquer approximation of ρ defined in Part III section 2.1

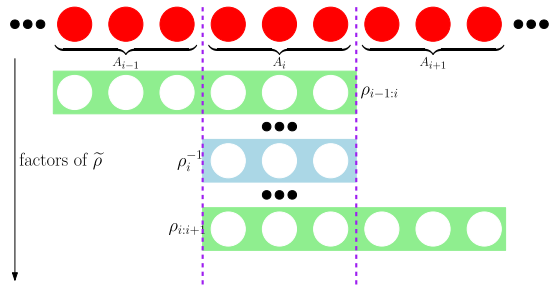

Figure 12: ρ_i appears between $\rho_{i-1:i}$ and $\rho_{i:i+1}$ in the factors of $\widetilde{\rho}$

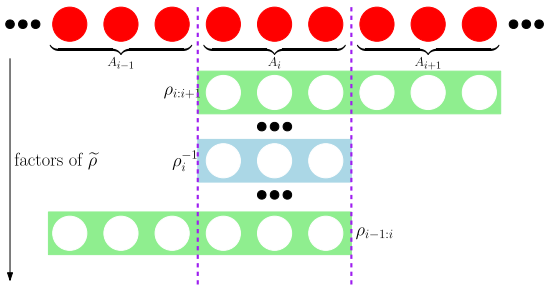

Figure 13: ρ_i appears between $\rho_{i:i+1}$ and $\rho_{i-1:i}$ in the factors of $\widetilde{\rho}$

We got for both times the same bound in order of magnitude: there exists a superexponentially decaying function $\epsilon(m)$ s.t.:

$$\|\widetilde{\rho_{dac}} - \rho\|_1, \|\widetilde{\rho_{lin}} - \rho\|_1 \leq (1+\epsilon(m))^{n-2} - 1$$

We can generalise these results:

Theorem 5. *Let $n \geq 2$ AO_n (Approximation in Order) be the subset of A_n s.t. $\forall \widetilde{\rho} \in A_n, \forall i \in [\![2, n-1]\!]$, the occurrence of ρ_i^{-1} in the factors of $\widetilde{\rho}$ comes between the occurrences of $\rho_{i:i+1}$ and $\rho_{i-1:i}$, in both orders. In other words, every $\widetilde{\rho} \in A_n$ can be written, for all i, either*

$$\widetilde{\rho} = A\rho_{i-1:i}B\rho_i^{-1}C\rho_{i:i+1}D \tag{5}$$

$$\text{or } \widetilde{\rho} = A\rho_{i:i+1}B\rho_i^{-1}C\rho_{i-1:i}D \tag{6}$$

As represented by the figures 12 and 13. Note that A, B, C, D are possibly equal to $\mathbb{1}$. Then, there exists a superexponentially decaying function $\epsilon(m)$ s.t.:

$$\forall \widetilde{\rho} \in AO_n, \|\widetilde{\rho} - \rho\|_1 \leq (1+\epsilon(m))^{n-2} - 1$$

Proof. We will in fact prove $P_n : \forall \widetilde{\rho_{1:n}} \in AO_n, \left\|\widetilde{\rho_{1:n}}\rho_{1:n}^{-1} - \mathbb{1}\right\| \leq (1+\epsilon(m))^{n-2} - 1$ by recursion of n:

- Case $n = 2$: $AO_2 = \{\rho_{1:2}\}$ so $\forall \widetilde{\rho} \in AO_2, \left\|\widetilde{\rho}\rho^{-1} - \mathbb{1}\right\| = \left\|\rho_{1:2}\rho_{1:2}^{-1} - \mathbb{1}\right\| = 0 = (1+\epsilon(m))^{n-2} - 1$
- Case $n \geq 3$: let us suppose that P_{n-1} is true. Let $\widetilde{\rho_{1:n}} \in AO_n$. Then:
  1. If $\widetilde{\rho_{1:n}} = A\rho_{n-2:n-1}B\rho_{n-1}^{-1}C\rho_{n-1:n}D$, then, as $\rho_{n-2:n-1}, \rho_{n-1}^{-1}, \rho_{n-1:n}$ appear exactly once in $\widetilde{\rho}$, the subproducts A, B, C, D live in $\mathcal{H}_{1:n-2}$. Thus, they commute with ρ_{n-1}^{-1} and $\rho_{n-1:n}$. So:

$$\begin{aligned}\widetilde{\rho_{1:n}}\rho_{1:n}^{-1} &= A\rho_{n-2:n-1}B\rho_{n-1}^{-1}C\rho_{n-1:n}D\rho_{1:n}^{-1} \\ &= A\rho_{n-2:n-1}BCD\rho_{n-1}^{-1}\rho_{n-1:n}\rho_{1:n}^{-1} \\ &= A\rho_{n-2:n-1}BCD\rho_{1:n-1}^{-1}\rho_{1:n-1}\rho_{n-1}^{-1}\rho_{n-1:n}\rho_{1:n}^{-1}\end{aligned}$$

  Note that $A\rho_{n-2:n-1}BCD \in AO_{n-1}$. So:

$$A\rho_{n-2:n-1}BCD\rho_{1:n-1}^{-1} \approx \mathbb{1}\left[(1+\epsilon(m))^{n-3} - 1\right] \text{ by } P_{n-1}$$

$$\rho_{1:n-1}\rho_{n-1}^{-1}\rho_{n-1:n}\rho_{1:n}^{-1} \approx \mathbb{1}\left[\epsilon(m)\right] \text{ by lemma 4}$$

And:

$$\begin{aligned}\widetilde{\rho_{1:n}}\rho_{1:n}^{-1} &\approx \mathbb{1}\left[\left(\left(1+\epsilon\left(m\right)\right)^{n-3}-1\right)\left(1+\epsilon\left(m\right)\right)-1\right]\\ &\approx \mathbb{1}\left[\left(1+\epsilon\left(m\right)\right)^{n-2}-1\right]\end{aligned}$$

2. If $\widetilde{\rho_{1:n}} = A\rho_{n-1:n}B\rho_{n-1}^{-1}C\rho_{n-2:n-1}D$, then, similarly:

$$\rho_{1:n}^{-1}\widetilde{\rho_{1:n}} = \rho_{1:n}^{-1}\rho_{n-1:n}\rho_{n-1}^{-1}\rho_{1:n-1}\rho_{1:n-1}^{-1}ABC\rho_{n-2:n-1}D$$

Note that, since every element of S_n is hermitian, $\rho_{1:n-1}^{-1}ABC\rho_{n-2:n-1}D = \left(D\rho_{n-2:n-1}CBA\rho_{1:n-1}^{-1}\right)^{\dagger}$, so:

$$\begin{aligned}\rho_{1:n-1}^{-1}ABC\rho_{n-2:n-1}D &\approx \mathbb{1}\left[\left(1+\epsilon\left(m\right)\right)^{n-3}-1\right] \text{ by } P_{n-1} \text{ and Prop. 3}\\ \rho_{1:n}^{-1}\rho_{n-1:n}\rho_{n-1}^{-1}\rho_{1:n-1} &\approx \mathbb{1}\left[\epsilon\left(m\right)\right] \text{ by a reversed version of lemma 4 and Prop. 3}\end{aligned}$$

And:

$$\widetilde{\rho_{1:n}}\rho_{1:n}^{-1} \approx \mathbb{1}\left[\left(1+\epsilon\left(m\right)\right)^{n-2}-1\right]$$

□

Proposition 5. $|AO_n| \leq 2^{n-2}$

Proof. The idea is that, for each $i \in [\![2, n-1]\!]$ (that is $n-2$ possibilities), we have to chose whether $\widetilde{\rho}$ can be written either as in eq.(5) or in eq.(6): that is 2 possibilities. As the factors could commute even if they live in the same Hilbert space, this results in *at most* 2^{n-2} possibilities. Formally for $n \geq 3$, the application:

$$\begin{aligned}f : AO_{n-1} \times \{0,1\} &\to AO_n\\ \left(\widetilde{\rho_{1:n-1}}, b\right) &\mapsto \begin{cases} A\rho_{n-2:n-1}\rho_{n-1}^{-1}\rho_{n-1:n}B & \text{if } b=0\\ A\rho_{n-1:n}\rho_{n-1}^{-1}\rho_{n-2:n-1}B & \text{if } b=1\end{cases}\end{aligned}$$

where A, B are s.t. $\widetilde{\rho_{1:n-1}} = A\rho_{n-2:n-1}B$ is a surjection:

1. **Function definition:** $\widetilde{\rho_{1:n-1}}$ can be decomposed into several equal expressions, so we need to make sure that $f\left(\widetilde{\rho_{1:n-1}}, \bullet\right)$ is well defined. Let us suppose that $\widetilde{\rho_{1:n-1}}$ can be decomposed in two different ways:

$$\widetilde{\rho_{1:n-1}} = A\rho_{n-2:n-1}B = A'\rho_{n-2:n-1}B'$$

Then:

$$\begin{aligned}A\rho_{n-2:n-1}\rho_{n-1}^{-1}\rho_{n-1:n}B &= A\rho_{n-2:n-1}B\rho_{n-1}^{-1}\rho_{n-1:n} \text{ because } B \in \mathcal{H}_{1:n-2}\\ &= A'\rho_{n-2:n-1}B'\rho_{n-1}^{-1}\rho_{n-1:n}\\ &= A'\rho_{n-2:n-1}\rho_{n-1}^{-1}\rho_{n-1:n}B'\end{aligned}$$

So $f\left(\widetilde{\rho_{1:n-1}}, 0\right)$ is well defined. The same goes for $f\left(\widetilde{\rho_{1:n-1}}, 1\right)$.

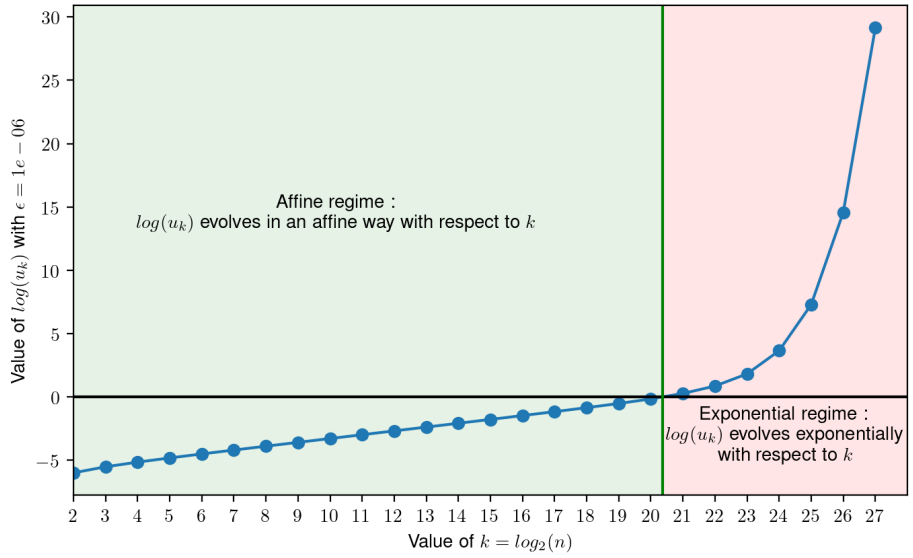

Figure 14: Evolution of the divide-and-conquer bound in terms of the number of particle chunks n

2. **Surjection:** let $\widetilde{\rho_{1:n}} \in AO_n$. By definition, there are two cases:

   (a) $\widetilde{\rho_{1:n}} = A\rho_{n-2:n-1}B\rho_{n-1}^{-1}C\rho_{n-1:n}D = A\rho_{n-2:n-1}\rho_{n-1}^{-1}\rho_{n-1:n}BCD = f\left(A\rho_{n-2:n-1}BCD, 0\right)$

   (b) $\widetilde{\rho_{1:n}} = A\rho_{n-1:n}B\rho_{n-1}^{-1}C\rho_{n-2:n-1}D = ABC\rho_{n-1:n}\rho_{n-1}^{-1}\rho_{n-2:n-1}D = f\left(ABC\rho_{n-2:n-1}D, 1\right)$

As a result, $|AO_n| \leq |AO_{n-1} \times \{0,1\}| = 2\,|AO_{n-1}|$. Moreover, $|AO_2| = |\{\rho_{1:2}\}| = 1$, hence the result. □

As a conclusion, we proved that the bound $(1+\epsilon\,(m))^{n-2} - 1$ holds for up to 2^{n-2} ways of approximating ρ. Numerical estimations on Matlab for some examples also suggest that the bound does not hold for $\widetilde{\rho} \in A_n \backslash AO_n$. For example, with CNOT nearest-neighbor interactions on a chain of 3 blocks of 3 particles each, we get :

$$\left\|\rho_{AB}\rho_B^{-1}\rho_{BC} - \rho_{ABC}\right\|_1 = 4.61e-3$$
$$\left\|\rho_{AB}\rho_{BC}\rho_B^{-1} - \rho_{ABC}\right\|_1 = 1.04$$

Part IV
Numerical estimations

1 Bounds plots

Part V
Personal experiences

My 3 months long internship in Tübingen was, from both scientific and personal points of view, a great experience. The fact that the lab was primarly oriented towards mathematics and physics made the internship interesting but also challenging. I spent the first few weeks becoming familiar with the subject and studying linear algebra, as well as installing and learning how to use the MATLAB programming language. Maple was also among the tools I used, particularly to conjecture recurrence relationships. Researching in exponentially-sized systems proved quite challenging and I spent a non negligible time on errors related to the difficulty of imagining the situation. Apart from that, the only difficulty I had was when I got sick for one full week at the end of the internship.

I also attended the Beyond IID in Quantum Information Theory at the start of August: one week of conferences hosted in Tübingen. Most of the talks were too hard for me to understand, but I appreciated the experience, some talks about quantum verification in computer science and quantum cryptography, as well as a public talk of Ignacio Cirac on the current state of the art of real-life quantum computing.

Part VI
Conclusion

We proposed a general quantum learning algorithm that can be implemented in up to 2^{n-2} ways. The bound we get is still exponential in n, but it is temperature independent. We were originally hoping to improve on the $(1+\epsilon(m))^{n-2}-1$ original bound of the linear-shaped approach. That failed but we still got an interesting equivalence of approaches when transposing the terms of the product. Because the number of particles simulable by a classical computer can not really exceed 18 even on a very powerful computer with careful optimisation, we couldn't make experiments with significantly different values of n to conjecture if the bound was tight for our approaches or not. That question is therefore open. I think subjectively that it has good chances of being tight in order of magnitude (that is, exponentially increasing with n and superexponentially decaying with m).

Another conjecture we made is the fact that, in section 3, the condition $\widetilde{\rho} \in AO_n$ is necessary.

In addition, it would be interesting and might prove useful to explore the MPO approximations of Gibbs states using marginals over more than two subsystems. For example, with terms consisting of marginals over two subsystems at the beginning of the particle chain and two other at another place.

More importantly, the biggest problem our approaches have is that they require exact marginals. A perhaps more realistic version of our problem would be: given $\widetilde{\rho_{1:2}}, \widetilde{\rho_{2:3}}, \widetilde{\rho_{3:4}}, \dots$ $\widetilde{\rho}_1, \widetilde{\rho}_2, \widetilde{\rho}_3, \dots$ s.t.

$$\forall i, \left\| \widetilde{\rho_{i:i+1}} - \rho_{i:i+1} \right\|_1 \leq \epsilon_{i:i+1}$$
$$\forall i, \left\| \widetilde{\rho}_i - \rho_i \right\|_1 \leq \epsilon_i$$

can we bound efficiently $\left\| \widetilde{\rho_{1:2}} \widetilde{\rho}_2^{-1} \widetilde{\rho_{2:3}} \widetilde{\rho}_3^{-1} \dots \widetilde{\rho_{n-1:n}} - \rho \right\|_1$? We don't have an answer to that question either.

Part VII
Bibliography

References

[1] Mohammad H. Amin, Evgeny Andriyash, Jason Rolfe, Bohdan Kulchytskyy, and Roger Melko. Quantum boltzmann machine. *Physical Review X*, 8(2), may 2018.

[2] Huzihiro Araki. Gibbs states of a one dimensional quantum lattice. *Communications in Mathematical Physics*, 14(2):120–157, Jun 1969.

[3] Andreas Bluhm, Ángela Capel, and Antonio Pérez-Hernández. Exponential decay of mutual information for gibbs states of local hamiltonians. *Quantum*, 6:650, feb 2022.

[4] Jean-Christophe Bourin and Eun-Young Lee. On the russo-dye theorem for positive linear maps, 2019.

[5] Hamza Fawzi, Omar Fawzi, and Samuel O. Scalet. A subpolynomial-time algorithm for the free energy of one-dimensional quantum systems in the thermodynamic limit. *Quantum*, 7:1011, may 2023.

[6] M. B. Hastings. Solving gapped hamiltonians locally. *Physical Review B*, 73(8), feb 2006.

[7] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert. Locality of temperature. *Physical Review X*, 4(3), jul 2014.

[8] Tomotaka Kuwahara, Á lvaro M. Alhambra, and Anurag Anshu. Improved thermal area law and quasilinear time algorithm for quantum gibbs states. *Physical Review X*, 11(1), mar 2021.

[9] Andras Molnar, Norbert Schuch, Frank Verstraete, and J. Ignacio Cirac. Approximating gibbs states of local hamiltonians efficiently with projected entangled pair states. *Physical Review B*, 91(4), jan 2015.

[10] Tongyang Li Xiaodi Wu. Lecture notes for quantum semidefinite programming solvers, November 2018.

[11] Hasan Yetis and Mehmet Karakoes. Investigation of noise effects for different quantum computing architectures in ibm-q at nisq level. In *2021 25th International Conference on Information Technology (IT)*, pages 1–4, 2021.

[12] Xiao Yuan. Lecture 11 trace distance and fidelity, November 2021.

[13] Álvaro M. Alhambra. Quantum many-body systems in thermal equilibrium, 2022.

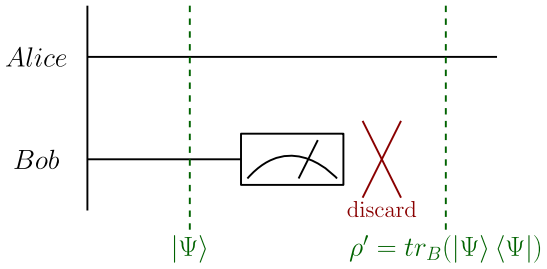

Figure 15: Partial trace over B

Part VIII
Appendix

1 Linear algebra facts

This part contains properties and linear algebra recalls that are necessary to make this report as precise and self-contained as possible. It can however be skipped.

Definition 6. Let $\mathcal{H}_{AB}$ be a bipartite system and $\rho = \rho_{AB}$. The *partial trace over the B subsystem*, denoted tr_B, is the linear map s.t. for every simple tensor $u \otimes v \in \mathcal{H}_{AB}$:

$$\mathrm{tr}_B(u \otimes v) = \mathrm{Tr}(v)\, u \in \mathcal{H}_A$$

Similarly, $\mathrm{tr}_A(u \otimes v) = \mathrm{Tr}(u)\, v \in \mathcal{H}_B$.

In addition, if A and B are separate (and not necessarily adjacent) subintervals of I, we denote by tr_{AB} the map $\mathrm{tr}_A \circ \mathrm{tr}_B$. Note that $\mathrm{tr}_A \circ \mathrm{tr}_B = \mathrm{tr}_B \circ \mathrm{tr}_A$. That way, we can take the partial trace of ρ over any subset $J = \{j_1, \ldots, j_k\} \subseteq I$ by defining:

$$\mathrm{tr}_J \coloneqq \mathrm{tr}_{\{j_1\}} \circ \cdots \circ \mathrm{tr}_{\{j_k\}}$$

Definition 7. Given a mixed quantum state represented by its density matrix ρ, its partial trace over B can be expressed as:

$$\mathrm{tr}_B(\rho) = \sum_{|i\rangle \text{ orthonormal basis of } \mathcal{H}_B} (I_A \otimes \langle i|)\, \rho\, (I_A \otimes |i\rangle)$$

Remark 3. Taking the partial trace of ρ_{AB} over B is the same as measuring the state in B and discarding the measure.

Proof. First, let us specify what we mean by that. Let $|\Psi\rangle$ be a pure bipartite state shared between Alice and Bob. let us suppose that Bob measures its share and discards the result (see Figure 15).

let us write $|\Psi\rangle = \sum_{i,j} \Psi_{i,j} |i\rangle \otimes |j\rangle \in \mathcal{H}_{AB}$.

For each j, Bob has a chance

$$p_j \coloneqq \sum_i |\Psi_{i,j}|^2 = \langle\Psi| (I_A \otimes |j\rangle\langle j|)^\dagger (I_A \otimes |j\rangle\langle j|) |\Psi\rangle$$

of measuring $|j\rangle$. In this case, the post-measurement state is

$$\left|\phi^{(j)}\right\rangle = \frac{1}{\sqrt{p_j}}\sum_i \Psi_{i,j}\,|i\rangle \otimes |j\rangle$$

Thus, the post-measurement state is a mixed state $\left\{\left(p_j, \left|\phi^{(j)}\right\rangle\right)\right\}_j$: it can be represented by the density matrix:

$$\begin{aligned}\rho' &\coloneqq \sum_j p_j \left|\phi^{(j)}\right\rangle \left\langle\phi^{(j)}\right| \\ &= \sum_{j,k,l} \Psi_{k,j}\Psi_{l,j}\left(|k\rangle\langle l|\right) \otimes \left(|j\rangle\langle j|\right)\end{aligned}$$

Discarding the component in $\mathcal{H}_B$ yields

$$\rho' = \sum_{j,k,l} \Psi_{k,j}\Psi_{l,j}\,|k\rangle\langle l|$$

On the other hand,

$$\begin{aligned}tr_B\left(|\Psi\rangle\langle\Psi|\right) &= \sum_j \left(I_A \otimes \langle j|\right)|\Psi\rangle\langle\Psi|\left(I_A \otimes |j\rangle\right) \\ &= \sum_{j,k,l,m,n} \Psi_{k,m}\Psi_{l,n}\left(I_A \otimes \langle j|\right)\left(|k\rangle \otimes |m\rangle\right)\left(\langle l| \otimes \langle n|\right)\left(I_A \otimes |j\rangle\right) \\ &= \sum_{j,k,l} \Psi_{k,j}\Psi_{l,j}\left(I_A \otimes \langle j|\right)\left(|k\rangle \otimes |j\rangle\right)\left(\langle l| \otimes \langle j|\right)\left(I_A \otimes |j\rangle\right) \\ &= \sum_{j,k,l} \Psi_{k,j}\Psi_{l,j}\,|k\rangle\langle l| \\ &= \rho'\end{aligned}$$

□

Note: this proof also works when $|\Psi\rangle$ is a mixed state.

Lemma 8. *The Gibbs state $\rho \coloneqq \frac{e^{-H}}{\mathrm{Tr}(e^{-H})}$ is a $2^{|I|} \times 2^{|I|}$ density matrix. Moreover, like all density matrices:*

$$\|\rho\|_1 = \mathrm{tr}\left(\rho\right) = 1$$

Proof. H being hermitian ($H^\dagger = H$), it is diagonalizable: there exists a unitary matrix U and a diagonal matrix $D = diag\left(d_1, \ldots, d_{2^N}\right)$ s.t. $H = UDU^\dagger$. Writing D in the orthonormal basis $\left(|i\rangle\right)_i$ of $\mathcal{H}_I$ gives $H = \sum_i d_i\,|i\rangle\langle i|$. As these terms commute:

$$\begin{aligned}e^{-H} &= \frac{1}{\mathrm{Tr}\left(e^{-H}\right)} U e^{-D} U^\dagger \\ &= \frac{1}{\mathrm{Tr}\left(e^{-H}\right)} U \left(\sum_i e^{-d_i}\,|i\rangle\langle i|\right) U^\dagger \\ &= \sum_i \underbrace{\frac{e^{-d_i}}{\mathrm{Tr}\left(e^{-H}\right)}}_{=:p_i} \left(\underbrace{U\,|i\rangle}_{\text{pure quantum state}}\right)\left(U\,|i\rangle\right)^\dagger\end{aligned}$$

With $\sum_i p_i = \sum_i \frac{e^{-d_i}}{\mathrm{Tr}(e^{-H})} = \frac{1}{\sum_i e^{-d_i}} \sum_i e^{-d_i} = 1$, which concludes our proof of ρ being a density matrix.

For the trace, let us rewrite $\rho = \sum_j p_j |\Psi_j\rangle \langle\Psi_j|$ where $|\Psi_j\rangle$ is a pure quantum state and $\sum_j p_j = 1$:

$$\begin{aligned}
\mathrm{tr}(\rho) &= \sum_i p_i \,\mathrm{tr}\left(|\Psi_j\rangle \langle\Psi_j|\right) \text{ by linearity} \\
&= \sum_i p_i \,\mathrm{tr}\left(\langle\Psi_j|\Psi_j\rangle\right) \\
&= \sum_i p_i = 1
\end{aligned}$$

For the 1-norm, $\|\rho\|_1 \underset{\text{def}}{=} \sum_i |\lambda_i|$ where $\{\lambda_i\}_i$ are the singular values (here, eigenvalues) of ρ. As the eigenvalues of a quantum state are real and positive, $\|\rho\|_1 = \mathrm{tr}(\rho) = 1$. □

2 Proofs

Proof. of Lemma 1: let us write I as the union of disjoint intervals $I = \overbrace{I_1 A \underbrace{I_2}_{J_2}}^{J_1} B I_3$. Then:

$$\begin{aligned}
\rho_A \rho_B = &\left(\frac{1}{Tr\left(\mathbb{1}_{I_1}\right) Tr\left(\mathbb{1}_{J_2}\right)} \mathbb{1}_{I_1} \otimes tr_{I_1 J_2}(\rho) \otimes \mathbb{1}_{J_2}\right) \cdot \\
&\left(\frac{1}{Tr\left(\mathbb{1}_{J_1}\right) Tr\left(\mathbb{1}_{I_3}\right)} \mathbb{1}_{J_1} \otimes tr_{J_1 I_3}(\rho) \otimes \mathbb{1}_{I_3}\right) \\
&= \frac{1}{Z} \left(\mathbb{1}_{I_1} \otimes tr_{I_1 J_2}(\rho) \otimes \mathbb{1}_{I_2} \otimes tr_{J_1 I_3}(\rho) \otimes \mathbb{1}_{I_3}\right) \\
&\text{where } Z \coloneqq Tr\left(\mathbb{1}_{I_1}\right) Tr\left(\mathbb{1}_{J_2}\right) Tr\left(\mathbb{1}_{J_1}\right) Tr\left(\mathbb{1}_{I_3}\right) \\
&= \left(\frac{1}{Tr\left(\mathbb{1}_{J_1}\right) Tr\left(\mathbb{1}_{I_3}\right)} \mathbb{1}_{J_1} \otimes tr_{J_1 I_3}(\rho) \otimes \mathbb{1}_{I_3}\right) \cdot \\
&\left(\frac{1}{Tr\left(\mathbb{1}_{I_1}\right) Tr\left(\mathbb{1}_{J_2}\right)} \mathbb{1}_{I_1} \otimes tr_{I_1 J_2}(\rho) \otimes \mathbb{1}_{J_2}\right) \\
&= \rho_B \rho_A
\end{aligned}$$

□

Proof. Sketch of proof of Lemma 2: the map

$$M \mapsto \mathrm{tr}_A\left(e^{-H_A} M\right)$$

is positive, so the Russo-Dye theorem (see [4]) applies, hence the conclusion □

Proof. of Lemma 3: for simplicity, suppose that $\mathcal{H}_I = \mathcal{H}_A \otimes \mathcal{H}_B$. let us decompose Z_{AB} as $Z_{AB} = \sum_{i,j} \lambda_{i,j} \phi_i \otimes \psi_j$ where $(\phi_i \otimes \psi_j)_{i,j}$ is a basis of $\mathcal{H}_I$. Then:

$$\begin{aligned} tr_B \left(Z_{AB} \left(Y_A \otimes \mathbb{1}_B \right) \right) &= \sum_{i,j} \lambda_{i,j} tr_B \left(\phi_i \otimes \psi_j \left(Y_A \otimes \mathbb{1}_B \right) \right) \\ &= \sum_{i,j} \lambda_{i,j} tr_B \left(\left(\phi_i Y_A \right) \otimes \psi_j \right) \\ &= \sum_{i,j} \lambda_{i,j} Tr \left(\psi_j \right) \phi_i Y_A \\ &= \left(\sum_{i,j} \lambda_{i,j} Tr \left(\psi_j \right) \phi_i \right) Y_A \\ &= tr_B \left(Z_{AB} \right) Y_A \end{aligned}$$

The proof follows exactly the same steps in the general case where $\mathcal{H}_I = \mathcal{H}_{I_1} \otimes \mathcal{H}_A \otimes \mathcal{H}_{I_2} \otimes \mathcal{H}_B \otimes \mathcal{H}_{I_3}$. □

Proof. of lemma 4. We will in fact bound $\left\| \mathbb{1} - \rho_{ABC} \rho_{BC}^{-1} \rho_B \rho_{AB}^{-1} \right\|$ and conclude with

$$\begin{aligned} \left\| \mathbb{1} - Q^{-1} \right\| &= \left\| \sum_{k \geq 1} \left(\mathbb{1} - Q \right)^k \right\| \\ &\leq \left\| \mathbb{1} - Q \right\| \sum_{k \geq 0} \left\| \mathbb{1} - Q \right\|^k \\ &\leq 2 \left\| \mathbb{1} - Q \right\| \end{aligned}$$

as long as $\left\| \mathbb{1} - Q \right\| \leq \frac{1}{2}$ which will be the case for m large enough. For small values of m, $\left\| \mathbb{1} - \rho_{ABC} \rho_{BC}^{-1} \rho_B \rho_{AB}^{-1} \right\|$ will be bounded by a constant.

First,

$$\begin{aligned} \rho_{ABC} \rho_{BC}^{-1} &= e^{-H_{ABC}} \operatorname{tr}_A \left(e^{-H_{ABC}} \right)^{-1} \\ &= e^{-H_{ABC}} e^{H_A + H_{BC}} e^{-H_A} \operatorname{tr}_A \left(e^{-H_{ABC}} e^{H_A + H_{BC}} e^{-H_A} \right)^{-1} \text{ because of lemma 3} \\ &= E_{A,BC} \rho^A \operatorname{tr} \left(E_{A,BC} \rho^A \right)^{-1} \text{ with the notation introduced in 2} \end{aligned}$$

Similarly,

$$\rho_B \rho_{AB}^{-1} = \operatorname{tr}_A \left(\widetilde{E_{A,BC}} \rho^A \right) \left(\rho_A \right)^{-1} \widetilde{E_{A,BC}}^{-1}$$

where $\widetilde{E_{A,BC}} \coloneqq \operatorname{tr}_C \left(e^{-H_{ABC}} \right) \operatorname{tr}_C \left(e^{-H_{BC}} \right)^{-1} e^{H_A}$ *looks like* $E_{A,BC}$

As a consequence

$$\rho_{ABC} \rho_{BC}^{-1} \rho_B \rho_{AB}^{-1} = E_{A,BC} \operatorname{tr} \left(E_{A,BC} \rho^A \right)^{-1} \operatorname{tr}_A \left(\widetilde{E_{A,BC}} \rho^A \right) \widetilde{E_{A,BC}}^{-1}$$

Claim 2. These terms are uniformly bounded by a constant $\hat{C} > 0$ depending on Γ, r but independent of A, B, C.

Claim 3. There exists a superexponentially decaying function $\delta(l)$ depending on Γ, r but independent of A, B, C s.t.

$$E_{A,BC} \approx \widetilde{E_{A,BC}}\left[\delta\left(|B|\right)\right]$$

Claim 3 in conjunction with 2 allows us to deduct

$$\mathrm{tr}_A\left(E_{A,BC}\rho^A\right) \approx \mathrm{tr}_A\left(\widetilde{E_{A,BC}}\rho^A\right)\left[\delta\left(l\right)\right]$$

Then :

$$\begin{aligned}
\left\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_A^{-1} - \mathbb{1}\right\| &= \left\|E_{A,BC}\rho^A \,\mathrm{tr}\left(E_{A,BC}\rho^A\right)\mathrm{tr}_A\left(\widetilde{E_{A,BC}}\rho^A\right)\widetilde{E_{A,BC}}^{-1} - \mathbb{1}\right\| \\
&= \left\|E_{A,BC}\rho^A \,\mathrm{tr}\left(E_{A,BC}\rho^A\right)\mathrm{tr}_A\left(\widetilde{E_{A,BC}}\rho^A\right)\left(E_{A,BC} - E_{A,BC} + \widetilde{E_{A,BC}}^{-1}\right) - \mathbb{1}\right\| \\
&= \hat{C}^3\left\|\widetilde{E_{A,BC}}^{-1} - E_{A,BC}\right\| + \left\|E_{A,BC}\rho^A \,\mathrm{tr}\left(E_{A,BC}\rho^A\right)\mathrm{tr}_A\left(\widetilde{E_{A,BC}}\rho^A\right)E_{A,BC} - \mathbb{1}\right\| \\
&= \hat{C}^3\delta\left(|B|\right) + \hat{C}\delta\left(|B|\right)
\end{aligned}$$

Proof. of Claim 2 : Araki's results in conjunction with lemma 2 give us directly

$$\left\|E_{A,BC}\right\|, \left\|\mathrm{tr}_C\left(E_{AB,C}\rho^C\right)\right\|, \left\|\mathrm{tr}_C\left(E_{B,C}\rho^C\right)^{-1}\right\| \leq \hat{C}$$

Then,

$$\begin{aligned}
\widetilde{E_{A,BC}} &= \mathrm{tr}_C\left(e^{-H_{ABC}}e^{H_{AB}}\right)e^{-H_{AB}}e^{H_A+H_B}\,\mathrm{tr}_C\left(e^{-H_{BC}}e^{H_B}\right)^{-1} \\
&= \mathrm{tr}_C\left(E_{AB,C}\rho^C\right)E_{A,B}\,\mathrm{tr}_C\left(E_{B,C}\rho^C\right)^{-1} \\
\left\|\widetilde{E_{A,BC}}\right\| &\leq \hat{C}^3
\end{aligned}$$

□

Proof. of claim 3 : let $l \coloneqq \left\lfloor \frac{|B|}{2} \right\rfloor$ and let us split B into two parts $B = B_l B_r$ with $|B_l|, |B_r| \geq l$. Then :

$$\begin{aligned}
\left\|\widetilde{E_{A,BC}} - E_{A,BC}\right\| &= \left(\mathrm{tr}_C\left(E_{AB,C}\rho^C\right)E_{A,B} - E_{A,BC}\,\mathrm{tr}_C\left(E_{B,C}\rho^C\right)\right)\left(\mathrm{tr}_C\left(E_{B,C}\rho^C\right)\right)^{-1} \\
&\leq C\left\|\mathrm{tr}_C\left(E_{AB,C}\rho^C\right)E_{A,B} - E_{A,BC}\,\mathrm{tr}_C\left(E_{B,C}\rho^C\right)\right\| \\
&\leq C\left\|\mathrm{tr}_C\left(E_{AB,C}\rho^C\right)E_{A,B} - \mathrm{tr}_C\left(E_{B_r,C}\rho^C\right)E_{A,B_l}\right\| \\
&+ C\left\|E_{A,B_l}\,\mathrm{tr}_C\left(E_{B_r,C}\rho^C\right) - E_{A,BC}\,\mathrm{tr}_C\left(E_{B,C}\rho^C\right)\right\| \\
&\leq C\left\|\mathrm{tr}_C\left(E_{AB,C}\rho^C\right)\right\|\left\|E_{A,B} - E_{A,B_l}\right\| \\
&+ C\left\|\mathrm{tr}_C\left(E_{AB,C}\rho^C\right) - \mathrm{tr}_C\left(E_{B_r,C}\rho^C\right)\right\|\left\|E_{A,B_l}\right\| \\
&+ C\left\|E_{A,B_l}\right\|\left\|\mathrm{tr}_C\left(E_{B_r,C}\rho^C\right) - \mathrm{tr}_C\left(E_{B,C}\rho^C\right)\right\| \\
&+ C\left\|E_{A,B_l} - E_{A,BC}\right\|\left\|\mathrm{tr}_C\left(E_{B,C}\rho^C\right)\right\| \\
&\leq C\left\|E_{AB,C}\right\|\left\|E_{A,B} - E_{A,B_l}\right\| + C\left\|E_{AB,C} - E_{B_r,C}\right\|\left\|E_{A,B_l}\right\| \\
&+ C\left\|E_{A,B_l}\right\|\left\|E_{B_r,C} - E_{B,C}\right\| + C\left\|E_{A,B_l} - E_{A,BC}\right\|\left\|E_{B,C}\right\| \\
&\leq \frac{4C\Delta^{1+l}}{\left(\left\lfloor \frac{l}{r} \right\rfloor + 1\right)!}
\end{aligned}$$

□

□

Proof. of lemma 5

$$\begin{aligned}
\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} &= \mathrm{tr}_D\left(\rho_{AB}\rho_B^{-1}\rho_{BCD}\right)\mathrm{tr}_D\left(\rho_{ABCD}\right)^{-1} \\
&= \mathrm{tr}_D\left(\rho_{AB}\rho_B^{-1}\rho_{BCD}\rho_{ABCD}^{-1}\rho_{ABCD}\right)\mathrm{tr}_D\left(\rho_{ABCD}\right)^{-1} \\
&= \mathrm{tr}_D\left(\rho_{AB}\rho_B^{-1}\rho_{BCD}\rho_{ABCD}^{-1}e^{-H_{ABCD}}e^{H_{ABC}}\right)\mathrm{tr}_D\left(e^{-H_{ABCD}}e^{H_{ABC}}\right)^{-1} \\
&= \mathrm{tr}_D\left(\rho_{AB}\rho_B^{-1}\rho_{BCD}\rho_{ABCD}^{-1}E_{ABC,D}e^{-H_D}\right)\mathrm{tr}_D\left(E_{ABC,D}e^{-H_D}\right)^{-1} \\
&= \mathrm{tr}_D\left(\rho^D\rho_{AB}\rho_B^{-1}\rho_{BCD}\rho_{ABCD}^{-1}E_{ABC,D}\right)\mathrm{tr}_D\left(\rho^D E_{ABC,D}\right)^{-1}
\end{aligned}$$

Thus

$$\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} - \mathbb{1} = \mathrm{tr}_D\left(\rho^D\left(\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} - \mathbb{1}\right)E_{ABC,D}\right)\mathrm{tr}_D\left(E_{ABC,D}\right)^{-1}$$

Araki's results allow us to bound the second term by a constant Δ, and the contractiveness lemma 2 gives us $\Delta\epsilon\left(m\right)$ as a bound for the first one. All in all

$$\rho_{AB}\rho_B^{-1}\rho_{BC}\rho_{ABC}^{-1} \approx \mathbb{1}\left[\underbrace{\Delta^2\epsilon\left(m\right)}_{\epsilon'(m)}\right]$$

□

Proof. of lemma 11: we will upper bound $\left\|\rho_{ABCD}\rho_{BC}^{-1}\rho_C\rho_{CD}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\right\|$. Let $E_{I,J} \coloneqq e^{-H_{IJ}}e^{H_I}e^{H_J}$. Notice that $E_{I,J} = E_{J,I}$. We can rewrite:

$$\begin{aligned}
\rho_B\rho_{AB}^{-1} &= \mathrm{tr}_{ACD}\left(e^{-H_{ABCD}}\right)\mathrm{tr}_{AB}\left(e^{-H_{ABCD}}\right)^{-1} \\
&= \mathrm{tr}_A\left(\mathrm{tr}_{CD}\left(e^{-H_{ABCD}}\right)\right)\mathrm{tr}_{AB}\left(e^{-H_{ABCD}}\right)^{-1} \\
&= \mathrm{tr}_A\left(\mathrm{tr}_{CD}\left(e^{-H_{ABCD}}\right)\mathrm{tr}_{CD}\left(e^{-H_{BCD}}\right)^{-1}e^{H_A}\rho^A\right)\left(\rho^A\right)^{-1}e^{-H_A}\,\mathrm{tr}_{CD}\left(e^{-H_{BCD}}\right)\mathrm{tr}_{CD}\left(e^{-H_{ABCD}}\right)^{-1} \\
&= \mathrm{tr}_A\left(\widetilde{E_{A,BCD}}\rho^A\right)\left(\rho^A\right)^{-1}\widetilde{E_{A,BCD}}^{-1} \text{ where } \widetilde{E_{A,BCD}} \coloneqq \mathrm{tr}_{CD}\left(e^{-H_{ABCD}}\right)\mathrm{tr}_{CD}\left(e^{-H_{BCD}}\right)^{-1}e^{H_A}
\end{aligned}$$

The idea is that $\widetilde{E_{A,BCD}}$ *"looks like"* $E_{A,BCD}$. Similarly,

$$\rho_C\rho_{CD}^{-1} = \mathrm{tr}_D\left(\widetilde{E_{D,ABC}}\rho^D\right)\left(\rho^D\right)^{-1}\widetilde{E_{D,ABC}}^{-1} \text{ where } \widetilde{E_{D,ABC}} \coloneqq \mathrm{tr}_{AB}\left(e^{-H_{ABCD}}\right)\mathrm{tr}_{AB}\left(e^{-H_{ABC}}\right)^{-1}e^{H_D}$$

Moreover,

$$\begin{aligned}\rho_{ABCD}\rho_{BC}^{-1} &= e^{-H_{ABCD}}\operatorname{tr}_{AD}\left(e^{-H_{ABCD}}\right)^{-1}\\ &= e^{-H_{ABCD}}e^{H_A+H_{BCD}}e^{-H_{BCD}}e^{H_{BC}+H_D}\rho^A\rho^D\\ &\operatorname{tr}_{AD}\left(e^{-H_{ABCD}}e^{H_A}e^{H_{BC}}e^{H_D}\rho^A\rho^D\right)^{-1}\\ &= E_{A,BCD}E_{BC,D}\rho^A\rho^D\operatorname{tr}_A\left(\operatorname{tr}_D\left(E_{A,BCD}E_{BC,D}\rho^D\right)\rho^A\right)^{-1}\end{aligned}$$

Finally:

$$\rho_{ABCD}\rho_{BC}^{-1}\rho_C\rho_{CD}^{-1}\rho_B\rho_{AB}^{-1} = \underbrace{E_{A,BCD}E_{BC,D}}_{X}\underbrace{\operatorname{tr}_A\left(\operatorname{tr}_D\left(E_{A,BCD}E_{BC,D}\rho^D\right)\rho^A\right)^{-1}}_{Y^{-1}}\cdot$$
$$\underbrace{\operatorname{tr}_D\left(\widetilde{E_{D,ABC}}\rho^D\right)\widetilde{E_{D,ABC}}^{-1}}_{\widetilde{Y}}\underbrace{\operatorname{tr}_A\left(\widetilde{E_{A,BCD}}\rho^A\right)\widetilde{E_{A,BCD}}^{-1}}_{\widetilde{X}^{-1}}$$

In a completely analogous way to claims used in the proof of 4, we conclude by bounding the terms and proving that $\left\|X-\widetilde{X}\right\|, \left\|Y-\widetilde{Y}\right\|, \left\|X^{-1}-\widetilde{X}^{-1}\right\|, \left\|Y^{-1}-\widetilde{Y}^{-1}\right\|$ are superexponentially decaying. □