Limit Theorems for Quantum Entropies

vorgelegt von

Diplom-Physiker

Igor Bjelaković

aus Sarajevo

von der Fakultät II – Mathematik- und Naturwissenschaften

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

-Dr. rer. nat.-

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. J.-D. Deuschel
Berichter:	Prof. Dr. R. Seiler
Berichter:	Dr. habil. T. Krüger

Tag der wissenschaftlichen Aussprache: 29. Januar 2004

Berlin 2004

D 83

Danksagung

Mein tiefster Dank gilt Tyll Krüger, Ruedi Seiler, Rainer Siegmund- Schultze und Arleta Szkoła, von und mit denen ich die Wahrscheinlichkets- und Informationstheorie gelernt habe. Der kollegiale und freundschaftliche Umgang innerhalb unserer Gruppe half mir auch, die Einarbeitungszeit durchzustehen, in der ”kein Land in Sicht” zu sein schien. Ohne die fast täglichen Diskussionen über die Thematik der Grenzwertsätze für Entropien wäre diese Arbeit wohl nie entstanden.
Besonders dankbar bin ich Rainer Siegmund-Schultze, der stets mit Hilfe und Rat zur Verfügung stand, und der mir mit Engelsgeduld das wissenschaftliche Schreiben beibrachte und immer noch beibringt. Er hat unermüdlich die verschiedenen Vorversionen dieser Dissertationsschrift kritisch durchgesehen und manchen ”Patzer” verhindert.
Zum Schluss möchte ich mich bei Prof. Jean-Dominique Deuschel bedanken, der sich schnell und unkompliziert bereit erklärt hat, die Funktion des Vorsitzenden des Prüfungsausschusses zu übernehmen.

Contents

1 Introduction **4**
- 1.1 Entropy 4
- 1.2 Relative Entropy or Kullback-Leibler Distance 6
- 1.3 Data Processing Inequality for the Relative Entropy 8

2 Quantum Version of the Shannon-McMillan Theorem **11**
- 2.1 The Set-up 12
- 2.2 An Ergodic Decomposition 14
- 2.3 Quantum Shannon-McMillan Theorem 19
  - 2.3.1 Proof of the Shannon-McMillan Theorem 20
- 2.4 Concluding Remarks 29

3 Limit Theorems for the Quantum Relative Entropy **31**
- 3.1 Quantum Stein's Lemma and the Relative AEP 31
  - 3.1.1 Proof of the Quantum Stein's Lemma 33
  - 3.1.2 Proof of the Quantum Relative AEP 44
- 3.2 Concluding Remarks 45

4 Monotonicity of the Quantum Relative Entropy **47**
- 4.1 Stein's Lemma: The i.i.d. Case 48
- 4.2 Monotonicity of the Quantum Relative Entropy 52

A The Hiai-Petz Approximation **54**

B Proof of Theorem 3.1.3 **58**

Chapter 1

Introduction

In this introduction we will explain the ideas behind the results of this thesis. We will restrict ourselves to the simplest case here in order to avoid technicalities, i.e. we shall consider only classical i.i.d. stochastic processes with finite alphabet. This is not a serious restriction since the interpretation of the results remains unchanged in the ergodic quantum case, too. At the same time instead of following always the shortest way of explanation we will try to stay as close as possible to the arguments used in the proofs of the main results in the quantum setting.
The classical results of this introduction are "folklore" of the information theory and full proofs can be found in almost every book on the subject.

1.1 Entropy

The entropy of a probability distribution P on a finite set A, introduced by Shannon, is defined as

$$H(P) := -\sum_{a\in A} P(a)\log P(a) = E_P(-\log P),$$

where E_P denotes the expectation with respect to the distribution P. Entropy is one of the fundamental quantities in the information theory because of its intimate connection to coding or data compression results. Following the fundamental idea of information theory that each information theoretical quantity obtains its full justification trough an appropriate limit theorem, we start with law of large numbers for entropy called *asymptotic equipartition property* (AEP) or Shannon's theorem. Consider the measurable space $[A^{\mathbb{Z}}, \mathfrak{A}^{\mathbb{Z}}]$ where $A^{\mathbb{Z}}$ is the set consisting of two-sided infinite sequences with components from A and $\mathfrak{A}^{\mathbb{Z}}$ denotes the σ-algebra generated by the cylinder sets. Let P_∞ be an i.i.d. probability measure on $[A^{\mathbb{Z}}, \mathfrak{A}^{\mathbb{Z}}]$ with generating distribution P. For each $\varepsilon > 0$ and each positive integer n we define the set of *entropy typical* sequences:

$$T_{n,\varepsilon} := \{a^n \in A^n : | - \frac{1}{n} \log P^{(n)}(a^n) - H(P)| < \varepsilon\}.$$

The law of large numbers guarantees that

$$\lim_{n\to\infty} P^{(n)}(T_{n,\varepsilon}) = 1. \tag{1.1}$$

Note that by definition of $T_{n,\varepsilon}$ we have

$$e^{-n(H(P)+\varepsilon)} < P^{(n)}(a^n) < e^{-n(H(P)-\varepsilon)} \quad \text{for all} \quad a^n \in T_{n,\varepsilon}, \tag{1.2}$$

and since bounds on probability imply bounds on cardinality we obtain

$$e^{n(H(P)-\varepsilon)} < \#T_{n,\varepsilon} < e^{n(H(P)+\varepsilon)}, \tag{1.3}$$

for sufficiently large n. The equations (1.1) and (1.3) state that the measure $P^{(n)}$ is essentially carried by the set $T_{n,\varepsilon}$ (see Figure 1.1) the cardinality of which is exponentially smaller than that of A^n (except for the case where P is the equidistribution). Moreover, the meaning of (1.2) is that the measure $P^{(n)}$ is "*equipartitioned*" over the set $T_{n,\varepsilon}$. From the data compression point of view the AEP states that an asymptotically error-free information compression at rate $H(P)$ is possible. On the other hand, (1.2) ensures that an asymptotically error-free compression at a rate below $H(P)$ is impossible since all sequences in $T_{n,\varepsilon}$ have probabilities of order $e^{-nH(P)}$. In fact, let us have a closer look at the data compression. We consider an encoder $f_n : A^n \to B^m$ and a decoder $g_n : B^m \to A^n$, where $B = \{0, 1\}$. The probability of error is given by $P^{(n)}(\{a^n : a^n \neq g_n \circ f_n(a^n)\})$. Our task is to find an encoder and a decoder such that the ratio $\frac{m}{n}$ is as small as possible while the probability of error should not exceed some prescribed $\varepsilon \in (0, 1)$. Observe now that the pairs (f_n, g_n) having error probabilities less or equal ε correspond in one-to-one manner to sets $A \subset A^n$ with $P^{(n)}(A) \geq 1 - \varepsilon$ and $\#A \leq 2^m$. Hence, in order to study asymptotics of such codes we have to consider the limiting behaviour of the following *covering exponents*

$$\alpha_{\varepsilon,n}(P_\infty) := \min\{\log_2 \#A : A \subset A^n, P^{(n)}(A) \geq 1 - \varepsilon\},$$

for $\varepsilon \in (0, 1)$. An easy application (cf. [24]) of the AEP shows that

$$\lim_{n\to\infty} \frac{1}{n} \alpha_{\varepsilon,n}(P_\infty) = H(P) \quad \text{for all} \quad \varepsilon \in (0, 1). \tag{1.4}$$

Setting $m(\varepsilon, n) = [\alpha_{\varepsilon,n}(P_\infty)] + 1$ we obtain

$$\lim_{n\to\infty} \frac{m(\varepsilon, n)}{n} = H(P),$$

i.e. the optimal asymptotic error-free coding scheme has the rate $H(P)$. This consideration makes clear that, looking trough the data compressors eyes, the AEP and (1.4) express the same fact. Indeed, as we indicated above the AEP implies (1.4). The converse implication will be demonstrated in proof of Lemma 2.3.5. We mention this fact because our proof of the Shannon-McMillan theorem in the quantum case will be based to a substantial part on the extension of the covering exponent idea to that more general situation. Another pillar in that proof will be the decomposition of the ergodic quantum state under consideration into ergodic components with respect to a subshift (cf. Theorem 2.2.1). This technique will help us to cope with ergodicity problems arising in an approximation of the quantum state by a appropriately chosen classical stochastic process.

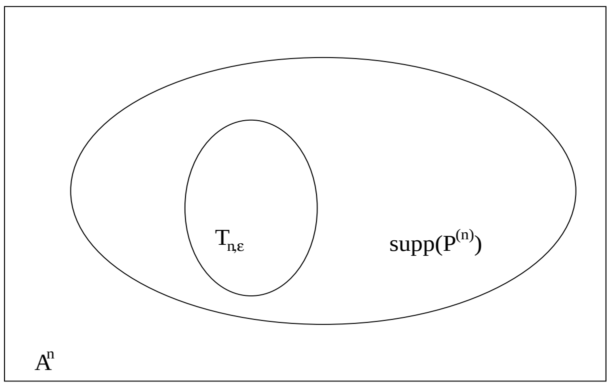

Figure 1.1: Entropy typical set $T_{n,\varepsilon}$

1.2 Relative Entropy or Kullback-Leibler Distance

The aim of this section is to describe, in an informal way, the main idea behind the ergodic theorem for relative entropy. Again, we consider only i.i.d. stochastic processes for simplicity.
When A is a finite set and P and Q are probability distributions on A the relative entropy of P with respect to Q, introduced by Kullback and Leibler, is given by

$$D(P,Q) := \sum_{a \in A} P(a) \log \frac{P(a)}{Q(a)}.$$

Our task in the remainder of this section is to understand intuitively the meaning of the relative entropy from the point of view of limit theorems. To this end we have to consider stochastic processes consisting of infinitely many independent repetitions of experiments governed by the probability distributions P and Q. Let P_∞ and Q_∞ be i.i.d. probability measures on

$[A^{\mathbb{Z}}, \mathfrak{A}^{\mathbb{Z}}]$ with generating distributions P resp. Q, where $A^{\mathbb{Z}}$ is the set of two-sided infinite sequences with components from A and $\mathfrak{A}^{\mathbb{Z}}$ is the σ-algebra generated by the cylinder sets. If we assume that $P \neq Q$ then a simple application of the law of large numbers shows that there is measurable set $S \subset A^{\mathbb{Z}}$ with the property that $P_\infty(S) = 1$ and $Q_\infty(S) = 0$, i. e. the measures P_∞ and Q_∞ are mutually singular. Hence we can see that the $Q^{(n)}$-probability of the set of sequences which are *frequency* typical with respect to $P^{(n)}$ approaches 0 as n becomes large. In other words, if we are given a long sample a^n drawn from the set A and we are asked to decide whether this sample is a realization of P_∞ or Q_∞ then, at least in principle, we can make the right decision with high probability if n is large enough.
Our next goal is to understand the role played by the relative entropy in the problem of discrimination between two i.i.d. probability measures described above. Consider a frequency typical sequence a^n with respect to $P^{(n)}$. Since $Q^{(n)}$ and $P^{(n)}$ are product measures we have

$$\begin{aligned} Q^{(n)}(a^n) &= \prod_{a \in A} Q(a)^{nP_{a^n}(a)} = e^{-n(H(P_{a^n}) + D(P_{a^n}, Q))}, \\ P^{(n)}(a^n) &= e^{-n(H(P_{a^n}) + D(P_{a^n}, P))}, \end{aligned} \tag{1.5}$$

where P_{a^n} denotes the empirical distribution or the type of the sequence a^n. We have $H(P_{a^n}) \simeq H(P)$, $D(P_{a^n}, Q) \simeq D(P, Q)$ and $D(P_{a^n}, P)) \simeq 0$ if n is large enough due to our assumption that a^n is frequency typical with respect to $P^{(n)}$. Hence we are led naturally to the set

$$A_{n,\varepsilon} := B_{n,\varepsilon} \cap T_{n,\varepsilon}(P),$$

for $\varepsilon > 0$, where

$$B_{n,\varepsilon} := \{a^n \in A^n : D(P, Q) - \varepsilon < \frac{1}{n} \log \frac{P^{(n)}(a^n)}{Q^{(n)}(a^n)} < D(P, Q) + \varepsilon\},$$

and $T_{n,\varepsilon}(P)$ denotes the set of entropy typical sequences with respect to P_∞ defined in the previous section. Then one immediately sees that

$$\lim_{n \to \infty} P^{(n)}(A_{n,\varepsilon}) = 1 \quad \text{and} \quad \# A_{n,\varepsilon} \simeq e^{-nH(P)}, \tag{1.6}$$

for n large enough. Moreover for each $a^n \in A_{n,\varepsilon}$ we have

$$Q^{(n)}(a^n) \simeq e^{-n(H(P) + D(P, Q))} \quad \text{and} \quad P^{(n)}(a^n) \simeq e^{-nH(P)}, \tag{1.7}$$

by definition of $A_{n,\varepsilon}$. Puting these pieces toghether we arrive at

$$Q^{(n)}(A_{n,\varepsilon}) \simeq e^{-nD(P,Q)}. \tag{1.8}$$

Clearly, we could derive the equations (1.6), (1.7) and (1.8) directly from (1.5) but in the quantum case the analogues of the sets $A_{n,\varepsilon}$ will serve as a substitute for frequency typical sets (even in the ergodic situation). The equations (1.6), (1.7) and (1.8) express the validity of the *AEP for the relative entropy*. Therefore, it is clear that there is no sequence (B_n) of sets contained in A^n, respectively, whose P-probabilities are close to 1 but having Q-probabilities asymptotically substantially smaller than the sequence $(A_{n,\varepsilon})$. To be precise, let $\varepsilon \in (0, 1)$ and set

$$\beta_{\varepsilon,n}(P_\infty, Q_\infty) := \min\{\log Q^{(n)}(B) : B \subset A^n, P^{(n)}(B) \geq 1 - \varepsilon\}. \tag{1.9}$$

Then the heuristics above can be expressed in the statement that

$$\lim_{n\to\infty} \frac{1}{n} \beta_{\varepsilon,n}(P_\infty, Q_\infty) = -D(P, Q), \tag{1.10}$$

for each pair of i.i.d. measures P_∞, Q_∞ and all $\varepsilon \in (0, 1)$. Sets obeying (1.8) and whose $P^{(n)}$-probability is close to 1 could be called *maximally separating subsets* (cf. Figure 1.2).

Maximally separating subsets have a natural interpretation in the context of statistical hypothesis testing. Suppose that a probability distribution of interest is either P or Q and we have to decide between P and Q on the basis of a sample of size n. A set $B \subset A^n$ serves as a test in the sense that we accept P if the sample belongs to B and else we accept Q. One important task is then to minimize the $Q^{(n)}$-probability of wrong decision subject to the constraint $P^{(n)}(B) \geq 1 - \varepsilon$. Hence we see that the maximally separating subsets are just asymptotically optimal tests (cf. (1.9) and (1.10)).

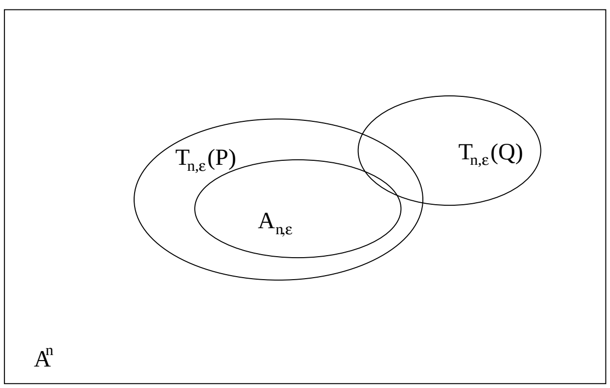

Figure 1.2: Maximally separating set $A_{n,\varepsilon}$

1.3 Data Processing Inequality for the Relative Entropy

Let A, B be finite sets and let $(W(b|a))_{a\in A, b\in B}$ be a stochastic matrix or a channel. Assume that the symbols from A are distributed either according to the distribution P or to Q. The output symbols, after passing the channel, are then distributed either according to $\tilde{P}$ or to $\tilde{Q}$, where

$$\tilde{P}(b) = \sum_{a \in A} P(a)W(b|a) \qquad \text{and} \qquad \tilde{Q}(b) = \sum_{a \in A} Q(a)W(b|a).$$

The data processing inequality for the relative entropy (or monotonicity of the relative entropy) states that

$$D(\tilde{P}, \tilde{Q}) \leq D(P, Q), \tag{1.11}$$

for all pairs of distributions P, Q on A and each channel W. In view of the interpretation of the relative entropy as a measure of distinguishability between two probability distributions described in the previous section the output distributions of the channel are easier confusable than the input distributions. Note that (1.11) is a trivial consequence of the elementary log-sum inequality

$$\sum_{i=1}^{l} u_i \log \frac{u_i}{v_i} \geq \sum_{i=1}^{l} u_i \log \frac{\sum_{i=1}^{l} u_i}{\sum_{i=1}^{l} v_i},$$

where u_i, v_i, $i = 1, \ldots, l$, are non-negative numbers.
The situation changes drastically if we pass to the quantum situation. Here, we consider two density operators D_1 and D_2 acting on the finite dimensional Hilbert space $\mathcal{H}_1$. The role of the stochastic matrix is played by a completely positive trace preserving map $T_\#$ transforming density operators on $\mathcal{H}_1$ into density operators on a finite dimensional Hilbert space $\mathcal{H}_2$. Lindblad proved in [15] that the monotonicity

$$S(T_\#(D_1), T_\#(D_2)) \leq S(D_1, D_2), \tag{1.12}$$

holds. Actually Lindblad showed that inequality for infinite dimensional Hilbert spaces but we pay attention only to finite dimensional case for simplicity. Lindblads proof uses the strong subadditivity of the von Neumann entropy, a deep property of the quantum entropy which was shown by Lieb and Ruskai in [14] using purely analytical tools. Subsequently many other proofs were discovered but all of them have in common the usage of more or less deep analytical considerations. In chapter 4 of this thesis we will present a new proof of (1.12) which is more natural from the information theoretical point of view. The crucial observation in our approach is that the quantum analogue of the limit theorem stated in (1.10) relates the monotonicity properties of the separation numbers $\beta_{\varepsilon,n}$ in (1.9) to the monotonicity of the relative entropy. This can be already demonstrated on the classical side, to some extent, as follows: Consider an algebra of sets $\mathfrak{B}$ which is coarser then the algebra consisting of all subsets of A. Then it is obvious that we increase the separation numbers $\beta_{\varepsilon,n}$ if we perform the minimization in (1.9) only over the sets belonging to $\mathfrak{B}$. The limit assertion (1.10) then shows that

$$D(P \restriction \mathfrak{B}, Q \restriction \mathfrak{B}) \leq D(P, Q).$$

Turning back to the quantum case we can see that the quantum analogue of (1.10) will provide a opportunity to show the monotonicity of the quantum relative entropy under restriction to subalgebras. But this will suffice since, according to a well known result of Stinespring, each completely positive trace preserving map can be represented as composition of three simple maps:

- tensoring with a fixed ancilla state,
- global unitary evolution of the system under consideration and ancilla state,
- restriction to a subalgebra (partial trace).

Observe that the relative entropy is invariant under the first two operations. Hence the crucial operation is the restriction to subalgebras, described in the third item, and we will be able to control the behaviour of the relative entropy under restrictions to subalgebras by means of the limit theorem that will be proved in chapter 4 below.

Chapter 2

Quantum Version of the Shannon-McMillan Theorem

In this chapter we will present the proof of the Shannnon-McMillan theorem for ergodic quantum states which appeared in paper [2] by T. Krüger, Ra. Siegmund-Schultze, A. Szkoła and the author of this thesis. This theorem states that, under the assumption of ergodicity, for large n-blocks the state of the system is concentrated on a subspace whose dimension is of order e^{ns} with s denoting the mean v. Neumann entropy of the system (to be defined in the following section). Moreover, each minimal projection dominated by the projection onto this subspace has expectation value with respect to the state under consideration of order e^{-ns}. Our proof of this theorem is inspired by the methods developed by Hiai and Petz in [8]. In that paper Hiai and Petz considered the problem of discriminating between a completely ergodic state (i.e. the state is ergodic with respect to each power of the shift) and a stationary product state, a problem closely related to the quantum Shannon-McMillan theorem. They succeeded to derive upper and lower bounds on the normalized separation numbers (or error probabilities concerning correct discrimination) depending on the mean quantum relative entropy and conjectured that the upper bound should be the limit of these numbers (A proof of this conjecture is reproduced in the third chapter of this thesis on the basis of joint work [4] with Ra. Siegmund-Schultze). Their method of proof was to show that the quantum relative entropy on a large l-block can be approximated well by the relative entropy of the states after their restriction to an abelian algebra living on that l-block. The abelian quasi-local algebra built up from this large abelian l-block-algebra and the restriction of the states to this algebra provide a classical stochastic process which is a good approximation of the original system. Now the assumption of *complete* ergodicity ensures that this stochastic process is ergodic and hence one can apply classical ergodic theorem for the relative entropy in order to derive the result of Hiai and Petz above.

Roughly, the Shannon McMillan theorem will be established if we can get rid of the restrictive assumption that the system is completely ergodic and if we can prove the conjecture of Hiai and Petz in this situaton for the stationary product state being the tracial state (see the third section of this chapter). Our basic tool to circumvent the assumption of complete ergodicity will be Theorem 2.2.1.

2.1 The Set-up

In this section we recall, in an informal way, the notion of quasilocal C^*-algebras over the lattice $\mathbb{Z}^\nu$. Moreover we fix the notations for this and the remaining chapters of this thesis.

We consider the lattice $\mathbb{Z}^\nu$. To each lattice site $\mathbf{x} \in \mathbb{Z}^\nu$ we assign a finite dimensional C^*-algebra $\mathcal{A}_{\mathbf{x}}$ being $*$-isomorphic to a fixed finite dimensional C^*-algebra $\mathcal{A}$. Recall that each finite dimensional C^*-algebra $\mathcal{A}$ can be thought of as $\bigoplus_{i=1}^n \mathcal{B}(\mathcal{H}_i)$ up to a $*$-isomorphism where the $\mathcal{H}_i$ are finite dimensional Hilbert spaces and $\mathcal{B}(\mathcal{H}_i)$ is the algebra of linear operators on $\mathcal{H}_i$. For a finite $\Lambda \subset \mathbb{Z}^\nu$ the algebra of local observables associated to Λ is defined by

$$\mathcal{A}_\Lambda := \bigotimes_{\mathbf{x}\in\Lambda} \mathcal{A}_{\mathbf{x}}.$$

Clearly, this definition implies that for $\Lambda \subset \Lambda'$ we have $\mathcal{A}_{\Lambda'} = \mathcal{A}_\Lambda \otimes \mathcal{A}_{\Lambda'\setminus\Lambda}$. Hence there is a canonical embedding of $\mathcal{A}_\Lambda$ into $\mathcal{A}_{\Lambda'}$ given by $a \mapsto a \otimes \mathbf{1}_{\Lambda'\setminus\Lambda}$ for $a \in \mathcal{A}_\Lambda$ where $\mathbf{1}_{\Lambda'\setminus\Lambda}$ denotes the identity of $\mathcal{A}_{\Lambda'\setminus\Lambda}$. The quasilocal C^*-algebra $\mathcal{A}^\infty$ is the norm completion of the normed algebra $\bigcup_{\Lambda\subset\mathbb{Z}^\nu} \mathcal{A}_\Lambda$ where the union is taken over all finite subsets Λ. For a precise definition of the quasilocal algebra we refer to [7] and [22]. The group $\mathbb{Z}^\nu$ acts in a natural way on the quasilocal algebra $\mathcal{A}^\infty$ as translations by $\mathbf{x} \in \mathbb{Z}^\nu$. A translation $T(\mathbf{x})$ by $\mathbf{x}$ associates to an $a \in \mathcal{A}_\Lambda$ the corresponding element $T(\mathbf{x})a \in \mathcal{A}_{\Lambda+\mathbf{x}}$. This mapping $T(\mathbf{x})$ extends canonically to a $*$-automorphism on $\mathcal{A}^\infty$ (cf. [7], [22]).

A state ψ on $\mathcal{A}^\infty$ is a linear, positive and unital mapping of $\mathcal{A}^\infty$ into $\mathbb{C}$. Each state ψ on $\mathcal{A}^\infty$ corresponds to a compatible family $\{\psi^{(\Lambda)}\}_{\Lambda\subset\mathbb{Z}^\nu,\#\Lambda<\infty}$ where each $\psi^{(\Lambda)}$ denotes the restriction of ψ to the local algebra $\mathcal{A}_\Lambda$. The compatibility means that $\psi^{(\Lambda')} \upharpoonright \mathcal{A}_\Lambda = \psi^{(\Lambda)}$ for $\Lambda \subset \Lambda'$. A state ψ is called stationary or translationally invariant if $\psi = \psi \circ T(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{Z}^\nu$. The convex set of stationary states is denoted by $\mathcal{T}(\mathcal{A}^\infty)$ and is compact in the weak*-topology. The extremal points in $\mathcal{T}(\mathcal{A}^\infty)$ are the *ergodic* states. The basic theory of ergodic states is described in [7] and [22].

In this thesis the local sets Λ will mainly be boxes in $\mathbb{Z}^\nu$ which are defined for $\mathbf{n} = (n_1, n_2, \ldots, n_\nu) \in \mathbb{N}^\nu$ by

$$\Lambda(\mathbf{n}) := \{0, \ldots, n_1 - 1\} \times \ldots \times \{0, \ldots, n_\nu - 1\},$$

and the cubic boxes of edge length $n \in \mathbb{N}$ will be denoted by $\Lambda(n)$. The local algebras associated with $\Lambda(\mathbf{n})$ will be denoted by $\mathcal{A}^{(\mathbf{n})}$. In a similar fashion we will denote the restriction of a given state ψ on $\mathcal{A}^\infty$ to the local algebra $\mathcal{A}^{(\mathbf{n})}$ by $\psi^{(\mathbf{n})}$. A state $\varphi \in \mathcal{T}(\mathcal{A}^\infty)$ is called a stationary product state if $\varphi^{(\mathbf{n})} = (\varphi^{(1)})^{\otimes \Lambda(\mathbf{n})}$ holds for each $\mathbf{n} \in \mathbb{N}$.

Recall that for every stationary state ψ on $\mathcal{A}^\infty$ the *mean von Neumann entropy* is defined by the limit

$$s(\psi) := \lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} S(\psi^{(\mathbf{n})}).$$

Here $S(\psi^{(\mathbf{n})})$ denotes the *von Neumann entropy* of the state $\psi^{(\mathbf{n})}$ which is given by

$$S(\psi^{(\mathbf{n})}) := -\mathrm{tr}_{\mathbf{n}}(D_{\psi^{(\mathbf{n})}} \log D_{\psi^{(\mathbf{n})}}),$$

where $\mathrm{tr}_{\mathbf{n}}$ denotes the trace on $\mathcal{A}_{\mathbf{n}}$. In the following we shall suppress the index $\mathbf{n}$ to the traces whenever it is clear which local algebra is considered.

The *mean relative entropy* of a stationary state ψ with respect to the stationary product state φ is defined by

$$s(\psi, \varphi) := \lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} S(\psi^{(\mathbf{n})}, \varphi^{(\mathbf{n})}) = \sup_{\Lambda(\mathbf{n})} \frac{1}{\#\Lambda(\mathbf{n})} S(\psi^{(\mathbf{n})}, \varphi^{(\mathbf{n})}). \quad (2.1)$$

The last equation is an easy consequence of the superadditivity of the relative entropy of the state $\psi^{(\mathbf{n})}$ with respect to the *product* state $\varphi^{(\mathbf{n})}$. For the convenience of the reader we recall the definition of the relative entropy of the state σ on a finite dimensional C^*-algebra $\mathcal{C}$ with respect to the state τ:

$$S(\sigma, \tau) := \begin{cases} \mathrm{tr}(D_\sigma(\log D_\sigma - \log D_\tau)) & \text{if } \mathrm{supp}(\sigma) \leq \mathrm{supp}(\tau) \\ \infty & \text{otherwise} \end{cases}$$

where D_σ respectively D_τ denote the density operators of σ respectively τ and $\mathrm{supp}(\sigma)$ is the support projection of the state σ. The important properties of the quantum relative entropy are summed up in the monograph [18] by Ohya and Petz.

Let $l \in \mathbb{N}$ and consider the subgroup $G_l := l \cdot \mathbb{Z}^\nu$ of $\mathbb{Z}^\nu$. Later on we will need the notion of the mean relative entropy of the G_l-stationary state ψ with respect to the G_l-stationary product state φ given by

$$s(\psi, \varphi, G_l) := \lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} S(\psi^{(l\mathbf{n})}, \varphi^{(l\mathbf{n})}). \quad (2.2)$$

Similarily, we set for a G_l-stationary state ψ

$$s(\psi, G_l) := \lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} S(\psi^{(l \cdot \mathbf{n})}). \quad (2.3)$$

Note that if all involved states are also stationary then the quantities $s(\psi, \varphi, G_l)$ resp. $s(\psi, G_l)$ equal $l^\nu s(\psi, \varphi)$ resp. $l^\nu \cdot s(\psi)$.

2.2 An Ergodic Decomposition

As mentioned in the introduction to this chapter our proofs of the limit theorems for the mean v. Neumann entropy as well as for the mean quantum relative entropy will be based on an approximation by classical systems on large boxes. If we would restrict our ergodic quantum state to this classical approximation algebra we would obtain in general a non-ergodic state. The reason for such behaviour is that an ergodic state may have periodicities with respect to a subshift (in our case the action of the group G_l). This is illustrated in the following

Examples: a) Let $B = \{0, 1\}$, consider $B^{\mathbb{N}}$ equipped with the shift T and the σ-algebra $\mathfrak{B}^{\mathbb{N}}$ generated by the cylinder sets. We then define the probability measure P by

$$P = \frac{1}{2}\delta_{(0,1,0,1,\ldots)} + \frac{1}{2}\delta_{(1,0,1,0,\ldots)},$$

where δ's on the right hand side denote Dirac probability measures carried by the indicated points. Then it is obvious that the measure P is stationary, non-ergodic with respect to T^2 and has T^2-ergodic components $\delta_{(0,1,0,1,\ldots)}$ and $\delta_{(1,0,1,0,\ldots)}$. Observe that each T-invariant set $A \in [\mathfrak{B}^{\mathbb{N}}, B^{\mathbb{N}}]$ containing $(0, 1, 0, 1, \ldots)$ for example contains also $(1, 0, 1, 0, \ldots)$ and vice versa. Hence, each T-invariant set A with $P(A) > 0$ fulfills $P(A) = 1$ which shows T-ergodicity of P.

b) We can give even a much simpler example. Consider to this end a positive integer n and the set $\{0, 1, \ldots, 2n-1\}$ with the dynamics given by $T(k) := k + 1 \pmod{2n}$ for $k \in \{0, 1, \ldots, 2n-1\}$. It is obvious that the uniform distribution P on $\{0, 1, \ldots, 2n-1\}$ is T-ergodic. Clearly, the sets O and E of odd and even numbers in $\{0, 1, \ldots, 2n-1\}$ are T^2-invariant and have probabilities which are neither 0 nor 1. The T^2-ergodic decomposition of P is given by

$$P(\,\cdot\,) = \frac{1}{2}P(\,\cdot\,|E) + \frac{1}{2}P(\,\cdot\,|O).$$

The following theorem provides a detailed description of ergodic states possessing periodicities with respect to a subshift: Each ergodic state on the algebra $\mathcal{A}^{\infty}$ can be decomposed into a *finite* number of G_l-ergodic components, G_l being the group $l \cdot \mathbb{Z}$, $l \in \mathbb{N}$, $l > 1$. Moreover, these ergodic components have equal mean v. Neumann entropies and mean quantum relative entropies (a fact we will need in the third chapter). The first three items of that theorem are from the paper [2] by T. Krüger, Ra. Siegmund-Schultze, A. Szkoła and the author. The last item is from the paper [4] by Ra. Siegmund-Schultze and this author.

Theorem 2.2.1 *Let ψ be an ergodic state on $\mathcal{A}^{\infty}$. Then for every subgroup $G_l := l \cdot \mathbb{Z}^{\nu}$, with $l > 1$ an integer, there exists a $\mathbf{k}(l) \in \mathbb{N}^{\nu}$ and a unique convex decomposition of ψ into G_l-ergodic states $\psi_{\mathbf{x}}$:*

$$\psi = \frac{1}{\#\Lambda(\mathbf{k}(l))} \sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \psi_{\mathbf{x}}. \tag{2.4}$$

The G_l-ergodic decomposition (2.4) has the following properties:

1. *$k_j(l) \leq l$ and $k_j(l)|l$ for all $j \in \{1, \ldots, \nu\}$*
2. *$\{\psi_{\mathbf{x}}\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} = \{\psi_0 \circ T(-\mathbf{x})\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))}$*
3. *For every G_l-ergodic state $\psi_{\mathbf{x}}$ in the convex decomposition (2.4) of ψ the mean entropy with respect to G_l, $s(\psi_{\mathbf{x}}, G_l)$, is equal to the mean entropy $s(\psi, G_l)$, i.e.*

   $$s(\psi_{\mathbf{x}}, G_l) = s(\psi, G_l) \tag{2.5}$$

   for all $\mathbf{x} \in \Lambda(\mathbf{k}(l))$.
4. *For each G_l-ergodic state $\psi_{\mathbf{x}}$ in the convex decomposition (2.4) of ψ and for every stationary product state φ with $s(\psi, \varphi) < \infty$, the mean relative entropy with respect to G_l, $s(\psi_{\mathbf{x}}, \varphi, G_l)$, fulfills*

   $$s(\psi_{\mathbf{x}}, \varphi, G_l) = s(\psi, \varphi, G_l) \tag{2.6}$$

   for all $\mathbf{x} \in \Lambda(\mathbf{k}(l))$.

Proof of Theorem 2.2.1: Let $(\mathcal{H}_\psi, \pi_\psi, \Omega_\psi, U_\psi)$ be the GNS representation of the C^*-dynamical system $(\mathcal{A}^\infty, \psi, \mathbb{Z}^\nu)$. U_ψ is the unitary representation of $\mathbb{Z}^\nu$ on $\mathcal{H}_\psi$. Note that the Hilbert space $\mathcal{H}_\psi$ is separable by GNS construction since the algebra $\mathcal{A}^\infty$ is separable. Recall the following properties of the GNS quadruple for every $\mathbf{x} \in \mathbb{Z}^\nu$:

$$U_\psi(\mathbf{x})\Omega_\psi = \Omega_\psi, \tag{2.7}$$
$$U_\psi(\mathbf{x})\pi_\psi(a)U_\psi^*(\mathbf{x}) = \pi_\psi(T(\mathbf{x})a), \qquad \forall a \in \mathcal{A}^\infty. \tag{2.8}$$

Define

$$\mathcal{N}_{\psi,G_l} := \pi_\psi(\mathcal{A}^\infty) \cup U_\psi(G_l), \tag{2.9}$$
$$\mathcal{P}_{\psi,G_l} := \{P \in \mathcal{N}'_{\psi,G_l} : \ P = P^* = P^2\}. \tag{2.10}$$

By $'$ we denote the commutant. Observe that $\mathcal{N}_{\psi,G_l}$ is selfadjoint. Thus $\mathcal{N}'_{\psi,G_l}$ (as the commutant of a selfadjoint set) is a von Neumann algebra. Further it is a known result that $\mathcal{N}'_{\psi,G_l}$ is abelian, (cf. Proposition 4.3.7. in [7] or Lemma IV.3.4 in [10]). This is essentially due to the fact that the quasilocal algebra is by construction G_l asymptotically abelian. The details can be found in the references quoted above.

Consider some $l > 1$ such that ψ is not G_l-ergodic. (If there is no such l the statement of the theorem is trivial.) This implies that $\mathcal{N}_{\psi,G_l}$ is reducible ande hence that

$$\mathcal{P}_{\psi,G_l}\backslash\{0,\mathbf{1}\} \neq \emptyset, \tag{2.11}$$

holds.

Let I be a countable index set. An implication of the $\mathbb{Z}^\nu$-ergodicity of the G_l-invariant ψ is the following:

$$\{Q_i\}_{i\in I} \text{ orthogonal partition of unity in } \mathcal{N}'_{\psi,G_l} \Longrightarrow |I| \leq l^\nu. \tag{2.12}$$

To see (2.12) observe at first that for any $Q \in \mathcal{P}_{\psi,G_l}\backslash\{0\}$ the projection $U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})$, $\mathbf{x} \in \Lambda(l)$, belongs to the abelian algebra $\mathcal{N}'_{\psi,G_l}$, namely

$$\begin{aligned} \pi_\psi(a)U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x}) &= U_\psi(\mathbf{x})\pi_\psi(T(-\mathbf{x})a)QU^*_\psi(\mathbf{x}) \quad (\text{by } (2.8)) \\ &= U_\psi(\mathbf{x})Q\pi_\psi(T(-\mathbf{x})a)U^*_\psi(\mathbf{x}) \\ &= U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})\pi_\psi(a) \quad (\text{by } (2.8)) \end{aligned}$$

holds for every $a \in \mathcal{A}^\infty$ and $[U_\psi(\mathbf{y}), U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})] = 0$ is obvious by $[U_\psi(\mathbf{y}), U_\psi(\mathbf{x})] = 0$ for all $\mathbf{y} \in G_l$ and $\mathbf{x} \in \mathbb{Z}^\nu$. Thus $\{U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})\}_{\mathbf{x}\in\Lambda(l)}$ is a family of mutually commuting projections. Define $\bar{Q} := \bigvee_{\mathbf{x}\in\Lambda(l)} U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})$. Note that if $Q \in \mathcal{P}_{\psi,G_l}\backslash\{0\}$ then for any $\mathbf{y} \in \mathbb{Z}^\nu$ we have $U_\psi(\mathbf{y})QU^*_\psi(\mathbf{y}) = U_\psi(\mathbf{y}(\text{mod } \mathbf{l}))QU^*_\psi(\mathbf{y}(\text{mod } \mathbf{l}))$, where $\mathbf{l} = (l,\ldots,l) \in \mathbb{Z}^\nu$. This means that $\{U_\psi(\mathbf{y})U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})U^*_\psi(\mathbf{y})\}_{\mathbf{x}\in\Lambda(l)} = \{U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})\}_{\mathbf{x}\in\Lambda(l)}$ and consequently $\bar{Q}$ is invariant under the action of $U_\psi(\mathbb{Z}^\nu)$. From the $\mathbb{Z}^\nu$-ergodicity of ψ we deduce that $\bar{Q} = \mathbf{1}$. The Gelfand isomorphism represents the projections $U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})$ as continuous characteristic functions 1_{Q_x} on some compact (totally disconnected) space. Thus $\bar{Q}$ has the representation as $\bigvee_{\mathbf{x}\in\Lambda(l)} 1_{Q_\mathbf{x}} = 1_{\bigcup_{\mathbf{x}\in\Lambda(l)} Q_\mathbf{x}}$. If we translate back the finite subadditivity of probability measures to the expectation values of the projections $U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})$ we obtain:

$$\begin{aligned} 1 = \langle\Omega_\psi, \bar{Q}\Omega_\psi\rangle &\leq \sum_{\mathbf{x}\in\Lambda(l)} \langle\Omega_\psi, U_\psi(\mathbf{x})QU^*_\psi(\mathbf{x})\Omega_\psi\rangle \\ &= l^\nu \cdot \langle\Omega_\psi, Q\Omega_\psi\rangle \quad (\text{by } (2.7)). \end{aligned}$$

Thus (2.12) is clear.
Combining the results (2.11) and (2.12) we get the existence of an orthogonal partition of unity $\{P_i\}_{i=0}^{n_l-1}$ in $n_l \leq l^\nu$ minimal projections $P_i \in \mathcal{P}_{\psi,G_l}\backslash\{0,\mathbf{1}\}$. The abelianess of $\mathcal{N}'_{\psi,G_l}$ implies the uniqueness of the orthogonal partition

of unity $\{P_i\}_{i=0}^{n_l-1}$. Further it follows that $\{P_i\}_{i=0}^{n_l-1}$ is a generating subset for $\mathcal{P}_{\psi,G_l}$ in the following sense:

$$Q \in \mathcal{P}_{\psi,G_l} \implies \exists \{P_{i_j}\}_{j=0}^{s\leq n_l-1} \subset \mathcal{P}_{\psi,G_l} \text{ such that } Q = \sum_{j=0}^{s} P_{i_j}. \tag{2.13}$$

Define $p_i := \langle \Omega_\psi, P_i \Omega_\psi \rangle$ and order the minimal projections P_i such that

$$p_0 \leq p_i, \qquad \forall i \in \{1, \dots, n_l - 1\}. \tag{2.14}$$

Let

$$G(P_0) := \{\mathbf{x} \in \mathbb{Z}^\nu : \; U(\mathbf{x}) P_0 U^*(\mathbf{x}) = P_0\}.$$

Note that $G(P_0)$ is a subgroup of $\mathbb{Z}^\nu$ and contains G_l, since $P_0 \in \mathcal{P}_{\psi,G_l}$. This leads to the representation

$$G(P_0) = \bigoplus_{j=1}^{\nu} k_j(l)\mathbb{Z}, \quad \text{with } k_j(l)|l \quad \text{ for all } j \in \{1, \dots, \nu\},$$

where the integers $k_j(l)$ are given by

$$k_j(l) := \min\{x_j : \; x_j \text{ is the j-th component of } \mathbf{x} \in G(P_0) \text{ and } x_j > 0\}.$$

For P_0, as an element of $\mathcal{P}_{\psi,G_l}$, $\{U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x})\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \subseteq \mathcal{P}_{\psi,G_l}$ for $\mathbf{k}(l) = (k_1(l), \dots, k_\nu(l))$. Thus by (2.13) each $U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x})$, $\mathbf{x} \in \Lambda(\mathbf{k}(l))$, can be represented as a sum of minimal projections. But then by linearity of the expectation values and the assumed ordering (2.14) each $U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x})$ must be a minimal projection for $\mathbf{x} \in \Lambda(\mathbf{k}(l))$. Otherwise there would be a contradiction to $\langle \Omega_\psi, U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x}) \Omega_\psi \rangle = p_0$. Consequently $\{U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x})\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \subseteq \{P_i\}_{i=0}^{n_l-1}$. Consider the projection $\bar{P}_0 = \sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x})$. Using the same argument as for $\bar{Q}$, defined a few lines above, we see that $\bar{P}_0$ is invariant under the action of $U_\psi(\mathbb{Z}^\nu)$ and because of the $\mathbb{Z}^\nu$-ergodicity of ψ

$$\bar{P}_0 = \mathbf{1}.$$

It follows by the uniqueness of the orthogonal partition of unity

$$\{U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x})\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} = \{P_j\}_{j=0}^{n_l-1}.$$

Obviously $n_l = \#\Lambda(\mathbf{k}(l))$ and for each P_i, $i \in \{0, \dots, n_l - 1\}$, there is only one $\mathbf{x} \in \Lambda(\mathbf{k}(l))$ such that

$$P_i = U_\psi(\mathbf{x}) P_0 U_\psi^*(\mathbf{x}) =: P_\mathbf{x}. \tag{2.15}$$

It follows $p_i = p_0$ for all $i \in \{0, \dots, n_l - 1\}$ and hence

$$p_i = \frac{1}{n_l} = \frac{1}{\#\Lambda(\mathbf{k}(l))}, \qquad i \in \{0, \dots, n_l - 1\}.$$

Finally, set for every $\mathbf{x} \in \Lambda(\mathbf{k}(l))$

$$\psi_{\mathbf{x}}(a) \quad := \quad \#\Lambda(\mathbf{k}(l))\langle \Omega_\psi, P_{\mathbf{x}}\pi_\psi(a)\Omega_\psi\rangle, \qquad a \in \mathcal{A}^\infty.$$

From (2.15), (2.7) and (2.8) we get

$$\begin{aligned}\psi_{\mathbf{x}}(a) &= \#\Lambda(\mathbf{k}(l))\langle \Omega_\psi, P_{\mathbf{x}}\pi_\psi(a)\Omega_\psi\rangle \\ &= \#\Lambda(\mathbf{k}(l))\langle \Omega_\psi, P_0\pi_\psi(T(-\mathbf{x})a)\Omega_\psi\rangle \\ &= \psi_0(T(-\mathbf{x})a), \qquad a \in \mathcal{A}^\infty,\end{aligned}$$

hence

$$\begin{aligned}\sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \langle \Omega_\psi, P_{\mathbf{x}}\pi_\psi(a)\Omega_\psi\rangle &= \langle \Omega_\psi, (\sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} P_{\mathbf{x}})\pi_\psi(a)\Omega_\psi\rangle \\ &= \psi(a).\end{aligned}$$

Thus we arrive at the convex decomposition of ψ:

$$\psi = \frac{1}{\#\Lambda(\mathbf{k}(l))} \sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \psi_0 \circ T(-\mathbf{x}).$$

By construction this is a G_l-ergodic decomposition of ψ. It remains to prove the fact that the mean entropies with respect to the lattice G_l are the same for all G_l-ergodic components $\psi_{\mathbf{x}}$.

Proof of item 3.: It is a well known result that the quantum mean entropy with respect to a given lattice G_l is affine on the convex set of G_l-invariant states, (cf. prop. 7.2.3 in [22]). Thus to prove (2.6) it is sufficient to show:

$$s(\psi_{\mathbf{x}}, G_l) = s(\psi_0, G_l), \qquad \forall \mathbf{x} \in \Lambda(\mathbf{k}(l)).$$

By the definition of the mean entropy this is equivalent to the statement

$$|S(\psi_{\mathbf{x}}^{(l\mathbf{n})}) - S(\psi_0^{(l\mathbf{n})})| = o(|\mathbf{n}|) \qquad \text{as } \mathbf{n} \to \infty. \tag{2.16}$$

This can be seen as follows: In view of the definition of $\psi_{\mathbf{x}}^{(l\mathbf{n})}$ we have $S(\psi_{\mathbf{x}}^{(l\mathbf{n})}) = S(\psi_{\mathbf{x}}^{(\Lambda(l\mathbf{n}))}) = S(\psi_0^{(\Lambda(l\mathbf{n})-\mathbf{x})})$. We introduce the box $\tilde{\Lambda}$ being concentric with $\Lambda(l\mathbf{n})$, with all edges enlarged by l on both directions, i.e. an l-neighborhood of $\Lambda(l\mathbf{n})$. The two expressions $S(\psi_{\mathbf{x}}^{(l\mathbf{n})})$ and $S(\psi_0^{(l\mathbf{n})})$ are von Neumann entropies of the restrictions of $\psi_0^{(\widetilde{\Lambda})}$ to the smaller sets $\Lambda(l\mathbf{n})$ and $\Lambda(l\mathbf{n}) - \mathbf{x}$, respectively. On the other hand we consider the box $\widehat{\Lambda}$ being concentric with $\Lambda(l\mathbf{n})$ with all edges shortened by l at both sides. $S(\psi_0^{(\widehat{\Lambda})})$ is the von Neumann entropy of $\psi_0^{(\Lambda(l\mathbf{n}))}$ and $\psi_{\mathbf{x}}^{(\Lambda(l\mathbf{n}))}$ after their restriction to the set $\widehat{\Lambda}$. $S(\psi_{\mathbf{x}}^{(\Lambda(l\mathbf{n}))})$ and $S(\psi_0^{(\Lambda(l\mathbf{n}))})$ can be estimated simultaneously using the subadditivity of the von Neumann entropy

$$S(\psi_0^{(\widetilde{\Lambda})}) - \log \operatorname{tr}_{\tilde{\Lambda}\setminus\Lambda(l\mathbf{n})}\mathbf{1} \le S(\psi_\natural^{(l\mathbf{n})}) \le S(\psi_0^{(\widehat{\Lambda})}) + \log \operatorname{tr}_{\Lambda(l\mathbf{n})\setminus\widehat{\Lambda}}\mathbf{1},$$

where $\natural \in \{\mathbf{x}, 0\}$. Thus (2.16) is immediate.

Proof of item 4.: The proof of the last item is based on the monotonicity of the relative entropy and the usage of item 2. of the theorem. For each $\mathbf{n} \in \mathbb{N}^\nu$ and $\mathbf{x} \in \Lambda(\mathbf{k}(l))$ we have $\psi_{\mathbf{x}}^{(\Lambda(l\mathbf{n}))} = \psi_0^{(\Lambda(l\mathbf{n})-\mathbf{x})}$ by the second item. We consider the box $\tilde{\Lambda}(\mathbf{n})$ containing $\Lambda(l\mathbf{n})$ and each $\Lambda(l\mathbf{n}) - \mathbf{x}$ defined by

$$\tilde{\Lambda}(\mathbf{n}) := \{-l, \dots, ln_1 - 1\} \times \dots \times \{-l, \dots, ln_\nu - 1\},$$

and the box $\hat{\Lambda}(\mathbf{n})$ contained in $\Lambda(l\mathbf{n})$ and $\Lambda(l\mathbf{n}) - \mathbf{x}$ given by

$$\hat{\Lambda}(\mathbf{n}) := \{0, \dots, l(n_1 - 1) - 1\} \times \dots \times \{0, \dots, l(n_\nu - 1) - 1\}.$$

The volumes of these boxes are asymptotically equivalent in the sense that the quotient tends to one. Hence using the observation above and twice the monotonicity of the relative entropy we obtain

$$\begin{aligned} S(\psi_0^{(\tilde{\Lambda}(\mathbf{n}))}, \varphi^{(\tilde{\Lambda}(\mathbf{n}))}) &\geq S(\psi_0^{(\Lambda(l\mathbf{n})-\mathbf{x})}, \varphi^{(\Lambda(l\mathbf{n})-\mathbf{x})}) \\ &= S(\psi_{\mathbf{x}}^{(\Lambda(l\mathbf{n}))}, \varphi^{(\Lambda(l\mathbf{n})-\mathbf{x})}) \\ &\geq S(\psi_{\mathbf{x}}^{(\hat{\Lambda}(\mathbf{n}))}, \varphi^{(\hat{\Lambda}(\mathbf{n}))}). \end{aligned}$$

After dividing by $\#\Lambda(\mathbf{n})$ and taking the limit $\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu$ this inequality chain shows that

$$s(\psi_0, \varphi, G_l) \geq s(\psi_{\mathbf{x}}, \varphi, G_l)$$

holds. A similar argument using $\psi_{\mathbf{x}}^{(\Lambda(l\mathbf{n})+\mathbf{x})} = \psi_0^{(\Lambda(l\mathbf{n}))}$ shows that the reverse inequality is also valid. Hence we established

$$s(\psi_0, \varphi, G_l) = s(\psi_{\mathbf{x}}, \varphi, G_l).$$

Since the mean relative entropy is affine in its first argument on the set of G_l-invariant states this implies

$$s(\psi_{\mathbf{x}}, \varphi, G_l) = s(\psi, \varphi, G_l) = l^\nu s(\psi, \varphi). \qquad \square$$

2.3 Quantum Shannon-McMillan Theorem

We are now prepared to state and prove the main theorem of this chapter (and perhaps of the whole thesis). This is the full extension of the fundamental Shannon-McMillan theorem of classical information theory to the quantum mechanical setting. Note that this theorem is sometimes refered to as *asymptotic equipartition property* (AEP).

Theorem 2.3.1 (Quantum Shannon-McMillan Theorem) *Let ψ be an ergodic state on $\mathcal{A}^\infty$ with mean entropy $s(\psi)$. Then for all $\delta > 0$ there is an $\mathbf{N}_\delta \in \mathbb{N}^\nu$ such that for all $\mathbf{n} \in \mathbb{N}^\nu$ with $\Lambda(\mathbf{n}) \supseteq \Lambda(\mathbf{N}_\delta)$ there exists an orthogonal projection $p_{\mathbf{n}}(\delta) \in \mathcal{A}^{(\mathbf{n})}$ such that*

1. $\psi^{(\mathbf{n})}(p_{\mathbf{n}}(\delta)) \geq 1-\delta$,

2. for all minimal projections $p \in \mathcal{A}^{(\mathbf{n})}$ *with* $p \leq p_{\mathbf{n}}(\delta)$

$$e^{-\#(\Lambda(\mathbf{n}))(s(\psi)+\delta)} < \psi^{(\mathbf{n})}(p) < e^{-\#(\Lambda(\mathbf{n}))(s(\psi)-\delta)},$$

3. $e^{\#(\Lambda(\mathbf{n}))(s(\psi)-\delta)} < \mathrm{tr}_{\mathbf{n}}(p_{\mathbf{n}}(\delta)) < e^{\#(\Lambda(\mathbf{n}))(s(\psi)+\delta)}$.

Remark: *In complete analogy to the classical case the above theorem, in particular item 2., expresses the AEP in a quantum context. As will be seen in the proof of this theorem the typical subspace given by the projection* $p_{\mathbf{n}}(\delta)$ *can be chosen as the linear hull of the eigenvectors of* $\psi^{(\mathbf{n})}$ *which have eigenvalues of order* $e^{-\#(\Lambda(\mathbf{n}))s(\psi)}$.

From the point of view of quantum data storage one is looking for essentially carrying subspaces the dimension of which is as small as possible. The circumstance that such subspaces have probability close to 1 ensures that the probability of decoding error vanishes asymptotically. The requirement of small dimensionality for essentially carrying subspaces leads to an reduction of resources needed for the data storage. Our next proposition is the crucial step towards the precise quantitative statement of the previously described intuition about quantum data compression. As the following proofs will show, this proposition is equivalent to the quantum Shannon-McMillan theorem.

For $\varepsilon \in (0,1)$ and $\mathbf{n} \in \mathbb{N}^{\nu}$ we define the *dimension covering exponent* by

$$\beta_{\varepsilon,\mathbf{n}}(\psi) := \min\{\log \mathrm{tr}(q) : \ q \in \mathcal{A}^{(\mathbf{n})} \text{ projection}, \ \psi^{(\mathbf{n})}(q) \geq 1-\varepsilon\}. \quad (2.17)$$

Proposition 2.3.2 *Let* ψ *be an ergodic state on* $\mathcal{A}^{\infty}$ *with mean entropy* $s(\psi)$*. Then for every* $\varepsilon \in (0,1)$

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^{\nu}} \frac{1}{\#(\Lambda(\mathbf{n}))}\beta_{\varepsilon,\mathbf{n}}(\psi) = s(\psi). \quad (2.18)$$

2.3.1 Proof of the Shannon-McMillan Theorem

A basic tool for the proof of the Shannon-McMillan theorem under the general assumption of ergodicity is the structural assertion Theorem 2.2.1. It is used to circumvent the complete ergodicity assumption made by Hiai and Petz. The substitution of the quantum system by a classical approximation on large boxes leads to an ergodicity problem for these classical approximations. Theorem 2.2.1 combined with the subsequent lemma allow to control not only the mean (per site limit) entropies of the ergodic components (with respect to the subshift generated by a large box), but also to cope with the obstacle that some of these components might have an atypical entropy on this large but finite box. Using these prerequisites, we prove Lemma 2.3.4 which is the extension of the Hiai/Petz upper bound result to ergodic states. Finally, from the simple probabilistic argument expressed in Lemma 2.3.5 we infer that the upper bound is really a limit.

In order to simplify our notation in the next lemma we introduce some abbreviations. We choose a positive integer l and consider the decomposition of an ergodic state $\psi \in \mathcal{T}(\mathcal{A}^\infty)$ into states $\psi_{\mathbf{x}}$ being ergodic with respect to the action of G_l, i.e. $\psi = \frac{1}{\#(\Lambda(\mathbf{k}(l)))} \sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \psi_{\mathbf{x}}$. Then we set

$$s := s(\psi, \mathbb{Z}^\nu) = s(\psi),$$

i.e. the mean entropy of the state ψ computed with respect to $\mathbb{Z}^\nu$. Moreover we set

$$s_{\mathbf{x}}^{(l)} := \frac{1}{\#(\Lambda(l))} S(\psi_{\mathbf{x}}^{(\Lambda(l))}) \quad \text{and} \quad s^{(l)} := \frac{1}{\#(\Lambda(l))} S(\psi^{(\Lambda(l))}).$$

From the third item of Theorem 2.2.1 we know that

$$s(\psi_{\mathbf{x}}, G_l) = s(\psi, G_l) = l^\nu \cdot s(\psi), \qquad \forall \mathbf{x} \in \Lambda(\mathbf{k}(l)). \tag{2.19}$$

For $\eta > 0$ let us introduce the following set

$$A_{l,\eta} := \{\mathbf{x} \in \Lambda(\mathbf{k}(l)) : \ s_{\mathbf{x}}^{(l)} \geq s + \eta\}. \tag{2.20}$$

By $A_{l,\eta}^c$ we denote its complement. The following lemma states that the density of G_l-ergodic components of ψ which have too large entropy on the box of side length l vanishes asymptotically in l.

Lemma 2.3.3 *If ψ is a $\mathbb{Z}^\nu$-ergodic state on $\mathcal{A}^\infty$, then*

$$\lim_{l\to\infty} \frac{\#A_{l,\eta}}{\#\Lambda(\mathbf{k}(l))} = 0$$

holds for every $\eta > 0$.

Proof of Lemma 2.3.3: We suppose on the contrary that there is some $\eta_0 > 0$ such that $\limsup_l \frac{\#A_{l,\eta_0}}{\#\Lambda(\mathbf{k}(l))} = a > 0$. Then there exists a subsequence (l_j) with the property

$$\lim_{j\to\infty} \frac{\#A_{l_j,\eta_0}}{\#\Lambda(\mathbf{k}(l_j))} = a.$$

By the concavity of the von Neumann entropy we obtain

$$\begin{aligned}
\#\Lambda(\mathbf{k}(l_j)) \cdot s^{(l_j)} &\geq \sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l_j))} s_{\mathbf{x}}^{(l_j)} \\
&= \sum_{\mathbf{x}\in A_{l_j,\eta_0}} s_{\mathbf{x}}^{(l_j)} + \sum_{\mathbf{x}\in A_{l_j,\eta_0}^c} s_{\mathbf{x}}^{(l_j)} \\
&\geq \#A_{l_j,\eta_0} \cdot (s+\eta_0) + \#A_{l_j,\eta_0}^c \cdot \min_{\mathbf{x}\in A_{l_j,\eta_0}^c} s_{\mathbf{x}}^{(l_j)}.
\end{aligned}$$

Here we made use of (2.20) at the last step. Using that for the mean entropy holds

$$s(\psi_{\mathbf{x}}, G_l) = \lim_{\Lambda(\mathbf{m})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{m})} S(\psi_{\mathbf{x}}^{(l\mathbf{m})}) = \inf_{\Lambda(\mathbf{m})} \frac{1}{\#\Lambda(\mathbf{m})} S(\psi_{\mathbf{x}}^{(l\mathbf{m})})$$

we obtain a further estimation for the second term on the right hand side:

$$\begin{aligned}\#A^c_{l_j,\eta_0} \cdot \min_{\mathbf{x}\in A^c_{l_j,\eta_0}} s_{\mathbf{x}}^{(l_j)} &\geq \#A^c_{l_j,\eta_0} \cdot \min_{\mathbf{x}\in A^c_{l_j,\eta_0}} \frac{1}{l_j^\nu} s(\psi_{\mathbf{x}}, G_{l_j}) \\ &= \#A^c_{l_j,\eta_0} \cdot s(\psi) \qquad (\text{by}(2.19)).\end{aligned}$$

After dividing $\#\Lambda(\mathbf{k}(l_j))\cdot s^{(l_j)} \geq \#A_{l_j,\eta_0}\cdot(s+\eta_0)+\#A^c_{l_j,\eta_0}\cdot s(\psi)$ by $\#\Lambda(\mathbf{k}(l_j))$ and taking limits we arrive at the following contradictory inequality:

$$s \geq a(s+\eta_0) + (1-a)s = s + a\eta_0 > s.$$

So, $a = 0$. $\square$

Lemma 2.3.4 *Let ψ be an ergodic state on $\mathcal{A}^\infty$. Then for every $\varepsilon \in (0,1)$*

$$\limsup_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} \beta_{\varepsilon,\mathbf{n}}(\psi) \leq s(\psi).$$

Proof of Lemma 2.3.4: We fix $\varepsilon > 0$ and choose arbitrary η, $\delta > 0$. Consider the G_l-ergodic decomposition

$$\psi = \frac{1}{\#\Lambda(\mathbf{k}(l))} \sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \psi_x$$

of ψ for integers $l \geq 1$. By Lemma 2.3.3 there is an integer $L \geq 1$ such that for any $l \geq L$

$$\frac{\varepsilon}{2} \geq \frac{1}{\#\Lambda(\mathbf{k}(l))} \#A_{l,\eta} \geq 0$$

holds, where $A_{l,\eta}$ is defined by (2.20). This inequality implies

$$\frac{1}{\#\Lambda(\mathbf{k}(l))} \#A^c_{l,\eta} \cdot (1 - \frac{\varepsilon}{2}) \geq 1 - \varepsilon. \tag{2.21}$$

On the other hand by

$$S(\psi^{(\mathbf{n})}) = \inf\{S(\psi^{(\mathbf{n})} \restriction \mathcal{B}) : \ \mathcal{B} \text{ maximal abelian } C^* - \text{subalgebra of } \mathcal{A}^{(\mathbf{n})}\}$$

(cf. Theorem 11.9 in [17] and use the one-to-one correspondence between maximal abelian $*$-subalgebras and orthogonal partitions of unity into minimal projections contained in $\mathcal{A}^{(\mathbf{n})}$) there exist maximal abelian C^*-subalgebras $\mathcal{B}_{\mathbf{x}}$ of $\mathcal{A}_{\Lambda(l)}$ with the property

$$\frac{1}{\#\Lambda(l)} S(\psi_{\mathbf{x}}^{(\Lambda(l))} \upharpoonright \mathcal{B}_{\mathbf{x}}) < s(\psi) + \eta, \qquad \forall\, \mathbf{x} \in A_{l,\eta}^c. \tag{2.22}$$

We fix an $l \geq L$ and consider the abelian quasilocal C^*-algebras $\mathcal{B}_{\mathbf{x}}^{\infty}$, constructed with $\mathcal{B}_{\mathbf{x}}$, as C^*-subalgebras of $\mathcal{A}^{\infty}$ and set

$$m_{\mathbf{x}} := \psi_{\mathbf{x}} \upharpoonright \mathcal{B}_{\mathbf{x}}^{\infty} \text{ and } m_{\mathbf{x}}^{(\mathbf{n})} := \psi_{\mathbf{x}} \upharpoonright \mathcal{B}_{\mathbf{x}}^{(\mathbf{n})}$$

for $\mathbf{x} \in A_{l,\eta}^c$ and $\mathbf{n} \in \mathbb{N}^{\nu}$. The states $m_{\mathbf{x}}$ are G_l-ergodic since they are restrictions of G_l-ergodic states $\psi_{\mathbf{x}}$ on a quasilocal algebra. This easily follows from Theorem 4.3.17. in [7]. Moreover, by the Gelfand isomorphism and Riesz representation theorem, we can identify the states $m_{\mathbf{x}}$ with probability measures on corresponding (compact) maximal ideal spaces of $\mathcal{B}_{\mathbf{x}}^{\infty}$. By commutativity and finite dimensionality of the algebras $\mathcal{B}_{\mathbf{x}}$ these compact spaces can be represented as $B_{\mathbf{x}}^{\mathbb{Z}^{\nu}}$ with finite sets $B_{\mathbf{x}}$ for all $\mathbf{x} \in A_{l,\eta}^c$. By the Shannon-McMillan-Breiman theorem (cf. [20], [16])

$$\lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^{\nu}} -\frac{1}{\#\Lambda(\mathbf{n})} \log m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}}) = h_{\mathbf{x}} \tag{2.23}$$

$m_{\mathbf{x}}$-almost surely and in $L^1(m_{\mathbf{x}})$ for all $\mathbf{x} \in A_{l,\eta}^c$, where $h_{\mathbf{x}}$ denotes the Kolmogorov-Sinai entropy of $m_{\mathbf{x}}$, and $\omega_{\mathbf{n}} \in {B_{\mathbf{x}}}^{\Lambda(\mathbf{n})}$ are the components of $\omega \in B_{\mathbf{x}}^{\mathbb{Z}^{\nu}}$ corresponding to the box $\Lambda(\mathbf{n})$. Actually, as we shall see, we need the theorem cited above only in its weaker form (convergence in probability) known as Shannon-McMillan theorem. For each n and $\mathbf{x} \in A_{l,\eta}^c$ let

$$\begin{aligned} C_{\mathbf{x}}^{(\mathbf{n})} &:= \{\omega_{\mathbf{n}} \in B_{\mathbf{x}}^{(\mathbf{n})} : \ | -\frac{1}{\#\Lambda(\mathbf{n})} \log m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}}) - h_{\mathbf{x}}| < \delta\} \\ &= \{\omega_{\mathbf{n}} \in B_{\mathbf{x}}^{(\mathbf{n})} : \ e^{-\#\Lambda(\mathbf{n})\cdot(h_{\mathbf{x}}+\delta)} < m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}}) < e^{-\#\Lambda(\mathbf{n})\cdot(h_{\mathbf{x}}-\delta)}\}. \end{aligned}$$

Since lower bounds on the probability imply upper bounds on the cardinality we obtain

$$\#C_{\mathbf{x}}^{(\mathbf{n})} = \mathrm{tr}_{\mathbf{n}} \left(p_{\mathbf{x}}^{(\mathbf{n})}\right) \leq e^{\#\Lambda(\mathbf{n})\cdot(h_{\mathbf{x}}+\delta)} \leq e^{\#\Lambda(\mathbf{n})\cdot(l^{\nu}(s(\psi)+\eta)+\delta)} \tag{2.24}$$

where $p_{\mathbf{x}}^{(\mathbf{n})}$ is the projection in $\mathcal{B}_{\mathbf{x}}^{(\mathbf{n})}$ corresponding to the function $1_{C_{\mathbf{x}}^{(\mathbf{n})}}$. In the last inequality we have used that $h_{\mathbf{x}} \leq S(\psi_{\mathbf{x}}^{(\Lambda(l))} \upharpoonright \mathcal{B}_{\mathbf{x}}) < l^{\nu}(s(\psi) + \eta)$ for all $\mathbf{x} \in A_{l,\eta}^c$ by

$$h_{\mathbf{x}} = \lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^{\nu}} \frac{1}{\#\Lambda(\mathbf{n})} H(m_{\mathbf{x}}^{(\mathbf{n})}) = \inf_{\Lambda(\mathbf{n})} \frac{1}{\#\Lambda(\mathbf{n})} H(m_{\mathbf{x}}^{(\mathbf{n})}),$$

(cf. [22]), and by (2.22). Here H denotes the Shannon entropy.
From (2.23) it follows that there is an $N \in \mathbb{N}$ (depending on l) such that for all $\mathbf{n} \in \mathbb{N}^\nu$ with $\Lambda(\mathbf{n}) \supset \Lambda(N)$

$$m_{\mathbf{x}}^{(\mathbf{n})}(C_{\mathbf{x}}^{(\mathbf{n})}) \geq 1 - \frac{\varepsilon}{2}, \qquad \forall\, \mathbf{x} \in A_{l,\eta}^c. \tag{2.25}$$

For each $\mathbf{y} \in \mathbb{N}^\nu$ with $y_i \geq Nl$ let $y_i = n_i l + j_i$, where $n_i \geq N$ and $0 \leq j_i < l$. We set

$$q^{(l\mathbf{n})} := \bigvee_{\mathbf{x} \in A_{l,\eta}^c} p_{\mathbf{x}}^{(\mathbf{n})}.$$

and denote by $q_{\mathbf{y}}$ the embedding of $q^{(l\mathbf{n})}$ in $\mathcal{A}^{(\mathbf{y})}$, i.e. $q_{\mathbf{y}} = q^{(l\mathbf{n})} \otimes \mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}$. By (2.25) and (2.21) we obtain

$$\begin{aligned}
\psi^{(\mathbf{y})}(q_{\mathbf{y}}) &= \frac{1}{\#\Lambda(\mathbf{k}(l))} \sum_{\mathbf{x} \in \Lambda(\mathbf{k}(l))} \psi_{\mathbf{x}}^{(\mathbf{y})}(q_{\mathbf{y}}) \\
&\geq \frac{1}{\#\Lambda(\mathbf{k}(l))} \# A_{l,\eta}^c \cdot (1 - \frac{\varepsilon}{2}) \geq (1 - \varepsilon).
\end{aligned}$$

Thus the condition in the definition of $\beta_{\varepsilon,\mathbf{y}}(\psi)$ is satisfied. Moreover by the definition of $q_{\mathbf{y}}$ and (2.24)

$$\begin{aligned}
\beta_{\varepsilon,\mathbf{y}}(\psi) &\leq \log \mathrm{tr}_{\mathbf{y}}(q_{\mathbf{y}}) = \log \mathrm{tr}_{l\mathbf{n}}(q^{(l\mathbf{n})}) + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}(\mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})})) \\
&\leq \log(\sum_{\mathbf{x} \in A_{l,\eta}^c} e^{\#\Lambda(\mathbf{n})\cdot(h_{\mathbf{x}}+\delta)}) + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}(\mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})})) \\
&\leq \log(\# A_{l,\eta}^c \cdot e^{\#\Lambda(\mathbf{n})(l^\nu(s(\psi)+\eta)+\delta)}) \qquad (\text{by } (2.24)) \\
&\quad + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}(\mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})})) \\
&\leq \log(\# A_{l,\eta}^c) + \#\Lambda(\mathbf{n})(l^\nu(s(\psi)+\eta)+\delta) \\
&\quad + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}(\mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})})) \\
&\leq \log(\# A_{l,\eta}^c) + \#\Lambda(l\mathbf{n})(s(\psi)+\eta+\frac{\delta}{l^\nu}) \\
&\quad + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}(\mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})})).
\end{aligned}$$

We can conclude from this that

$$\limsup_{\Lambda(\mathbf{y})\nearrow\mathbb{N}^\nu} \frac{1}{\#(\Lambda(\mathbf{y}))} \beta_{\varepsilon,\mathbf{y}}(\psi) \leq s(\psi) + \eta + \frac{\delta}{l^\nu},$$

because $\# A_{l,\eta}^c$ does not depend on $\mathbf{n}$ and $\Lambda(\mathbf{y}) \nearrow \mathbb{N}^\nu$ if and only if $\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu$. This leads to

$$\limsup_{\Lambda(\mathbf{y})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{y})} \beta_{\varepsilon,\mathbf{y}}(\psi) \leq s(\psi),$$

since η, $\delta > 0$ were chosen arbitrarily. $\square$

Let $\nu \in \mathbb{N}$. For $\mathbf{n} = (n_1, \ldots, n_\nu) \in \mathbb{N}^\nu$ we define $|\mathbf{n}| := \prod_{i=1}^{\nu} n_i$ and write $\mathbf{n} \to \infty$ alternatively for $\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu$. Further we introduce the notation

$$\mathbf{n} \geq \mathbf{m} \;:\Longleftrightarrow\; n_i \geq m_i, \quad \forall i \in \{1, \ldots, \nu\}.$$

Recall that for a probability distribution P on a finite set A the Shannon entropy is defined by

$$H(P) := -\sum_{a \in A} P(a) \log P(a).$$

Lemma 2.3.5 *Let $D > 0$ and $\{(A^{(\mathbf{n})}, P^{(\mathbf{n})})\}_{\mathbf{n} \in \mathbb{N}^\nu}$ be a family, where each $A^{(\mathbf{n})}$ is a finite set with $\frac{1}{|\mathbf{n}|} \log \# A^{(\mathbf{n})} \leq D$ for all $\mathbf{n} \in \mathbb{N}^\nu$ and $P^{(\mathbf{n})}$ is a probability distribution on $A^{(\mathbf{n})}$. Define*

$$\alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})}) := \min\{\log \# \Omega : \; \Omega \subset A^{(\mathbf{n})}, P^{(\mathbf{n})}(\Omega) \geq 1 - \varepsilon\}. \tag{2.26}$$

If $\{(A^{(\mathbf{n})}, P^{(\mathbf{n})})\}_{\mathbf{n} \in \mathbb{N}^\nu}$ satisfies the following two conditions:

1. $\lim_{\mathbf{n} \to \infty} \frac{1}{|\mathbf{n}|} H(P^{(\mathbf{n})}) = h < \infty$
2. $\limsup_{\mathbf{n} \to \infty} \frac{1}{|\mathbf{n}|} \alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})}) \leq h, \qquad \forall \varepsilon \in (0, 1)$

then for every $\varepsilon \in (0, 1)$

$$\lim_{\mathbf{n} \to \infty} \frac{1}{|\mathbf{n}|} \alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})}) = h. \tag{2.27}$$

Note that we do not expect either $\{A^{(\mathbf{n})}\}_{\mathbf{n} \in \mathbb{N}^\nu}$ or $\{P^{(\mathbf{n})}\}_{\mathbf{n} \in \mathbb{N}^\nu}$ to fulfill any consistency conditions. We will see later on that this is the important point for why Lemma 2.3.5 will be useful in the non-commutative setting.

Proof of Lemma 2.3.5: Let $\delta, \sigma > 0$ and define

$$\begin{aligned} A_1^{(\mathbf{n})}(\sigma) &:= \left\{ a \in A^{(\mathbf{n})} : P^{(\mathbf{n})}(a) > e^{-|\mathbf{n}|(h - \sigma)} \right\}, \\ A_2^{(\mathbf{n})}(\delta, \sigma) &:= \left\{ a \in A^{(\mathbf{n})} : e^{-|\mathbf{n}|(h + \delta)} \leq P^{(\mathbf{n})}(a) \leq e^{-|\mathbf{n}|(h - \sigma)} \right\}, \\ A_3^{(\mathbf{n})}(\delta) &:= \left\{ a \in A^{(\mathbf{n})} : P^{(\mathbf{n})}(a) < e^{-|\mathbf{n}|(h + \delta)} \right\}. \end{aligned}$$

To see that $\lim_{\mathbf{n} \to \infty} P^{(\mathbf{n})}\left(A_3^{(\mathbf{n})}(\delta)\right) = 0$ assume the contrary and observe that then there would exist infinitely many $\mathbf{n} \in \mathbb{N}^\nu$ such that $P^{(\mathbf{n})}(A_3^{(\mathbf{n})}(\delta_0)) > \varepsilon$, for some $\delta_0 > 0$ and some $\varepsilon \in (0, 1)$. Consider the subset $\Omega_{\frac{\varepsilon}{2}, \mathbf{n}}$ of $A^{(\mathbf{n})}$ with

$\Omega_{\frac{\varepsilon}{2},\mathbf{n}} = \arg\min \alpha_{\frac{\varepsilon}{2},\mathbf{n}}(P^{(\mathbf{n})})$. Then we would have $P^{(\mathbf{n})}(A_3^{(\mathbf{n})}(\delta_0) \cap \Omega_{\frac{\varepsilon}{2},\mathbf{n}}) \geq \frac{\varepsilon}{2}$ and hence

$$\#\Omega_{\frac{\varepsilon}{2},\mathbf{n}} \geq \#(A_3^{(\mathbf{n})}(\delta_0) \cap \Omega_{\frac{\varepsilon}{2},\mathbf{n}}) \geq \frac{\varepsilon}{2} \cdot e^{|\mathbf{n}|(h+\delta_0)},$$

for infinitely many $\mathbf{n} \in \mathbb{N}^\nu$, which contradicts condition 2 in the lemma. Furthermore the set $A_3^{(\mathbf{n})}(\delta)$ cannot asymptotically contribute to the mean entropy h since

$$\begin{aligned} & -\frac{1}{|\mathbf{n}|} \sum_{a \in A_3^{(\mathbf{n})}(\delta)} P^{(\mathbf{n})}(a) \log P^{(\mathbf{n})}(a) \\ \leq & -\frac{1}{|\mathbf{n}|} \sum_{a \in A_3^{(\mathbf{n})}(\delta)} P^{(\mathbf{n})}(a) \log \frac{1}{\#A_3^{(\mathbf{n})}(\delta)} P^{\mathbf{n}}(A_3^{\mathbf{n}}(\delta)) \end{aligned}$$

and

$$\begin{aligned} & \lim_{\mathbf{n}\to\infty} -\frac{1}{|\mathbf{n}|} \sum_{a \in A_3^{(\mathbf{n})}(\delta)} P^{(\mathbf{n})}(a) \log \frac{1}{\#A_3^{(\mathbf{n})}(\delta)} P^{(\mathbf{n})}(A_3^{(\mathbf{n})}(\delta)) \\ = & \lim_{\mathbf{n}\to\infty} \left(P^{(\mathbf{n})}(A_3^{(\mathbf{n})}(\delta)) \frac{\log \#A_3^{(\mathbf{n})}(\delta)}{|\mathbf{n}|} - \frac{P^{(\mathbf{n})}(A_3^{(\mathbf{n})}(\delta)) \log P^{(\mathbf{n})}(A_3^{(\mathbf{n})}(\delta))}{|\mathbf{n}|} \right) \\ = & \ 0. \end{aligned}$$

Here we used the fact that $\frac{\log \#A^{(\mathbf{n})}}{|\mathbf{n}|}$ stays bounded and $-\sum p_i \log p_i \leq -\sum p_i \log q_i$ for finite vectors $(p_i), (q_i)$ with $\sum_i p_i = \sum_i q_i \leq 1$ and $p_i, q_i \geq 0$. Since $A_3^{(\mathbf{n})}(\delta)$ does not contribute to the entropy one easily concludes that $\lim_{\mathbf{n}\to\infty} P^{(\mathbf{n})}\left(A_1^{(\mathbf{n})}(\sigma)\right) = 0$ for all $\sigma > 0$ because otherwise $\liminf_{\mathbf{n}\to\infty} \frac{1}{|\mathbf{n}|} H\left(P^{(\mathbf{n})}\right) < h$ would hold. Recall that $\delta, \sigma > 0$ were chosen arbitrarily.
Thus setting $\delta = \sigma$ we obtain

$$\lim_{\mathbf{n}\to\infty} P^{(\mathbf{n})}\left(A_2^{(\mathbf{n})}(\delta)\right) = 1, \qquad \forall \delta > 0. \tag{2.28}$$

Consequently the lemma follows since $P^{(\mathbf{n})}(\Omega) \geq 1 - \varepsilon$ implies $P^{(\mathbf{n})}\left(\Omega \cap A_2^{(\mathbf{n})}(\delta)\right) \geq (1-\varepsilon)^2$ for $|\mathbf{n}|$ sufficiently large and one needs at least $(1-\varepsilon)^2 \cdot e^{|\mathbf{n}|(h-\delta)}$ elements from $A_2^{(\mathbf{n})}(\delta)$ to cover $\Omega \cap A_2^{(\mathbf{n})}(\delta)$ and δ can be chosen arbitrarily small. $\square$

Proof of Proposition 2.3.2: $\mathcal{A}^{(\mathbf{n})}$ as a finite dimensional C^*-algebra is isomorphic to a finite direct sum $\bigoplus_{j=1}^M \mathcal{B}(\mathcal{H}_j^{(\mathbf{n})})$, where each $\mathcal{H}_j^{(\mathbf{n})}$ is a Hilbert space with $\dim \mathcal{H}_j^{(\mathbf{n})} = d_j^{(\mathbf{n})} < \infty$ and any minimal projection in $\mathcal{A}^{(\mathbf{n})}$ is represented by a one-dimensional projection on $\mathcal{H}^{(\mathbf{n})} := \bigoplus_{j=1}^M \mathcal{H}_j^{(\mathbf{n})}$

with $\dim \mathcal{H}^{(\mathbf{n})} = \sum_{j=1}^{M} d_j^{(\mathbf{n})} =: d_{\mathbf{n}}$. Note that $\bigoplus_{j=1}^{M} \mathcal{B}(\mathcal{H}_j^{(\mathbf{n})}) \subset \mathcal{B}(\mathcal{H}^{(\mathbf{n})})$. Consider the spectral representation of the density operator $D_{\mathbf{n}}$ of $\psi^{(\mathbf{n})}$ in $\mathcal{B}(\mathcal{H}^{(\mathbf{n})})$:

$$D_{\mathbf{n}} = \sum_{i=1}^{d_{\mathbf{n}}} \lambda_i^{(\mathbf{n})} q_i^{(\mathbf{n})},$$

where each $q_i^{(\mathbf{n})}$ is an one-dimensional projection. For $\mathbf{n} = (n_1, \dots, n_\nu) \in \mathbb{N}^\nu$ let $A^{(\mathbf{n})}$ be the finite set consisting of these eigenprojections $q_i^{(\mathbf{n})}$ of $\psi^{(\mathbf{n})}$, i.e.

$$A^{(\mathbf{n})} := \{q_i^{(\mathbf{n})}\}_{i=1}^{d_{\mathbf{n}}}. \tag{2.29}$$

Let $P^{(\mathbf{n})}$ be the probability distribution on $A^{(\mathbf{n})}$ given by:

$$P^{(\mathbf{n})}(q_i^{(\mathbf{n})}) := \psi^{(\mathbf{n})}(q_i^{(\mathbf{n})}) = \lambda_i^{(\mathbf{n})}. \tag{2.30}$$

Recall that $|\mathbf{n}| = \prod_{i=1}^{\nu} n_i$. Let $D := \log(\dim \mathcal{H}^{(0)})$, then $\frac{1}{|\mathbf{n}|} \log \# A^{(\mathbf{n})} \leq D$ for all $\mathbf{n} \in \mathbb{N}^\nu$. We show that the family $\{(A^{(\mathbf{n})}, P^{(\mathbf{n})})\}_{\mathbf{n} \in \mathbb{N}^\nu}$ fulfills both conditions in Lemma 2.3.5 and consequently

$$\lim_{\mathbf{n} \to \infty} \frac{1}{|\mathbf{n}|} \alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})}) = \lim_{\mathbf{n} \to \infty} \frac{1}{|\mathbf{n}|} H(P^{(\mathbf{n})}), \qquad \forall \varepsilon \in (0,1). \tag{2.31}$$

It is clear that $H(P^{(\mathbf{n})}) = -\sum_{i=1}^{d_{\mathbf{n}}} \lambda_i^{(\mathbf{n})} \log \lambda_i^{(\mathbf{n})} = S(\psi^{(\mathbf{n})})$. Thus

$$h := \lim_{\mathbf{n} \to \infty} \frac{1}{|\mathbf{n}|} H(P^{(\mathbf{n})}) = s(\psi). \tag{2.32}$$

We assume that for each $\mathbf{n} \in \mathbb{N}^\nu$ the eigenvalues are listed in decreasing order of magnitude, i.e.

$$\lambda_1^{(\mathbf{n})} \geq \lambda_2^{(\mathbf{n})} \geq \ldots \geq \lambda_{d_{\mathbf{n}}}^{(\mathbf{n})},$$

and define for $\varepsilon \in (0,1)$

$$k_{\varepsilon, \mathbf{n}} := \min\{k \in \{1, \dots, d_{\mathbf{n}}\} : \sum_{j=1}^{k} \lambda_j^{(\mathbf{n})} \geq 1 - \varepsilon\}.$$

Thus $\alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})}) = \log \#(\{q_i^{(\mathbf{n})}\}_{i=1}^{k_{\varepsilon, \mathbf{n}}}) = \log k_{\varepsilon, \mathbf{n}}$. We claim :

$$\alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})}) = \beta_{\varepsilon, \mathbf{n}}(\psi^{(\mathbf{n})}), \qquad \forall \varepsilon \in (0,1). \tag{2.33}$$

From $\psi^{(\mathbf{n})}(\sum_{i=1}^{k_{\varepsilon, \mathbf{n}}} q_i^{(\mathbf{n})}) \geq 1 - \varepsilon$ and $\mathrm{tr}(\sum_{i=1}^{k_{\varepsilon, \mathbf{n}}} q_i^{(\mathbf{n})}) = k_{\varepsilon, \mathbf{n}}$ it is obvious that $\beta_{\varepsilon, \mathbf{n}}(\psi^{(\mathbf{n})}) \leq \alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})})$.
Assume $\beta_{\varepsilon, \mathbf{n}}(\psi^{(\mathbf{n})}) < \alpha_{\varepsilon, \mathbf{n}}(P^{(\mathbf{n})})$. Then there exists a projection $q \in \mathcal{A}^{(\mathbf{n})}$

with $\psi^{(\mathbf{n})}(q) \geq 1-\varepsilon$ such that $m := \mathrm{tr}(q) < k_{\varepsilon,\mathbf{n}}$. Applying Ky Fan's maximum principle (cf. [1]) to the density operator $D_{\mathbf{n}}$ we infer the following contradiction:

$$1-\varepsilon \leq \psi^{(\mathbf{n})}(q) = \mathrm{tr}(D_{\mathbf{n}} q) \leq \sum_{i=1}^{m} \lambda_i^{(\mathbf{n})} < 1-\varepsilon.$$

ψ is ergodic. Thus we can apply Lemma 2.3.4:

$$\limsup_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#(\Lambda(\mathbf{n}))}\beta_{\varepsilon,\mathbf{n}}(\psi^{(\mathbf{n})}) \leq s(\psi), \qquad \forall\varepsilon\in(0,1). \tag{2.34}$$

Taking into account (2.33), (2.32), (2.34) and using that $\#\Lambda(\mathbf{n}) = |\mathbf{n}|$ we obtain

$$\limsup_{\mathbf{n}\to\infty} \frac{1}{|\mathbf{n}|}\alpha_{\varepsilon,\mathbf{n}}(P^{(\mathbf{n})}) \leq h, \qquad \forall\varepsilon\in(0,1). \tag{2.35}$$

With (2.32) and (2.35) both conditions in Lemma 2.3.5 are satisfied. It follows that (2.31) holds. Finally, the relations (2.33), (2.32) and (2.31) lead to

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#(\Lambda(\mathbf{n}))}\beta_{\varepsilon,\mathbf{n}}(\psi) = s(\psi), \qquad \forall\varepsilon\in(0,1). \qquad \square$$

Proof of the Quantum Shannon-McMillan Theorem:
Fix $\delta > 0$. We continue to use the family $\{(A^{(\mathbf{n})}, P^{(\mathbf{n})})\}_{\mathbf{n}\in\mathbb{N}^\nu}$ and further notations from the proof of Proposition 2.3.2. Choose some $\delta' < \delta$. Let $A_2^{(\mathbf{n})}(\delta')$ be the subset of $A^{(\mathbf{n})}$ defined in the proof of Lemma 2.3.5 with $h = s(\psi)$ (cf. (2.32)). Let $I_{\mathbf{n}}(\delta') := \{i \in \{1,\ldots,d_{\mathbf{n}}\} :\ q_i^{\mathbf{n}} \in A_2^{(\mathbf{n})}(\delta')\}$. Set

$$p_{\mathbf{n}}(\delta) = \sum_{I_{\mathbf{n}}(\delta')} q_i^{\mathbf{n}}.$$

By (2.28) there exists an $\mathbf{N}_\delta \in \mathbb{N}^\nu$ such that $p_{\mathbf{n}}(\delta)$ is a projection with

$$\psi^{(\mathbf{n})}(p_{\mathbf{n}}(\delta)) = P^{(\mathbf{n})}(A_2^{(\mathbf{n})}(\delta')) \geq 1-\delta, \qquad \forall \mathbf{n} \geq \mathbf{N}_\delta.$$

Any minimal projection $p \in \mathcal{A}^{(\mathbf{n})}$ with $p \leq p_{\mathbf{n}}(\delta)$ is represented as a one-dimensional projection $p_f, f \in \mathcal{H}^{(\mathbf{n})}, \|f\| = 1$, such that $f = \sum_{i\in I_{\mathbf{n}}(\delta')} \gamma_i e_i^{\mathbf{n}}$, $\sum_{i\in I_{\mathbf{n}}(\delta')} |\gamma_i|^2 = 1$, where each $e_i^{\mathbf{n}}$ is a normalized vector in the range of the one-dimensional projections $q_i^{\mathbf{n}}$. Hence

$$\psi^{(\mathbf{n})}(p) = \sum_{i\in I_{\mathbf{n}}(\delta')} |\gamma_i|^2 \lambda_i^{(\mathbf{n})}$$

is a weighted average of the eigenvalues $\lambda_i^{(\mathbf{n})}$ corresponding to the set $A_2^{(\mathbf{n})}(\delta')$. Thus we obtain by the definition of this set

$$e^{-\#\Lambda(\mathbf{n})(s(\psi)+\delta)} < \psi^{(\mathbf{n})}(p) < e^{-\#\Lambda(\mathbf{n})(s(\psi)-\delta)}. \tag{2.36}$$

Using the linearity of $\psi^{(\mathbf{n})}$ and applying (2.36) to the projections $q_i^{(\mathbf{n})}$ we arrive at the following estimation

$$e^{\#\Lambda(\mathbf{n})(s(\psi)-\delta)} < \operatorname{tr}(p_{\mathbf{n}}(\delta)) < e^{\#\Lambda(\mathbf{n})(s(\psi)+\delta)},$$

if $\mathbf{n}$ is large enough. We have shown all assertions of the theorem. □

2.4 Concluding Remarks

As mentioned in section 2.3, Proposition 2.3.2 should be seen just as a reformulation of the Shannon-McMillan theorem suited for application to problems of data compression. In fact, using the results of this chapter it had been shown by A. Szkoła and the author in [6] that there exist asymptotically error-free compression schemes for ergodic quantum sources at the rate which is given by the mean von Neumann entropy. Moreover it had been shown there that compression at a rate below mean von Neumann entropy results in a decoding error (defined by an appropriate notion of fidelity) approaching 1. This extends the quantum data compression results of Jozsa and Schumacher in [11], Petz and Mosonyi in [21] to ergodic sources. On the other hand Kaltchenko and Yang [12] could show, on the basis of the results and techniques presented in this chapter, the existence of projections typical for all ergodic states the mean von Neumann entropy of which does not exceed some prescribed number r. This in turn demonstrates the possibility of universal quantum data compression at rate r for all sources with mean von Neumann entropy below r. These compression results are theoretical in spirit and far from being algorithmically implemented.

Note that there is a sharper version of the quantum Shannon-McMillan theorem which reduces to the famous Shannon-McMillan-Breiman theorem in the classical case. The main idea behind this theorem is that the pointwise (almost-sure) convergence can be reformulated in a finitary way: The Shannon-McMillan theorem asserts the convergence of the normalized logarithms of individual probabilities to $-h$ in probability, where h denotes mean entropy. Breiman's sharpening is obtained if we are able to show that the entropy typical subsets, obtained in Shannon-McMillan theorem, can be chained in the sense that projecting the entropy typical set of the block of length $n+1$ onto the first n coordinates produces the entropy typical set of the n-block. It can be shown that Breiman's theorem is equivalent to the existence of chained high probability subsets fulfilling only the upper estimate of the AEP concerning probabilities of the minimal typical projections while having cardinalities of the order e^{nh} of typical n-blocks, where h denotes the entropy of the process. This reformulation is suited for extending of Breiman's theorem to the quantum situation where the notion of the realization (or trajectory) of the process does not make sense. A quantum version of Breiman's theorem was proved by T. Krüger, Ra. Siegmund-Schultze, A Szkoła and the author in [3]

Chapter 3

Limit Theorems for the Quantum Relative Entropy

In this chapter we shall give a proof of the Hiai-Petz conjecture we described in the introduction to chapter 2. Moreover from our proof of that conjecture we will infer the validity of the *asymptotic equipartition property* for the mean quantum relative entropy (quantum relative AEP) containing the quantum Shannon-McMillan theorem as a special case. The results in this chapter are from the paper [4] by Ra. Siegmund-Schultze and the author.

3.1 Quantum Stein's Lemma and the Relative AEP

The basic quantity in this section is given by

$$\beta_{\varepsilon,\mathbf{n}}(\psi,\varphi) := \min\{\log \varphi^{(\mathbf{n})}(q) : q \in \mathcal{A}^{(\mathbf{n})} \text{ projection}, \ \psi^{(\mathbf{n})}(q) \geq 1-\varepsilon\}, \quad (3.1)$$

where $\varepsilon \in (0,1)$ and $\psi, \varphi \in \mathcal{T}(\mathcal{A}^{\infty})$. The quantities appearing in (3.1) have a natural interpretation from the point of view of quantum statistical hypothesis testing. Assume that under the hypothesis H_1 the state of the system is given by $\psi^{(\mathbf{n})}$ and under the hypothesis H_2 the state is given by $\varphi^{(\mathbf{n})}$. A projection $q \in \mathcal{A}^{(\mathbf{n})}$ can be thought of as a decision rule: If a measurement of the projection q on the system under consideration has outcome 1, then the hypothesis H_1 is accepted and the measurement outcome 0 implies that H_2 is accepted. The quantities $\psi^{(\mathbf{n})}(1-q)$ and $\varphi^{(\mathbf{n})}(q)$ respectively describe the probabilities for the occurrence of the eigenvalue 0 respectively 1 if the hypothesis H_1 respectively H_2 was accepted, i.e. the error probabilities. The following theorem had been conjectured by Hiai and Petz in [8].

Theorem 3.1.1 *Let ψ be an ergodic state and let φ be a stationary product state on $\mathcal{A}^{\infty}$ with mean relative entropy $s(\psi,\varphi) < \infty$. Then for all $\varepsilon \in (0,1)$*

we have

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})}\beta_{\varepsilon,\mathbf{n}}(\psi,\varphi) = -s(\psi,\varphi).$$

The fact that in the case where both states are stationary product states the limit of $\frac{1}{\#\Lambda(\mathbf{n})}\beta_{\varepsilon,\mathbf{n}}(\psi,\varphi)$ exists and equals $-s(\psi,\varphi)$, had been shown by Ogawa and Nagaoka in [19]. Their proof requires some deeper properties of so called quasi-entropies which are quantities defined analogously to the relative entropy where an arbitrary operator convex function is used in their definition instead of the function $-\log x$. We shall derive now the existence of the limit in the general case. As an instrument we will use some abelian approximation of all relevant quantities.

Projections q in (3.1) for which $\varphi^{(\mathbf{n})}(q)$ is of the order $e^{-\#\Lambda(\mathbf{n})s(\psi,\varphi)}$ and for which $\psi(q) \geq 1-\varepsilon$ could be called *maximally separating projections* and their range could be denoted as *maximally separating subspace*. The proof of this theorem reveals some additional information about maximally separating projections, which is collected in the following

Theorem 3.1.2 (quantum relative AEP) *Let ψ be an ergodic state with mean entropy $s(\psi)$ and let φ be a stationary product state on $\mathcal{A}^\infty$ with mean relative entropy $s(\psi,\varphi) < \infty$. Then for all $\varepsilon > 0$ there is an $\mathbf{N}_\varepsilon \in \mathbb{N}^\nu$ such that for all $\mathbf{n} \in \mathbb{N}^\nu$ with $\Lambda(\mathbf{n}) \supseteq \Lambda(\mathbf{N}_\varepsilon)$ there exists an orthogonal projection $p_{\mathbf{n}}(\varepsilon) \in \mathcal{A}^{(\mathbf{n})}$ such that*

1. $\psi^{(\mathbf{n})}(p_{\mathbf{n}}(\varepsilon)) \geq 1-\varepsilon$,

2. *for all minimal projections $p \in \mathcal{A}^{(\mathbf{n})}$ with $p \leq p_{\mathbf{n}}(\varepsilon)$ we have*

$$e^{-\#(\Lambda(\mathbf{n}))(s(\psi)+\varepsilon)} < \psi^{(\mathbf{n})}(p) < e^{-\#(\Lambda(\mathbf{n}))(s(\psi)-\varepsilon)},$$

and consequently $e^{\#(\Lambda(\mathbf{n}))(s(\psi)-\varepsilon)} < \mathrm{tr}_{\mathbf{n}}(p_{\mathbf{n}}(\varepsilon)) < e^{\#(\Lambda(\mathbf{n}))(s(\psi)+\varepsilon)}$.

3. *for all minimal projections $p \in \mathcal{A}^{(\mathbf{n})}$ with $p \leq p_{\mathbf{n}}(\varepsilon)$ we have*

$$e^{-\#(\Lambda(\mathbf{n}))(s(\psi)+s(\psi,\varphi)+\varepsilon)} < \varphi^{(\mathbf{n})}(p) < e^{-\#(\Lambda(\mathbf{n}))(s(\psi)+s(\psi,\varphi)-\varepsilon)},$$

and consequently

$$e^{-\#(\Lambda(\mathbf{n}))(s(\psi,\varphi)+\varepsilon)} < \varphi^{(\mathbf{n})}(p_{\mathbf{n}}(\varepsilon)) < e^{-\#(\Lambda(\mathbf{n}))(s(\psi,\varphi)-\varepsilon)}.$$

In the case that the state φ is the tracial state on $\mathcal{A}^\infty$ Theorem 3.1.2 above is equivalent to the quantum version of the Shannon-McMillan theorem proved in chapter 2 of this thesis.

The proof of Theorem 3.1.1 will make use of a classical law of large numbers for the (classical) mean relative entropy.

Theorem 3.1.3 *Let A be a finite set and P respectively Q be an ergodic respectively an i.i.d. probability measure on $[A^{\mathbb{Z}^\nu}, \mathfrak{A}^{\mathbb{Z}^\nu}]$, where $\mathfrak{A}^{\mathbb{Z}^\nu}$ is the σ-algebra generated by the cylinder sets. We have*

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} \log \frac{P^{(\mathbf{n})}(\omega_{\mathbf{n}})}{Q^{(\mathbf{n})}(\omega_{\mathbf{n}})} = D_M(P,Q) \qquad P-\text{almost surely}, \tag{3.2}$$

where $D_M(P,Q)$ denotes the mean relative entropy of P with respect to Q and $\omega_{\mathbf{n}} \in A^{\Lambda(\mathbf{n})}$ are the components of $\omega \in A^{\mathbb{Z}^\nu}$ corresponding to the box $\Lambda(\mathbf{n})$.

The proof of this classical assertion is an elementary application of the Shannon-McMillan-Breiman theorem and the individual ergodic theorem. The higher dimensional versions of these theorems needed in the present situation can be found in the article [20] by Ornstein and Weiss. A full proof of Theorem 3.1.3 is given in appendix B.
Convention: Due to our assumption $s(\psi,\varphi) < \infty$ we suppose w.l.o.g. in the following proofs that the state $\varphi^{(1)}$ is faithful.

3.1.1 Proof of the Quantum Stein's Lemma

An important ingredient in our proof of the Theorem 3.1.1 is a result proved by Hiai and Petz in [8]. The starting point is the spectral decomposition of the density operator $D_{\varphi^{(1)}}$ corresponding to the state $\varphi^{(1)}$ on $\mathcal{A}$:

$$D_{\varphi^{(1)}} = \sum_{i=1}^{d} \lambda_i e_i,$$

where e_i are one-dimensional projections. Clearly, the spectral representation of the tensor product of $D_{\varphi^{(1)}}$ over a box $\Lambda(\mathbf{y})$ can be written as

$$D_{\varphi^{(\mathbf{y})}} = \sum_{i_1,\ldots,i_{N(\mathbf{y})}=1}^{d} \lambda_{i_1}\cdots\lambda_{i_{N(\mathbf{y})}} e_{i_1}\otimes\ldots\otimes e_{i_{N(\mathbf{y})}},$$

where we have chosen some enumeration $\{1,\ldots,N(\mathbf{y})\}$ of the points belonging to the box $\Lambda(\mathbf{y})$. We have $N(\mathbf{y}) = y_1\cdot\ldots\cdot y_\nu$. Collecting all one dimensional projections $e_{i_1}\otimes\ldots\otimes e_{i_{N(\mathbf{y})}}$ which correspond to the same eigenvalue of $D_{\varphi^{(\mathbf{y})}}$ we can rewrite the last expression as

$$D_{\varphi^{(\mathbf{y})}} = \sum_{n_1,\ldots n_d : n_1+\ldots+n_d = N(\mathbf{y})} \left(\prod_{k=1}^{d} \lambda_k^{n_k}\right) p_{n_1,\ldots n_d}, \tag{3.3}$$

with

$$p_{n_1,\ldots n_d} := \sum_{(i_1,\ldots,i_{N(\mathbf{y})})\in I_{n_1,\ldots n_d}} e_{i_1}\otimes\ldots\otimes e_{i_{N(\mathbf{y})}},$$

where

$$I_{n_1,\ldots n_d} := \{(i_1,\ldots,i_{N(\mathbf{y})}) : \#\{j : i_j = k\} = n_k \text{ for } 1 \le k \le d\}.$$

We define the conditional expectation with respect to the trace by

$$E_{\mathbf{y}} : \mathcal{A}^{(\mathbf{y})} \to \bigoplus_{n_1,\ldots,n_d:\ n_1+\ldots+n_d=N(\mathbf{y})} p_{n_1,\ldots,n_d}\mathcal{A}^{(\mathbf{y})}p_{n_1,\ldots,n_d},$$

$$E_{\mathbf{y}}(a) := \sum_{n_1,\ldots,n_d:\ n_1+\ldots+n_d=N(\mathbf{y})} p_{n_1,\ldots,n_d}a\ p_{n_1,\ldots,n_d}. \tag{3.4}$$

We are prepared to state the announced important result of Hiai and Petz, Lemma 3.1 and Lemma 3.2 in [8].

Theorem 3.1.4 *If ψ is a stationary state on $\mathcal{A}^\infty$ and $\mathcal{D}_{\mathbf{y}}$ is the abelian subalgebra of $\mathcal{A}^{(\mathbf{y})}$ generated by $\{p_{n_1,\ldots n_d}D_{\psi^{(\mathbf{y})}}p_{n_1,\ldots n_d}\}_{n_1,\ldots,n_d} \cup \{p_{n_1,\ldots n_d}\}_{n_1,\ldots,n_d}$ then*

$$S(\psi^{(\mathbf{y})},\varphi^{(\mathbf{y})}) = S(\psi^{(\mathbf{y})} \upharpoonright \mathcal{D}_{\mathbf{y}}, \varphi^{(\mathbf{y})} \upharpoonright \mathcal{D}_{\mathbf{y}}) + S(\psi^{(\mathbf{y})} \circ E_{\mathbf{y}}) - S(\psi^{(\mathbf{y})}),$$

and

$$S(\psi^{(\mathbf{y})} \circ E_{\mathbf{y}}) - S(\psi^{(\mathbf{y})}) \le d\log(\#\Lambda(\mathbf{y}) + 1). \tag{3.5}$$

Consequently we have

$$\lim_{\Lambda(\mathbf{y})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{y})} S(\psi^{(\mathbf{y})} \upharpoonright \mathcal{D}_{\mathbf{y}}, \varphi^{(\mathbf{y})} \upharpoonright \mathcal{D}_{\mathbf{y}}) = s(\psi,\varphi). \tag{3.6}$$

Remark This theorem had been proved by Hiai and Petz in [8] for the one-dimensional lattice. A simplified proof appeared in the work [5] by Ra. Siegmund-Schultze and this author which is reproduced in appendix A. However, both proofs extend canonically to the present situation.
Any abelian algebra $\mathcal{D}_{\mathbf{y}}$ in Theorem 3.1.4 can be represented as

$$\mathcal{D}_{\mathbf{y}} = \bigoplus_{i=1}^{a_{\mathbf{y}}} \mathbb{C}\cdot f_{\mathbf{y},i} \tag{3.7}$$

where $\{f_{\mathbf{y},i}\}_{i=1}^{a_{\mathbf{y}}}$ is the set of orthogonal minimal projections in $\mathcal{D}_{\mathbf{y}}$. For any $\mathbf{y}$ we introduce a maximally abelian refinement $\mathcal{B}_{\mathbf{y}}$ of $\mathcal{D}_{\mathbf{y}}$ by splitting each $f_{\mathbf{y},i}$ into a sum of orthogonal and minimal (in the sense of the algebra $\mathcal{A}^{(\mathbf{y})}$) projections $g_{\mathbf{y},i,j}$ which leads to the representation

$$\mathcal{B}_{\mathbf{y}} = \bigoplus_{i=1}^{a_{\mathbf{y}}}\bigoplus_{j=1}^{b_{\mathbf{y},i}} \mathbb{C}\cdot g_{\mathbf{y},i,j}. \tag{3.8}$$

By the monotonicity of the relative entropy we get

$$S(\psi^{(\mathbf{y})} \upharpoonright \mathcal{D}_{\mathbf{y}}, \varphi^{(\mathbf{y})} \upharpoonright \mathcal{D}_{\mathbf{y}}) \leq S(\psi^{(\mathbf{y})} \upharpoonright \mathcal{B}_{\mathbf{y}}, \varphi^{(\mathbf{y})} \upharpoonright \mathcal{B}_{\mathbf{y}}) \leq S(\psi^{(\mathbf{y})}, \varphi^{(\mathbf{y})}), \quad (3.9)$$

from which we deduce

$$\lim_{\Lambda(\mathbf{y}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{y})} S(\psi^{(\mathbf{y})} \upharpoonright \mathcal{B}_{\mathbf{y}}, \varphi^{(\mathbf{y})} \upharpoonright \mathcal{B}_{\mathbf{y}}) = s(\psi, \varphi), \quad (3.10)$$

by (3.6).
We choose a positive integer l and consider the decomposition of ψ into states $\psi_{\mathbf{x}}$ being ergodic with respect to the action of G_l, i.e. $\psi = \frac{1}{\#\Lambda(\mathbf{k}(l))} \sum_{\mathbf{x} \in \Lambda(\mathbf{k}(l))} \psi_{\mathbf{x}}$ in accordance with Theorem 2.2.1. Moreover, we consider a stationary product state φ. Note that this state is G_l-ergodic for each $l \in \mathbb{Z}$. In order to keep our notation transparent we agree on the following abbreviations:

$$s := s(\psi, \varphi, \mathbb{Z}^\nu) = s(\psi, \varphi),$$

i.e. the mean relative entropy of the state ψ computed with respect to $\mathbb{Z}^\nu$. We write $\mathcal{B}_l$ for $\mathcal{B}_{(l,l,\ldots,l)}$ and set

$$s_{\mathbf{x}}^{(l)} := \frac{1}{\#\Lambda(l)} S(\psi_{\mathbf{x}}^{(\Lambda(l))} \upharpoonright \mathcal{B}_l, \varphi^{(\Lambda(l))} \upharpoonright \mathcal{B}_l)$$

and

$$s^{(l)} := \frac{1}{\#\Lambda(l)} S(\psi^{(\Lambda(l))} \upharpoonright \mathcal{B}_l, \varphi^{(\Lambda(l))} \upharpoonright \mathcal{B}_l).$$

From the Theorem 2.2.1 we know that

$$s(\psi_{\mathbf{x}}, \varphi, G_l) = s(\psi, \varphi, G_l) = l^\nu \cdot s(\psi, \varphi), \qquad \forall \mathbf{x} \in \Lambda(\mathbf{k}(l)). \quad (3.11)$$

For an arbitrarily chosen $\eta > 0$ let us define the set

$$A_{l,\eta} := \{\mathbf{x} \in \Lambda(\mathbf{k}(l)) : \ s_{\mathbf{x}}^{(l)} < s - \eta\}. \quad (3.12)$$

By $A_{l,\eta}^c$ we denote its complement. In the next lemma we shall show that the essential part of G_l-ergodic components of ψ have the entropy per site close to s as l becomes large, even if we restrict them to the abelian algebras $\mathcal{B}_l$.

Lemma 3.1.5 *If ψ is a $\mathbb{Z}^\nu$-ergodic state and φ a stationary product state on $\mathcal{A}^\infty$ and if $s(\psi, \varphi) < \infty$, then*

$$\lim_{l \to \infty} \frac{\#A_{l,\eta}}{\#\Lambda(\mathbf{k}(l))} = 0$$

holds for every $\eta > 0$.

Proof of Lemma 3.1.5: We suppose, on the contrary, that there exists some $\eta_0 > 0$ such that $\limsup_l \frac{\# A_{l,\eta_0}}{\# \Lambda(\mathbf{k}(l))} = a > 0$. Then we can select a subsequence, which we denote again by (l) for simplicity, with the property

$$\lim_{l \to \infty} \frac{\# A_{l,\eta_0}}{\# \Lambda(\mathbf{k}(l))} = a.$$

Using joint convexity of the relative entropy we have the following estimate:

$$\begin{aligned} \# \Lambda(\mathbf{k}(l)) \cdot s^{(l)} &\leq \sum_{\mathbf{x} \in \Lambda(\mathbf{k}(l))} s_{\mathbf{x}}^{(l)} \\ &= \sum_{\mathbf{x} \in A_{l,\eta_0}} s_{\mathbf{x}}^{(l)} + \sum_{\mathbf{x} \in A_{l,\eta_0}^c} s_{\mathbf{x}}^{(l)} \\ &< \# A_{l,\eta_0} \cdot (s - \eta_0) + \# A_{l,\eta_0}^c \cdot \max_{\mathbf{x} \in A_{l,\eta_0}^c} s_{\mathbf{x}}^{(l)} \quad \text{(by 3.12)} \end{aligned} \tag{3.13}$$

Employing that for the mean entropy

$$s_{\mathbf{x}}^{(l)} \leq \frac{1}{\# \Lambda(l)} S(\psi_{\mathbf{x}}^{(\Lambda(l))}, \varphi^{(\Lambda(l))}) \qquad \text{(by the monotonicity)}$$

and

$$\begin{aligned} s(\psi_{\mathbf{x}}, \varphi, G_l) &= \lim_{\Lambda(\mathbf{m}) \nearrow \mathbb{N}^\nu} \frac{1}{\# \Lambda(\mathbf{m})} S(\psi_{\mathbf{x}}^{(l\mathbf{m})}, \varphi^{(l\mathbf{m})}) \\ &= \sup_{\Lambda(\mathbf{m})} \frac{1}{\# \Lambda(\mathbf{m})} S(\psi_{\mathbf{x}}^{(l\mathbf{m})}, \varphi^{(l\mathbf{m})}) \end{aligned}$$

are fulfilled, we obtain an upper bound for the last term in (3.13):

$$\begin{aligned} \# A_{l,\eta_0}^c \cdot \max_{\mathbf{x} \in A_{l,\eta_0}^c} s_{\mathbf{x}}^{(l)} &\leq \# A_{l,\eta_0}^c \cdot \max_{\mathbf{x} \in A_{l,\eta_0}^c} \frac{1}{l^\nu} s(\psi_{\mathbf{x}}, \varphi, G_l) \\ &= \# A_{l,\eta_0}^c \cdot s(\psi, \varphi) \qquad \text{(by (3.11))}. \end{aligned}$$

Inserting this in (3.13) and dividing by $\# \Lambda(\mathbf{k}(l))$ we obtain

$$s^{(l)} < \frac{\# A_{l,\eta_0}}{\# \Lambda(\mathbf{k}(l))} (s - \eta_0) + \frac{\# A_{l,\eta_0}^c}{\# \Lambda(\mathbf{k}(l))} s.$$

And after taking limits we arrive at the following contradictory inequality:

$$s \leq a(s - \eta_0) + (1 - a)s = s - a\eta_0 < s,$$

since $\lim_{l \to \infty} s^{(l)} = s$ by Theorem 3.1.4 and $s < \infty$. Hence $a = 0$. $\square$

Lemma 3.1.6 *Let ψ be an ergodic state on $\mathcal{A}^\infty$ and let φ be a stationary product state on $\mathcal{A}^\infty$ fulfilling $s(\psi,\varphi)<\infty$. Then for every $\varepsilon\in(0,1)$*

$$\limsup_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu}\frac{1}{\#\Lambda(\mathbf{n})}\beta_{\varepsilon,\mathbf{n}}(\psi,\varphi)\le -s(\psi,\varphi).$$

Proof of Lemma 3.1.6: We fix $\varepsilon>0$ and choose arbitrary η, $\delta>0$. Consider the G_l-ergodic decomposition

$$\psi=\frac{1}{\#\Lambda(\mathbf{k}(l))}\sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))}\psi_{\mathbf{x}}$$

of ψ for integers $l\ge 1$. By Lemma 3.1.5 there is an integer $L\ge 1$ such that for any $l\ge L$

$$\frac{\varepsilon}{2}\ge\frac{1}{\#\Lambda(\mathbf{k}(l))}\#A_{l,\eta}\ge 0$$

holds, where $A_{l,\eta}$ is defined by (3.12). This inequality implies

$$\frac{1}{\#\Lambda(\mathbf{k}(l))}\#A^c_{l,\eta}\cdot(1-\frac{\varepsilon}{2})\ge 1-\varepsilon. \tag{3.14}$$

Recall that by definition of $A_{l,\eta}$ we have

$$\frac{1}{\#\Lambda(l)}S(\psi_{\mathbf{x}}^{(\Lambda(l))}\upharpoonright\mathcal{B}_l,\varphi^{(\Lambda(l))}\upharpoonright\mathcal{B}_l)\ge s(\psi,\varphi)-\eta\qquad\text{for all }\mathbf{x}\in A^c_{l,\eta}. \tag{3.15}$$

We fix an $l\ge L$ and consider the abelian quasilocal C^*-algebra $\mathcal{B}_l^\infty$ built up from $\mathcal{B}_l$. $\mathcal{B}_l^\infty$ is clearly a C^*-subalgebra of $\mathcal{A}^\infty$. We set

$$m_{\mathbf{x}}:=\psi_{\mathbf{x}}\upharpoonright\mathcal{B}_l^\infty\text{ and }m_{\mathbf{x}}^{(\mathbf{n})}:=\psi_{\mathbf{x}}\upharpoonright\mathcal{B}_l^{(\mathbf{n})}.$$

Moreover we define

$$p:=\varphi\upharpoonright\mathcal{B}_l^\infty\text{ and }p^{(\mathbf{n})}:=\varphi\upharpoonright\mathcal{B}_l^{(\mathbf{n})}.$$

The state p is a G_l-stationary product state. On the other hand, Theorem 4.3.17. in [7] shows that the states $m_{\mathbf{x}}$ are G_l-ergodic. Due to the Gelfand isomorphism and the Riesz representation theorem we can (and shall) identify all the states above with the probability measures on the corresponding maximal ideal space of $\mathcal{B}_l^\infty$. Since the algebra $\mathcal{B}_l$ is abelian and finite dimensional this compact maximal ideal space can be thought of as $B_l^{\mathbb{Z}^\nu}$ for an appropriately chosen finite set B_l. This is essentially the well-known Kolmogorov representation of a classical dynamical system.

By the definition of the measures $m_{\mathbf{x}}$ and p, monotonicity of the relative entropy and the fourth item in Theorem 2.2.1 we have

$$D_M(m_{\mathbf{x}},p)\le s(\psi_{\mathbf{x}},\varphi,G_l)=l^\nu s(\psi,\varphi)<\infty, \tag{3.16}$$

where $D_M(m_{\mathbf{x}}, p)$ denotes the mean relative entropy of $m_{\mathbf{x}}$ with respect to p. Using Theorem 3.1.3 we see that

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\Lambda(\mathbf{n})} \log \frac{m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}})}{p^{(\mathbf{n})}(\omega_{\mathbf{n}})} = D_M(m_{\mathbf{x}}, p) =: D_{M,\mathbf{x}} \tag{3.17}$$

$m_{\mathbf{x}}$-almost surely for all $\mathbf{x} \in \Lambda(\mathbf{k}(l))$, where $\omega_{\mathbf{n}} \in B_l^{\Lambda(\mathbf{n})}$ are the components of $\omega \in B_l^{\mathbb{Z}^\nu}$ corresponding to the box $\Lambda(\mathbf{n})$. For each $\mathbf{n}$ and $\mathbf{x} \in A_{l,\eta}^c$ let

$$\begin{aligned} C_{\mathbf{x}}^{(\mathbf{n})} &:= \left\{ \omega_{\mathbf{n}} \in B_l^{(\mathbf{n})} : \left|\frac{1}{\#\Lambda(\mathbf{n})} \log \frac{m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}})}{p^{(\mathbf{n})}(\omega_{\mathbf{n}})} - D_{M,\mathbf{x}}\right| < \delta \right\} \\ &= \left\{ \omega_{\mathbf{n}} \in B_l^{(\mathbf{n})} : e^{\#\Lambda(\mathbf{n})\cdot(D_{M,\mathbf{x}}-\delta)} < \frac{m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}})}{p^{(\mathbf{n})}(\omega_{\mathbf{n}})} < e^{\#\Lambda(\mathbf{n})\cdot(D_{M,\mathbf{x}}+\delta)} \right\}. \end{aligned}$$

The indicator function of $C_{\mathbf{x}}^{(\mathbf{n})}$ corresponds to the projection $q_{\mathbf{x}}^{(\mathbf{n})} \in \mathcal{B}_l^{(\mathbf{n})}$. The definition of the set $C_{\mathbf{x}}^{(\mathbf{n})}$ implies a bound on the probability of this set with respect to the probability measure p:

$$\begin{aligned} p^{(\mathbf{n})}(C_{\mathbf{x}}^{(\mathbf{n})}) = \varphi(q_{\mathbf{x}}^{(\mathbf{n})}) &\leq m_{\mathbf{x}}^{(\mathbf{n})}(C_{\mathbf{x}}^{(\mathbf{n})}) e^{-\#\Lambda(\mathbf{n})\cdot(D_{M,\mathbf{x}}-\delta)} \\ &\leq e^{-\#\Lambda(\mathbf{n})\cdot(D_{M,\mathbf{x}}-\delta)} \\ &\leq e^{-\#\Lambda(\mathbf{n})\cdot(\#\Lambda(l)(s(\psi,\varphi)-\eta)-\delta)}, \end{aligned} \tag{3.18}$$

because for all $\mathbf{x} \in A_{l,\eta}^c$ we have

$$\begin{aligned} D_{M,\mathbf{x}} &\geq D(m_{\mathbf{x}}^{(1)}, p^{(1)}) = S(\psi_{\mathbf{x}}^{(\Lambda(l))} \restriction \mathcal{B}_l, \varphi^{(\Lambda(l))} \restriction \mathcal{B}_l) \\ &\geq \#\Lambda(l)(s(\psi,\varphi) - \eta), \end{aligned} \tag{3.19}$$

by (2.1) and (3.15). The limit assertion (3.17) implies the existence of an $N \in \mathbb{N}$ such that for any $\mathbf{n} \in \mathbb{N}^\nu$ with $\Lambda(\mathbf{n}) \supset \Lambda(N)$

$$m_{\mathbf{x}}^{(\mathbf{n})}(C_{\mathbf{x}}^{(\mathbf{n})}) \geq 1 - \frac{\varepsilon}{2}, \qquad \forall\, \mathbf{x} \in A_{l,\eta}^c. \tag{3.20}$$

For each $\mathbf{y} \in \mathbb{N}^\nu$ with $y_i \geq Nl$ let $y_i = n_i l + j_i$, where $n_i \geq N$ and $0 \leq j_i < l$. We set

$$q_{l\mathbf{n}} := \bigvee_{\mathbf{x}\in A_{l,\eta}^c} q_{\mathbf{x}}^{(\mathbf{n})}$$

and denote by $q_{\mathbf{y}}$ the embedding of $q_{l\mathbf{n}}$ in $\mathcal{A}^{(\mathbf{y})}$. By (3.20) and (3.14) we obtain

$$\begin{aligned} \Psi^{(\mathbf{y})}(q_{\mathbf{y}}) &= \frac{1}{\#\Lambda(k(l))} \sum_{\mathbf{x}\in\Lambda(k(l))} \psi_{\mathbf{x}}^{(\mathbf{y})}(q_{\mathbf{y}}) \\ &\geq \frac{1}{\#\Lambda(k(l))} \sum_{\mathbf{x}\in\Lambda(k(l))} \psi_{\mathbf{x}}^{(\mathbf{y})}(q_{\mathbf{x}}^{(\mathbf{n})}) \\ &\geq \frac{1}{\#\Lambda(k(l))} \#A_{l,\eta}^c \cdot (1 - \frac{\varepsilon}{2}) \geq 1 - \varepsilon. \end{aligned}$$

Thus the condition in the definition of $\beta_{\varepsilon,\mathbf{y}}(\psi,\varphi)$ is satisfied. We are now able to estimate $\beta_{\varepsilon,\mathbf{y}}(\psi,\varphi)$ using (3.18):

$$\begin{aligned}
\beta_{\varepsilon,\mathbf{y}}(\psi,\varphi) &\leq \log \varphi^{(\mathbf{y})}(q_{\mathbf{y}}) \\
&\leq \log \sum_{\mathbf{x}\in A^c_{l,\eta}} e^{-\#\Lambda(\mathbf{n})\cdot(D_{M,\mathbf{x}}-\delta)} \\
&\leq \log(\#A^c_{l,\eta}\cdot e^{-\#\Lambda(\mathbf{n})\cdot(\#\Lambda(l)(s(\psi,\varphi)-\eta)-\delta)}) \\
&= \log(\#A^c_{l,\eta}) - \#\Lambda(l\mathbf{n})(s(\psi,\varphi)-\eta-\frac{\delta}{\#\Lambda(l)}).
\end{aligned}$$

This leads to

$$\limsup_{\Lambda(\mathbf{y})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{y})}\beta_{\varepsilon,\mathbf{y}}(\psi,\varphi) \leq -s(\psi,\varphi)+\eta+\delta,$$

since $\#A^c_{l,\eta}$ does not depend on $\mathbf{n}$ and $\Lambda(\mathbf{y})\nearrow\mathbb{N}^\nu$ if and only if $\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu$. Since η, $\delta>0$ were chosen arbitrarily we have

$$\limsup_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})}\beta_{\varepsilon,\mathbf{n}}(\psi,\varphi) \leq -s(\psi,\varphi). \qquad \square$$

Suppose we are given a sequence $(p_{\mathbf{n}})$ of projections in $\mathcal{A}^{(\mathbf{n})}$ and a stationary state ψ on $\mathcal{A}^\infty$ with mean entropy $s(\psi)$. We consider the positive operators

$$p_{\mathbf{n}}D_{\psi^{(\mathbf{n})}}p_{\mathbf{n}} = \sum_{i=1}^{d(\mathbf{n})}\lambda_{\mathbf{n},i}r_{\mathbf{n},i} \qquad d(\mathbf{n}) := \mathrm{tr}(p_{\mathbf{n}}), \tag{3.21}$$

where the numbers $\lambda_{\mathbf{n},i}$ are the eigenvalues and the $r_{\mathbf{n},i}$ form a complete set of eigen-projections of $p_{\mathbf{n}}D_{\psi^{(\mathbf{n})}}p_{\mathbf{n}}$. We set

$$T_{\mathbf{n},\delta} := \{i\in\{1,\dots,d(\mathbf{n})\}:\ \lambda_{\mathbf{n},i}\leq e^{-\#\Lambda(\mathbf{n})(s(\psi)-\delta)}\} \qquad \text{for } \delta>0, \tag{3.22}$$

and denote by $T^c_{\mathbf{n},\delta}$ the complement of this set.

Lemma 3.1.7 *Let ψ be an ergodic state on $\mathcal{A}^\infty$ with mean entropy $s(\psi)$ and let $(p_{\mathbf{n}})$ be a sequence of projections in $\mathcal{A}^{(\mathbf{n})}$, respectively. If $p_{T^c_{\mathbf{n},\delta}}$ is the projection corresponding to the set $T^c_{\mathbf{n},\delta}$ then*

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu}\psi^{(\mathbf{n})}(p_{T^c_{\mathbf{n},\delta}}) = 0$$

for all $\delta>0$.

Proof of Lemma 3.1.7: We have

$$1 \geq \sum_{i \in T^c_{\mathbf{n},\delta}} \lambda_{\mathbf{n},i} > e^{-\#\Lambda(\mathbf{n})(s(\psi)-\delta)} \# T^c_{\mathbf{n},\delta} = e^{-\#\Lambda(\mathbf{n})(s(\psi)-\delta)} \mathrm{tr}(p_{T^c_{\mathbf{n},\delta}}),$$

and, consequently

$$\frac{1}{\#\Lambda(\mathbf{n})} \log(\mathrm{tr}(p_{T^c_{\mathbf{n},\delta}})) < s(\psi) - \delta.$$

If we would have $\psi^{(\Lambda(\mathbf{n}))}(p_{T^c_{\mathbf{n},\delta_0}}) \geq a > 0$ for infinitely many $\mathbf{n}$ and some $a > 0$ there would be a contradiction to Proposition 2.3.2, which implies that there is no sequence $(q_{\mathbf{n}})$ of projections in $\mathcal{A}^{(\mathbf{n})}$ with $\psi^{(\mathbf{n})}(q_{\mathbf{n}}) \geq a > 0$ and $\limsup_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} \log(\mathrm{tr}(q_{\mathbf{n}})) < s(\psi)$. □

Lemma 3.1.8 *Let ψ be an ergodic state on $\mathcal{A}^\infty$ and let φ be a stationary product state on $\mathcal{A}^\infty$. Suppose that $s(\psi, \varphi) < \infty$ holds, then for every $\varepsilon \in (0, 1)$*

$$\liminf_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} \beta_{\varepsilon,\mathbf{n}}(\psi, \varphi) \geq -s(\psi, \varphi).$$

Proof of Lemma 3.1.8: Let $(t_{\mathbf{y}})$ be a sequence of projections, $t_{\mathbf{y}} \in \mathcal{A}^{(\mathbf{y})}$, with $\Lambda(\mathbf{y}) \nearrow \mathbb{N}^\nu$ and

$$\liminf_{\Lambda(\mathbf{y}) \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{y})} \log \varphi^{(\mathbf{y})}(t_{\mathbf{y}}) < -s(\psi, \varphi).$$

Then there exists an $a > 0$ and a subsequence, which we denote by $(t_{\mathbf{y}})$ for notational simplicity, fulfilling

$$\varphi^{(\mathbf{y})}(t_{\mathbf{y}}) < e^{-\#\Lambda(\mathbf{y})(s(\psi,\varphi)+a)}. \tag{3.23}$$

We consider an integer $l \geq 1$ and the G_l−ergodic decomposition of ψ:

$$\psi = \frac{1}{\#\Lambda(\mathbf{k}(l))} \sum_{\mathbf{x} \in \Lambda(\mathbf{k}(l))} \psi_{\mathbf{x}}.$$

As in the proof of Lemma 3.1.6 for each $\varepsilon, \eta > 0$ we can choose l in such a way that we have

$$\frac{\#A^c_{l,\eta}}{\#\Lambda(\mathbf{k}(l))}(1 - \frac{\varepsilon}{2}) \geq 1 - \varepsilon, \tag{3.24}$$

where the set $A_{l,\eta}$ was defined by (3.12). Recall that we have by definition,

$$\frac{1}{\#\Lambda(l)} S(\psi^{(l)}_{\mathbf{x}} \upharpoonright \mathcal{B}_l, \varphi^{(l)} \upharpoonright \mathcal{B}_l) \geq s(\psi, \varphi) - \eta \quad \text{for all } \mathbf{x} \in A^c_{l,\eta}.$$

We consider again the abelian quasilocal algebra $\mathcal{B}_l^\infty$, which will be identified with the algebra of continuous functions on the maximal ideal space $B_l^\infty := B_l^{\mathbb{Z}^\nu}$, bearing in mind that the restrictions of the G_l-ergodic components of ψ and φ to this algebra are G_l-ergodic. We denote those restrictions by $m_{\mathbf{x}}$, $\mathbf{x} \in \Lambda(\mathbf{k}(l))$, and p. As in the proof of Lemma 3.1.6 we can show that for $\delta > 0$ the sets

$$C_{\mathbf{x},\delta}^{(\mathbf{n})} := \{\omega_{\mathbf{n}} \in B_l^{(\mathbf{n})} : \ e^{\#\Lambda(\mathbf{n})(D_{M,\mathbf{x}} - \frac{\delta}{2})} < \frac{m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}})}{p^{(\mathbf{n})}(\omega_{\mathbf{n}})} < e^{\#\Lambda(\mathbf{n})(D_{M,\mathbf{x}} + \frac{\delta}{2})}\}$$

fulfil

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} m_{\mathbf{x}}^{(\mathbf{n})}(C_{\mathbf{x},\delta}^{(\mathbf{n})}) = 1 \quad \text{for all } \mathbf{x} \in \Lambda(\mathbf{k}(l)). \tag{3.25}$$

In a similar way, by employing the classical Shannon-McMillan theorem, we can see that for

$$F_{\mathbf{x},\delta}^{(\mathbf{n})} := \{\omega_{\mathbf{n}} \in B_l^{(\mathbf{n})} : e^{-\#\Lambda(\mathbf{n})(h_{\mathbf{x}} + \frac{\delta}{2})} < m_{\mathbf{x}}^{(\mathbf{n})}(\omega_{\mathbf{n}}) < e^{-\#\Lambda(\mathbf{n})(h_{\mathbf{x}} - \frac{\delta}{2})}\},$$

we have

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} m_{\mathbf{x}}^{(\mathbf{n})}(F_{\mathbf{x},\delta}^{(\mathbf{n})}) = 1 \quad \text{for all } \mathbf{x} \in \Lambda(\mathbf{k}(l)), \tag{3.26}$$

where $h_{\mathbf{x}}$ denotes the Shannon mean entropy of $m_{\mathbf{x}}$. Each $\omega_{\mathbf{n}} \in B_l^{(\mathbf{n})}$ corresponds to a minimal projection $q_{\mathbf{n}} \in \mathcal{B}_l^{(\mathbf{n})} \subset \mathcal{A}^{(l\mathbf{n})}$. So, for all $\mathbf{x} \in A_{l,\eta}^c$ and for any $\omega_{\mathbf{n}} \in C_{\mathbf{x},\delta}^{(\mathbf{n})} \cap F_{\mathbf{x},\delta}^{(\mathbf{n})}$ and corresponding $q_{\mathbf{n}}$ we have

$$\begin{aligned} \varphi^{(l\mathbf{n})}(q_{\mathbf{n}}) &= p^{(\mathbf{n})}(\omega_{\mathbf{n}}) \\ &> e^{-\#\Lambda(\mathbf{n})(D_{M,\mathbf{x}} + h_{\mathbf{x}} + \delta)} \\ &\geq e^{-\#\Lambda(\mathbf{n})(s(\psi,\varphi)l^\nu + h_{\mathbf{x}} + \delta)}. \end{aligned} \tag{3.27}$$

In fact, the first inequality follows from the definitions of the sets $C_{\mathbf{x},\delta}^{(\mathbf{n})}$ and $F_{\mathbf{x},\delta}^{(\mathbf{n})}$ and the assumption that $\omega_{\mathbf{n}} \in C_{\mathbf{x},\delta}^{(\mathbf{n})} \cap F_{\mathbf{x},\delta}^{(\mathbf{n})}$. The second inequality is a consequence of (3.16). Moreover, by (3.19) we get the upper estimate

$$\begin{aligned} \varphi^{(l\mathbf{n})}(q_{\mathbf{n}}) &= p^{(\mathbf{n})}(\omega_{\mathbf{n}}) \\ &< e^{-\#\Lambda(\mathbf{n})(D_{M,\mathbf{x}} + h_{\mathbf{x}} - \delta)} \\ &\leq e^{-\#\Lambda(\mathbf{n})((s(\psi,\varphi)-\eta)l^\nu + h_{\mathbf{x}} - \delta)}. \end{aligned} \tag{3.28}$$

Next, observe that the representation (3.8) of $\mathcal{B}_l$ implies that

$$D_{\psi_{\mathbf{x}}^{(l)}\upharpoonright\mathcal{B}_l} = \sum q_{l,i} D_{\psi_{\mathbf{x}}^{(l)}} q_{l,i},$$

where $(q_{l,i})$ is a complete set of minimal eigen-projections of $D_{\varphi^{(l)}}$. Observe that from the fact that the maximal abelian algebra $\mathcal{B}_l$ is generated by eigen-projections of $D_{\varphi^{(l)}}$ it follows that $D_{\varphi^{(l)}\upharpoonright\mathcal{B}_l} = D_{\varphi^{(l)}}$. We get

$$\begin{aligned}
\mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}\upharpoonright\mathcal{B}_l} \log D_{\varphi^{(l)}\upharpoonright\mathcal{B}_l}) &= \mathrm{tr}(\sum q_{l,i} D_{\psi_{\mathbf{x}}^{(l)}} q_{l,i} \log D_{\varphi^{(l)}}) \\
&= \mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}} \sum q_{l,i} (\log D_{\varphi^{(l)}}) q_{l,i}) \\
&= \mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}} \log D_{\varphi^{(l)}}). \qquad (3.29)
\end{aligned}$$

Finally, by our assumption that $s(\psi,\varphi) < \infty$ holds we know that $s(\psi,\varphi) = -s(\psi) - \mathrm{tr}(D_{\psi^{(1)}} \log D_{\varphi^{(1)}})$. Using the product structure of the state φ and the fact just derived we obtain

$$\begin{aligned}
S(\psi_{\mathbf{x}}^{(l\mathbf{n})} \upharpoonright \mathcal{B}_l^{(\mathbf{n})}, \varphi^{(l\mathbf{n})} \upharpoonright \mathcal{B}_l^{(\mathbf{n})}) &= -H(m_{\mathbf{x}}^{(\mathbf{n})}) - \mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l\mathbf{n})}\upharpoonright\mathcal{B}_l^{\mathbf{n}}} \log D_{\varphi^{(l\mathbf{n})}}) \\
&= -H(m_{\mathbf{x}}^{(\mathbf{n})}) - \#\Lambda(\mathbf{n}) \mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}\upharpoonright\mathcal{B}_l} \log D_{\varphi^{(l)}}) \\
&= -H(m_{\mathbf{x}}^{(\mathbf{n})}) - \#\Lambda(\mathbf{n})(\mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}} \log D_{\varphi^{(l)}})(\text{by } (3.29)). \qquad (3.30)
\end{aligned}$$

On the other hand we have for all $\mathbf{x} \in A_{l,\eta}^c$

$$\begin{aligned}
S(\psi_{\mathbf{x}}^{(l\mathbf{n})} \upharpoonright \mathcal{B}_l^{(\mathbf{n})}, \varphi^{(l\mathbf{n})} \upharpoonright \mathcal{B}_l^{(\mathbf{n})}) &\geq \#\Lambda(\mathbf{n})(-H(m_{\mathbf{x}}^{(1)}) - \mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}\upharpoonright\mathcal{B}_l} \log D_{\varphi^{(l)}})) \\
&\quad \text{(by the subadditivity of the entropy)} \\
&= \#\Lambda(\mathbf{n}) S(\psi_{\mathbf{x}}^{(l)} \upharpoonright \mathcal{B}_l, \varphi^{(l)} \upharpoonright \mathcal{B}_l) \\
&\geq \#\Lambda(\mathbf{n}) l^\nu (s(\psi,\varphi) - \eta) \\
&\quad \text{(by the choice of algebra } \mathcal{B}_l) \\
&= \#\Lambda(\mathbf{n}) l^\nu (-s(\psi) - \mathrm{tr}(D_{\psi^{(1)}} \log D_{\varphi^{(1)}}) - \eta). \qquad (3.31)
\end{aligned}$$

The equation chain (3.30) and the inequality chain (3.31) imply that

$$\begin{aligned}
-H(m_{\mathbf{x}}^{(\mathbf{n})}) &- \#\Lambda(\mathbf{n})(\mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}} \log D_{\varphi^{(l)}})) \\
&\geq \#\Lambda(\mathbf{n}) l^\nu (-s(\psi) - \mathrm{tr}(D_{\psi^{(1)}} \log D_{\varphi^{(1)}}) - \eta). \qquad (3.32)
\end{aligned}$$

Note that the third and the fourth item of Theorem 2.2.1 show that

$$\mathrm{tr}(D_{\psi_{\mathbf{x}}^{(l)}} \log D_{\varphi^{(l)}}) = l^\nu \mathrm{tr}(D_{\psi^{(1)}} \log D_{\varphi^{(1)}}),$$

in the case where $s(\psi,\varphi) < \infty$. Dividing (3.32) by $-\#\Lambda(\mathbf{n})$ and taking the limit $\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu$ leads to

$$l^\nu s(\psi) \leq h_{\mathbf{x}} \leq l^\nu (s(\psi) + \eta), \qquad (3.33)$$

where the lower bound simply follows from the fact that the entropy on a maximally abelian subalgebra is not less than the entropy on the algebra. This inequality, (3.27) and (3.28) imply:

$$e^{-\#\Lambda(l\mathbf{n})(s(\psi,\varphi)+s(\psi)+\eta+\delta/l^{\nu})} \leq \varphi^{(l\mathbf{n})}(q_{\mathbf{n}}) \leq e^{-\#\Lambda(l\mathbf{n})(s(\psi,\varphi)+s(\psi)-\eta-\delta/l^{\nu})}. \quad (3.34)$$

Let $p_{l\mathbf{n}}$ denote the projection corresponding to the set $\bigcup_{\mathbf{x}\in A^c_{l,\eta}} C^{(\mathbf{n})}_{\mathbf{x},\delta} \cap F^{(\mathbf{n})}_{\mathbf{x},\delta}$ and $p_{\mathbf{n},\mathbf{x}}$ be the projection which corresponds to $C^{(\mathbf{n})}_{\mathbf{x},\delta} \cap F^{(\mathbf{n})}_{\mathbf{x},\delta}$. For sufficiently large $\mathbf{n} \in \mathbb{N}^{\nu}$ we have

$$\psi_{\mathbf{x}}(p_{\mathbf{n},\mathbf{x}}) \geq 1 - \frac{\varepsilon}{2} \qquad \text{by (3.25) and (3.26)},$$

and hence by (3.24)

$$\begin{aligned} \psi(p_{l\mathbf{n}}) &\geq \frac{1}{\#\Lambda(\mathbf{k}(l))} \sum_{\mathbf{x}\in A^c_{l,\eta}} \psi_{\mathbf{x}}(p_{\mathbf{n},\mathbf{x}}) \\ &\geq \frac{1}{\#\Lambda(\mathbf{k}(l))} \#A^c_{l,\eta}(1-\frac{\varepsilon}{2}) \geq 1-\varepsilon. \end{aligned} \quad (3.35)$$

Any $t_{\mathbf{y}}$ fulfilling (3.23) can be embedded in an appropriately chosen $\mathcal{A}^{(l\mathbf{n})}$. Indeed, choose the unique $\mathbf{n} \in \mathbb{N}^{\nu}$ such that $(n_i-1)l \leq y_i < n_i l$ for all $i \in \{1,\ldots,\nu\}$ and set

$$e_{l\mathbf{n}} := t_{\mathbf{y}} \otimes \mathbf{1}_{\Lambda(l\mathbf{n})\setminus\Lambda(\mathbf{y})}.$$

Then we have

$$\psi^{(l\mathbf{n})}(e_{l\mathbf{n}}) = \psi^{(\mathbf{y})}(t_{\mathbf{y}}), \quad (3.36)$$

and

$$\begin{aligned} \psi^{(l\mathbf{n})}(e_{l\mathbf{n}}) &= \psi^{(l\mathbf{n})}(e_{l\mathbf{n}}p_{l\mathbf{n}}) + \psi^{(l\mathbf{n})}(e_{l\mathbf{n}}(\mathbf{1}-p_{l\mathbf{n}})) \\ &\leq \psi^{(l\mathbf{n})}(e_{l\mathbf{n}}p_{l\mathbf{n}}) + \varepsilon \qquad (\text{by (3.35)}). \end{aligned}$$

Applying this argument once more we obtain

$$\begin{aligned} \psi^{(l\mathbf{n})}(e_{l\mathbf{n}}) &\leq \psi^{(l\mathbf{n})}(p_{l\mathbf{n}}e_{l\mathbf{n}}p_{l\mathbf{n}}) + 2\varepsilon \\ &= \mathrm{tr}(p_{l\mathbf{n}}D_{\psi^{(l\mathbf{n})}}p_{l\mathbf{n}}e_{l\mathbf{n}}) + 2\varepsilon. \end{aligned} \quad (3.37)$$

In the final step we will prove that the first term in the last line above can be made arbitrarily small. Using the notation from (3.21), (3.22) and applying Lemma 3.1.7 we know that

$$\begin{aligned} \mathrm{tr}(p_{l\mathbf{n}}D_{\psi^{(l\mathbf{n})}}p_{l\mathbf{n}}e_{l\mathbf{n}}) &\leq e^{-\#\Lambda(l\mathbf{n})(s(\psi)-\delta)} \sum_{i\in T_{l\mathbf{n},\delta}} \mathrm{tr}(r_{l\mathbf{n},i}e_{l\mathbf{n}}) + \varepsilon \\ &\leq e^{-\#\Lambda(l\mathbf{n})(s(\psi)-\delta)} \sum_{i=1}^{\mathrm{tr}(p_{l\mathbf{n}})} \mathrm{tr}(r_{l\mathbf{n},i}e_{l\mathbf{n}}) + \varepsilon \\ &= e^{-\#\Lambda(l\mathbf{n})(s(\psi)-\delta)}\mathrm{tr}(p_{l\mathbf{n}}e_{l\mathbf{n}}) + \varepsilon, \end{aligned} \quad (3.38)$$

for sufficiently large $\mathbf{n} \in \mathbb{N}^\nu$. We represent the projection $p_{l\mathbf{n}}$ as a sum of unique minimal projections in $\mathcal{B}_l^{(\mathbf{n})}$:

$$p_{l\mathbf{n}} = \sum_{i=1}^{\mathrm{tr}(p_{l\mathbf{n}})} q_{\mathbf{n},i}.$$

Hence by (3.34) we have

$$\begin{aligned}
\mathrm{tr}(p_{l\mathbf{n}} e_{l\mathbf{n}}) &\leq e^{\#\Lambda(l\mathbf{n})(s(\psi,\varphi)+s(\psi)+\eta+\delta/l^\nu)} \sum_{i=1}^{\mathrm{tr}(p_{l\mathbf{n}})} \varphi^{(l\mathbf{n})}(q_{\mathbf{n},i})\mathrm{tr}(q_{\mathbf{n},i} e_{l\mathbf{n}}) \\
&= e^{\#\Lambda(l\mathbf{n})(s(\psi,\varphi)+s(\psi)+\eta+\delta/l^\nu)} \varphi^{(l\mathbf{n})}(p_{l\mathbf{n}} e_{l\mathbf{n}}) \\
&\leq e^{\#\Lambda(l\mathbf{n})(s(\psi,\varphi)+s(\psi)+\eta+\delta/l^\nu)} \varphi^{(\mathbf{y})}(t_{\mathbf{y}}) \\
&\leq e^{\#\Lambda(l\mathbf{n})(s(\psi,\varphi)+s(\psi)+\eta+\delta/l^\nu)} e^{-\#\Lambda(\mathbf{y})(s(\psi,\varphi)+a)} \text{ (by (3.23))},
\end{aligned} \tag{3.39}$$

since the minimal projections $q_{\mathbf{n},i}$ correspond to the eigen-vectors of $D_{\varphi^{(l\mathbf{n})}}$. Inserting this into (3.38) we obtain

$$\mathrm{tr}(p_{l\mathbf{n}} D_{\psi^{(l\mathbf{n})}} p_{l\mathbf{n}} e_{l\mathbf{n}}) \leq e^{-\#\Lambda(l\mathbf{n})((\frac{\#\Lambda(\mathbf{y})}{\#\Lambda(l\mathbf{n})}-1)s(\psi,\varphi)+\frac{\#\Lambda(\mathbf{y})}{\#\Lambda(l\mathbf{n})}a-\delta-\eta-\frac{\delta}{l^\nu})} + \varepsilon.$$

Note that $\lim_{\Lambda(\mathbf{y})\nearrow\mathbb{N}^\nu} \frac{\#\Lambda(\mathbf{y})}{\#\Lambda(l\mathbf{n})} = 1$, and that $a > 0$. Hence if we choose δ, η small enough and $\mathbf{n}$ large enough we can achieve that the exponent in the last inequality becomes negative eventually. This inequality toghether with (3.37) and (3.36) shows that

$$\lim_{\Lambda(\mathbf{y})\nearrow\mathbb{N}^\nu} \psi(t_{\mathbf{y}}) = 0. \qquad \square$$

3.1.2 Proof of the Quantum Relative AEP

To derive Theorem 3.1.2, consider the projections $p_{l\mathbf{n}}$ constructed in the proof of Lemma 3.1.8. They can be written as the sum of minimal projections $q_{\mathbf{n}}$ fulfilling (3.34). From the fact, that these are minimal eigen-projections for $D_{\varphi^{(\mathbf{n})}}$ we derive that (3.34) is valid even if we replace $q_{\mathbf{n}}$ by any minimal projection which is dominated by $p_{l\mathbf{n}}$. So we see that, if we would choose the projections $p_{\mathbf{y}}(\varepsilon)$ to be specified for Theorem 3.1.2 just as $p_{l\mathbf{n}}$ for boxes with edge lengths being multiples of l, item 3 would be satisfied for large $\mathbf{n}$, supposed l is (fixed but) large enough. Item 1 is fulfilled for large $\mathbf{n}$ in view of (3.35). Observe now that embedding the projections $p_{l\mathbf{n}}$ into $\mathcal{A}^{(\mathbf{y})}$ for $0 \leq \mathbf{r} := \mathbf{y} - l\mathbf{n} < (l, l, \dots, l)$, i.e. $\mathbf{y} = l\mathbf{n} + \mathbf{r}$ and $0 \leq \mathbf{r} < (l, l, \dots, l)$, leads us to a family of projections $(p_{\mathbf{y}})$ still fulfilling item 1. Item 3 is satisfied by this family, too, since $\varphi^{(1)}$ was supposed faithful. Indeed, let $\lambda_{\min} > 0$ denote the smallest eigenvalue of the density operator $D_{\varphi^{(1)}}$ and let $\{e_{\mathbf{r},j}\}_{j=1}^{\#\Lambda(\mathbf{r})}$ be a complete set of minimal eigen-projections of $D_{\varphi^{(\mathbf{r})}}$. Then taking into account (3.34) and that $\varphi^{(\mathbf{y})}$ is a product state we can infer that

$$\begin{aligned}\varphi^{(\mathbf{y})}(q_{\mathbf{n},i}\otimes e_{\mathbf{r},j}) &\geq \varphi^{(l\mathbf{n})}(q_{\mathbf{n},i})e^{\#\Lambda(\mathbf{r})\log\lambda_{\min}}\\ &\geq e^{-\#\Lambda(l\mathbf{n})(s(\psi,\varphi)+s(\psi)+\eta+\delta/l^{\nu})}e^{\#\Lambda(\mathbf{r})\log\lambda_{\min}}\\ &> e^{-\#\Lambda(\mathbf{y})(s(\psi,\varphi)+s(\psi)+\varepsilon)},\end{aligned}$$

holds for sufficiently small η and δ and $\mathbf{n}$ large enough. Using the largest eigenvalue $\lambda_{\max}$ of $D_{\varphi^{(1)}}$ a similar argument shows that

$$\varphi^{(\mathbf{y})}(q_{\mathbf{n},i}\otimes e_{\mathbf{r},j}) < e^{-\#\Lambda(\mathbf{y})(s(\psi,\varphi)+s(\psi)-\varepsilon)},$$

also holds for sufficiently large $\mathbf{y}$. Note that each $q_{\mathbf{n},i}\otimes e_{\mathbf{r},j}$ is a minimal eigen-projection of $D_{\varphi^{(\mathbf{y})}}$ which implies that item 3 holds for each minimal projection dominated by $p_{\mathbf{y}}$.
In order to ensure that item 2 is fulfilled, the constructed family of projections $(p_{\mathbf{y}})$ has to be modified. Represent $p_{\mathbf{y}}$ as a sum of eigen-projections of the operator $p_{\mathbf{y}}D_{\psi^{(\mathbf{y})}}p_{\mathbf{y}}$. Now from the definition of the sets $F^{(\mathbf{n})}_{\mathbf{x},\delta}$ we easily conclude that $\mathrm{tr}(p_{\mathbf{y}}) < e^{\#\Lambda(\mathbf{y})(s(\psi)+\varepsilon)}$ can be guaranteed for large $\mathbf{y}$. So the asymptotic contribution to $\psi^{(\mathbf{y})}(p_{\mathbf{y}})$ of eigen-values of $p_{\mathbf{y}}D_{\psi^{(\mathbf{y})}}p_{\mathbf{y}}$ of magnitude exponentially smaller than $e^{-\#\Lambda(\mathbf{y})(s(\psi)+\varepsilon)}$ can be neglected. The asymptotic contribution to $\psi^{(\mathbf{y})}(p_{\mathbf{y}})$ of eigen-values of $p_{\mathbf{y}}D_{\psi^{(\mathbf{y})}}p_{\mathbf{y}}$ of magnitude exponentially larger than $e^{-\#\Lambda(\mathbf{y})(s(\psi)-\varepsilon)}$ can be neglected, too, because of Lemma 3.1.7. So we may omit the eigen-projections corresponding to either too large or too small eigenvalues from the sum, getting a modified family $(p'_{\mathbf{y}})$, which additionally fulfils item 2. This proves Theorem 3.1.2.

3.2 Concluding Remarks

Note that in [23] Shields had given an example of an B-process (i.e. a stationary encoding of an i.i.d. process) which has the property that upper resp. lower limit of the relative entropies per site of a certain ergodic process with respect to that B-process are ∞ resp. 0. Hence, the assumption in Stein's lemma that the reference state should be a stationary product state, even in the classical situation, cannot be weakened essentially. However, there is a promising indication (cf. [9]) that one could extend Stein's lemma to reference states which are strongly mixing and algebraic, a class of states which is closely related to strongly mixing Markov states. The essential property of these states is that they are, in some sense (cf. [9]), almost product states. In [9] Hiai and Petz could show that for each completely ergodic state ψ and each strongly mixing algebraic state φ the mean relative entropy exists and that

$$\limsup_{n\to\infty}\frac{1}{n}\beta_{\varepsilon,n}(\psi,\varphi)\leq -s(\psi,\varphi),$$

$$\liminf_{n\to\infty} \frac{1}{n}\beta_{\varepsilon,n}(\psi,\varphi) \geq -\frac{1}{1-\varepsilon}s(\psi,\varphi),$$

hold for each $\varepsilon \in (0,1)$. They also conjectured that the limit of the normalized separation numbers $\frac{1}{n}\beta_{\varepsilon,n}$ exist and equals $-s(\psi,\varphi)$, even for ergodic states ψ.

Chapter 4

Monotonicity of the Quantum Relative Entropy

The monotonicity of the quantum relative entropy is one of its deeper properties. It has been proved by Lindblad in [15] for finite quantum systems and by Uhlmann in [27] for the most general situation. Lindblad based his proof on the strong subadditivity of the von Neumann entropy which was shown by Lieb and Ruskai in [14]. At the present time there are several proofs of these entropic properties. The most prominent proof is perhaps Simon's elegant derivation of the famous Lieb's concavity (cf. [25]), which implies the joint convexity of the quantum relative entropy (cf. [17]). This in turn suffices to show the monotonicity of the quantum relative entropy and the strong subadditivity of the von Neumann entropy (cf. [17]). All these proofs are analytical in spirit. In this chapter we will provide another proof of the monotonicity in the finite dimensional situation which could be more appealing from the point of view of quantum information theory. The guiding idea in our approach is that (in complete analogy to classical information theory) the notion of relative entropy obtains its full information theoretical justification through an appropriate law of large numbers or ergodic theorem. Such ergodic theorem is given by the Stein's lemma of the previous chapter. Ogawa and Nagaoka (cf. [19]) proved that result in the case of stationary product states. However, they used deep properties of the so called quasi-entropies, which are analogous to the monotonicity of the relative entropy. Note that we used the monotonicity and the joint convexity in the proof of the ergodic version of Stein's lemma in the last chapter, but these properties were only needed to estimate entropies of l-ergodic components on local algebras, which is unnecessary for the stationary product case. We present a complete proof here in the product situation which does not make use of these properties.

Once we have this theorem in the stationary product situation it is almost trivial to obtain the monotonicity. In fact, according to this result the relative entropy equals a thermodynamic parameter which trivially decreases if the underlying algebra of observables is replaced by a subalgebra. In view of a well-known representation theorem for completely positive unital maps discovered by Stinespring (cf. [26]) this implies monotonicity. The derivation of the monotonicity is the content of the last section of this chapter.

4.1 Stein's Lemma: The i.i.d. Case

Our proof of the monotonicity of the relative entropy will be based on a law of large numbers for the quantum relative entropy. Let $\psi_\infty, \varphi_\infty$ be stationary product states and let $\varepsilon \in (0,1)$. Recall the definition of the separation numbers $\beta_{\varepsilon,n}$ for the states $\psi_\infty, \varphi_\infty$:

$$\beta_{\varepsilon,n}(\psi_\infty, \varphi_\infty) := \min\{\log \varphi^{(n)}(q) : q \in \mathcal{A}^{(n)} \text{ projection}, \ \psi^{(n)}(q) \geq 1 - \varepsilon\}. \tag{4.1}$$

The projection $q_0^{(n)}$ at which the minimum in (4.1) is attained represents an ideal measurement that distinguishes the states $\psi^{(n)}$ and $\varphi^{(n)}$ in an optimal way. The relationship of the separation numbers defined in (4.1) and the quantum relative entropy is the content of the following theorem.

Theorem 4.1.1 *Let $\psi_\infty, \varphi_\infty$ be stationary product states and let $\varepsilon \in (0,1)$. Suppose that $S(\psi, \varphi) < \infty$ holds. Then we have*

$$\lim_{n \to \infty} \frac{1}{n} \beta_{\varepsilon,n}(\psi_\infty, \varphi_\infty) = -S(\psi, \varphi), \tag{4.2}$$

for all $\varepsilon \in (0,1)$.

Remark Recall that the mean relative entropy of the stationary product state ψ_∞ with respect to the stationary product state φ_∞ fulfills

$$s(\psi_\infty, \varphi_\infty) = S(\psi, \varphi).$$

Proof of Theorem 4.1.1: We choose an $\eta > 0$. Then according to Theorem A.0.2 in the appendix A there exists an integer $l \geq 1$ such that for the abelian algebra $\mathcal{D}_l$ considered in that theorem

$$\frac{1}{l} S(\psi^{(l)} \upharpoonright \mathcal{D}_l, \varphi^{(l)} \upharpoonright \mathcal{D}_l) \geq S(\psi, \varphi) - \eta, \tag{4.3}$$

holds. We consider the quasilocal algebra $\mathcal{D}_l^\infty$ and define

$$P := \psi_\infty \upharpoonright \mathcal{D}_l^\infty \text{ and } Q := \varphi_\infty \upharpoonright \mathcal{D}_l^\infty.$$

Note that these states are i.i.d. states on $\mathcal{D}_l^\infty$. We will think of this abelian algebra as continuous functions on the maximal ideal space X_l^∞ which in turn may be represented as the infinite cartesian product of a finite set X_l. Let D_M denote the classical relative mean entropy of the probability measure P with respect to the probability measure Q. Then we have by (A.3)

$$D_M = S(\psi^{(l)} \upharpoonright \mathcal{D}_l, \varphi^{(l)} \upharpoonright \mathcal{D}_l) \leq S(\psi^{(l)}, \varphi^{(l)}) = lS(\psi, \varphi) < \infty, \tag{4.4}$$

since the projective measurements increase the von Neumann entropy (cf. [17]). We choose a $\delta > 0$ and define

$$C^{(n)} := \{\omega_n \in X_l^n : e^{n(D_M - \delta)} < \frac{P^{(n)}(\omega_n)}{Q^{(n)}(\omega_n)} < e^{n(D_M + \delta)}\}. \tag{4.5}$$

By the law of large numbers we have $\lim_{n\to\infty} P^{(n)}(C^{(n)}) = 1$. Consider the projection $r_{ln} \in \mathcal{A}^{(ln)}$ corresponding to the characteristic function $1_{C^{(n)}}$. Then as in the proof of Lemma 3.1.6 we can show that

$$\limsup_{k\to\infty} \frac{1}{k} \beta_{\varepsilon,k}(\psi_\infty, \varphi_\infty) \leq -S(\psi, \varphi), \tag{4.6}$$

for all $\varepsilon \in (0, 1)$.

In the next step we show the existence of the limit in (4.6) and that this limit equals $-S(\psi, \varphi)$. Let (t_y) be a sequence of projections in $\mathcal{A}^{(y)}$ such that $y \to \infty$ and

$$\liminf_{y\to\infty} \frac{1}{y} \log \varphi^{(y)}(t_y) < -S(\psi, \varphi).$$

Then there exists an $a > 0$ and a subsequence, which we denote by (t_y) again for notational simplicity, with the property

$$\varphi^{(y)}(t_y) < e^{-y(S(\psi,\varphi)+a)}. \tag{4.7}$$

Let $h = H(P^{(1)})$ denote the mean entropy of the i.i.d. measure P and consider the set

$$F^{(n)} := \{\omega_n \in X_l^n : e^{-n(h+\delta)} < P^{(n)}(\omega_n) < e^{-n(h-\delta)}\}, \tag{4.8}$$

then we have

$$\lim_{n\to\infty} P^{(n)}(F^{(n)}) = 1 \qquad \text{by the law of large numbers.} \tag{4.9}$$

Let q_n be the projection in $\mathcal{A}^{(ln)}$ corresponding to some $\omega_n \in C^{(n)} \cap F^{(n)}$ (cf. (4.5)) then we have

$$\begin{aligned} \varphi^{(ln)}(q_n) &= Q^{(n)}(\omega_n) > e^{-n(D_M + h + 2\delta)} \text{ (by definition of } C^{(n)} \text{ and } F^{(n)}) \\ &> e^{-n(lS(\psi,\varphi) + l(S(\psi)+\eta) - E_l + 2\delta)}, \end{aligned} \tag{4.10}$$

where $E_l := \sum_{i=1}^{a_l} \mathrm{tr}(D_{\psi^{(l)}} q_{i,l}) \log \mathrm{tr}(q_{i,l})$ and a_l and $q_{i,l}$ were defined in (A.6). The validity of the last inequality follows on the one hand from (4.4). On the other hand we have

$$\begin{aligned} h &= S(\psi^{(l)} \restriction \mathcal{D}_l) = -S(\psi^{(l)} \restriction \mathcal{D}_l, \varphi^{(l)} \restriction \mathcal{D}_l) - \mathrm{tr}_{\mathcal{D}_l}(D_{\psi^{(l)} \restriction \mathcal{D}_l} \log D_{\varphi^{(l)} \restriction \mathcal{D}_l}) \\ &\leq -l(S(\psi, \varphi) - \eta) - \mathrm{tr}(D_{\psi^{(l)}} \log D_{\varphi^{(l)}}) - E_l \quad \text{(by (4.3) and (A.9))} \\ &= l(S(\psi) + \eta) - E_l, \end{aligned}$$

which shows (4.10). Note that each minimal projection $s_n \in \mathcal{D}_l^{(n)}$ can be written as $\otimes_{j=1}^n q_{i_j,l}$. We define the following set:

$$L^{(n)} := \{s_n \in \mathcal{D}_l^{(n)} : s_n \text{ minimal projection}, \frac{1}{n} \log \mathrm{tr}(s_n) < E_l + \delta\}. \quad (4.11)$$

Then the law of large numbers implies that

$$\lim_{n \to \infty} P^{(n)}(L^{(n)}) = 1, \quad (4.12)$$

after obvious identification of the minimal projections in $\mathcal{D}_l^{(n)}$ with elementary outcomes of X_l^n. (Observe that by definition E_l is the expectation of $\log \mathrm{tr}(q_{(.),l})$ with respect to P.) Let $p_{ln} \in \mathcal{A}^{(ln)}$ denote the projection corresponding to the characteristic function of the set $C^{(n)} \cap F^{(n)} \cap L^{(n)}$. Then for all sufficiently large n we have $\psi(p_{ln}) \geq 1 - \varepsilon$ by (4.9), (4.12) and (4.5). Consider the embedding e_{ln} of the projections t_y fulfilling (4.7) into an appropriately chosen $\mathcal{A}^{(ln)}$. For example we choose the unique $n \in \mathbb{N}$ with $(n-1)l \leq y < nl$. Then we have

$$\psi^{(ln)}(e_{ln}) = \psi^{(y)}(t_y) \quad (4.13)$$

and as in the proof of Lemma 3.1.8 we can infer

$$\psi^{(ln)}(e_{ln}) = \mathrm{tr}(p_{ln} D_{\psi^{(ln)}} p_{ln} e_{ln}) + 2\varepsilon. \quad (4.14)$$

In the next step we will estimate the first term in the last line above. Consider the spectral decomposition of $p_{ln} D_{\psi^{(ln)}} p_{ln}$ into one-dimensional projections

$$p_{ln} D_{\psi^{(ln)}} p_{ln} = \sum_{i=1}^{b_{ln}} \lambda_{i,ln} r_{i,ln},$$

where $b_{ln} = \mathrm{tr}(p_{ln})$. For $\delta > 0$ we define the set

$$T_{\delta,ln} := \{i \in \{1, \ldots, b_{ln}\} : \lambda_{i,ln} \leq e^{-ln(S(\psi) - \delta)}\},$$

and

$$r_{\delta,ln} := \sum_{i \in T_{\delta,ln}} r_{i,ln}.$$

Then we have

$$\lim_{n\to\infty} \psi^{(ln)}(p_{ln} - r_{\delta,ln}) = 0 \qquad \text{for all } \delta > 0, \tag{4.15}$$

by Lemma 3.1.7. The limit assertion (4.15) implies the following estimate for sufficiently large $n \in \mathbb{N}$:

$$\begin{aligned} \operatorname{tr}(p_{ln} D_{\psi^{(ln)}} p_{ln} e_{ln}) &\leq e^{-ln(S(\psi)-\delta)} \sum_{i\in T_{\delta,ln}} \operatorname{tr}(r_{i,ln} e_{ln}) + \varepsilon \\ &\leq e^{-ln(S(\psi)-\delta)} \sum_{i=1}^{b_{ln}} \operatorname{tr}(r_{i,ln} e_{ln}) + \varepsilon \\ &= e^{-ln(S(\psi)-\delta)} \operatorname{tr}(p_{l\mathbf{n}} e_{l\mathbf{n}}) + \varepsilon. \end{aligned} \tag{4.16}$$

Note that each projection p_{ln} can be represented as a sum of minimal projections $s_{i,n} \in \mathcal{D}_l^{(n)}$:

$$p_{ln} = \sum_{i\in I(p_{ln})} s_{i,n},$$

for some index set $I(p_{ln})$. Recall that by construction the operators $s_{i,n}$ are projections onto eigensubspaces of $D_{\varphi^{(ln)}}$. Let $\{\mu_{i,n}\}_{i=1}^{a_l^n}$ be the eigen-values corresponding to the set of minimal projections $\{s_{i,n}\}_{i=1}^{a_l^n}$. Then by (4.10) we have

$$\begin{aligned} \operatorname{tr}(p_{ln} e_{ln}) &\leq e^{n(S(\psi,\varphi)l+(S(\psi)+\eta)l-E_l+2\delta)} \sum_{i\in I(p_{ln})} \varphi^{(ln)}(s_{i,n}) \operatorname{tr}(s_{i,n} e_{ln}) \\ &= e^{n(S(\psi,\varphi)l+(S(\psi)+\eta)l-E_l+2\delta)} \sum_{i\in I(p_{ln})} e^{n(\frac{1}{n}\log \operatorname{tr}(s_{i,n}))} \mu_{i,n} \operatorname{tr}(s_{i,n} e_{ln}) \\ &\leq e^{n(S(\psi,\varphi)l+(S(\psi)+\eta)l+3\delta)} \varphi^{(ln)}(p_{ln} e_{ln}) \qquad (\text{by } (4.11)) \\ &\leq e^{n(S(\psi,\varphi)l+(S(\psi)+\eta)l+3\delta)} \varphi^{(ln)}(e_{ln}) \\ &= e^{n(S(\psi,\varphi)l+(S(\psi)+\eta)l+3\delta)} \varphi^{(y)}(t_y) \qquad (\text{by } (4.13)) \\ &\leq e^{n(S(\psi,\varphi)l+(S(\psi)+\eta)l+3\delta)} e^{-y(S(\psi,\varphi)+a)} \qquad (\text{by } (4.7)), \end{aligned} \tag{4.17}$$

Taking into account (4.16) and (4.17) we obtain

$$\operatorname{tr}(p_{ln} D_{\psi^{(ln)}} p_{ln} e_{ln}) \leq e^{-ln((\frac{y}{ln}-1)S(\psi,\varphi)+\frac{y}{ln}a-\delta-\eta-\frac{3\delta}{l})} + \varepsilon.$$

Observe that $\lim_{y\to\infty} \frac{y}{ln} = 1$ and $a > 0$ hold. Hence if $\delta, \eta > 0$ are chosen appropriately small we can achieve that the exponent in the last inequality becomes negative for large n. This fact combined with (4.14) and (4.13) shows that

$$\lim_{y\to\infty} \psi(t_y) = 0$$

holds. This limit assertion toghether with (4.6) shows (4.2). □

4.2 Monotonicity of the Quantum Relative Entropy

Let $T : \mathcal{A} \to \mathcal{A}$ be a unital, completely positive linear map of a finite dimensional C^*-algebra $\mathcal{A}$ into itself. There is a dual map $T_\#$ with respect to the inner product defined by the unique trace tr (which extends canonically to the state space $\mathcal{S}(\mathcal{A})$) given by

$$\mathrm{tr}(D_\psi T(a)) = \mathrm{tr}(T_\#(D_\psi)a) \quad \text{for all } a \in \mathcal{A} \text{ and all } \psi \in \mathcal{S}(\mathcal{A}).$$

In [26] Stinespring showed that for each unital, completely positive map T there exist a finite dimensional Hilbert space $\mathcal{H}_a$, a state σ on $\mathcal{B}(\mathcal{H}_a)$ and a unitary operator $U : \mathcal{H} \otimes \mathcal{H}_a \to \mathcal{H} \otimes \mathcal{H}_a$ such that

$$T_\#(D_\psi) = \mathrm{tr}_{\mathcal{H}_a}(U(D_\psi \otimes D_\sigma)U^*), \tag{4.18}$$

where $\mathrm{tr}_{\mathcal{H}_a}$ denotes the partial trace with respect to the Hilbert space $\mathcal{H}_a$.

Theorem 4.2.1 (Monotonicity of the relative entropy) *Let ψ, φ be states on $\mathcal{A}$ and let $T : \mathcal{A} \to \mathcal{A}$ be a unital, completely positive map of $\mathcal{A}$ into itself. Then it holds*

$$S(\psi, \varphi) \geq S(\psi \circ T, \varphi \circ T), \tag{4.19}$$

or equivalently

$$S(D_\psi, D_\varphi) \geq S(T_\#(D_\psi), T_\#(D_\varphi)). \tag{4.20}$$

Proof of Theorem 4.2.1: The monotonicity holds trivially in the case $S(\psi, \varphi) = \infty$. We consider the case $S(\psi, \varphi) < \infty$ and define two states $\tilde{\psi}, \tilde{\varphi}$ on $\mathcal{A} \otimes \mathcal{B}(\mathcal{H}_a)$ by their density operators:

$$D_{\tilde{\psi}} := U(D_\psi \otimes D_\sigma)U^* \text{ and } D_{\tilde{\varphi}} := U(D_\varphi \otimes D_\sigma)U^*,$$

where the unitary operator $U : \mathcal{H} \otimes \mathcal{H}_a \to \mathcal{H} \otimes \mathcal{H}_a$ and the density operator $D_\sigma \in \mathcal{B}(\mathcal{H}_a)$ are chosen according to the representation (4.18) of the completely positive, unital map T. Observe that according to the representation (4.18) we have

$$D_{\tilde{\psi}\restriction\mathcal{A}} = \mathrm{tr}_{\mathcal{H}_a}(U(D_\psi \otimes D_\sigma)U^*) = T_\#(D_\psi). \tag{4.21}$$

and

$$D_{\tilde{\varphi}\restriction\mathcal{A}} = T_\#(D_\varphi). \tag{4.22}$$

Consider the quasilocal algebra $(\mathcal{A} \otimes \mathcal{B}(\mathcal{H}_a))^\infty$ and the stationary product states $\tilde{\psi}_\infty$ with $\tilde{\psi}^{(1)} = \tilde{\psi}$, $\tilde{\varphi}_\infty$ with $\tilde{\varphi}^{(1)} = \tilde{\varphi}$ on $(\mathcal{A} \otimes \mathcal{B}(\mathcal{H}_a))^\infty$. Note that the algebra $\mathcal{A}^\infty$ is a subalgebra of $(\mathcal{A} \otimes \mathcal{B}(\mathcal{H}_a))^\infty$. The definition (4.1) implies that

$$\beta_{\varepsilon,n}(\tilde{\psi}_\infty \restriction \mathcal{A}^\infty, \tilde{\varphi}_\infty \restriction \mathcal{A}^\infty) \geq \beta_{\varepsilon,n}(\tilde{\psi}_\infty, \tilde{\varphi}_\infty),$$

and hence by Theorem 4.1.1

$$\begin{aligned} -S(\tilde{\psi} \restriction \mathcal{A}, \tilde{\varphi} \restriction \mathcal{A}) &= \lim_{n\to\infty} \frac{1}{n} \beta_{\varepsilon,n}(\tilde{\psi}_\infty \restriction \mathcal{A}^\infty, \tilde{\varphi}_\infty \restriction \mathcal{A}^\infty) \\ &\geq \lim_{n\to\infty} \frac{1}{n} \beta_{\varepsilon,n}(\tilde{\psi}_\infty, \tilde{\varphi}_\infty) \\ &= -S(\tilde{\psi}, \tilde{\varphi}). \end{aligned}$$

Using (4.21) and (4.22) this last inequality chain shows that

$$\begin{aligned} S(T_\#(D_\psi), T_\#(D_\varphi)) &\leq S(\tilde{\psi}, \tilde{\varphi}) \\ &= S(U(D_\psi \otimes D_\sigma)U^*, U(D_\varphi \otimes D_\sigma)U^*) \\ &= S(D_\psi, D_\varphi), \end{aligned}$$

since the relative entropy is invariant under the action of the unitaries and simultaneous tensor multiplication with D_σ. $\square$

Appendix A

The Hiai-Petz Approximation

In this appendix we give a proof of the result established by Hiai and Petz in [8]. At the same time we simplify slightly their arguments (cf.[5]). Let φ be a stationary product state on the quasilocal algebra $\mathcal{A}^\infty$. Consider the spectral resolution of $D_{\varphi^{(1)}}$

$$D_{\varphi^{(1)}} = \sum_{i=1}^{d} \lambda_i e_i,$$

where λ_i are the eigenvalues and e_i are projections onto a complete set of eigenvectors of $D_{\varphi^{(1)}}$. The density operator $D_{\varphi^{(n)}}$ can be written as

$$D_{\varphi^{(n)}} = \sum_{i_1,\ldots,i_n=1}^{d} \lambda_{i_1} \cdot \ldots \cdot \lambda_{i_n} e_{i_1} \otimes \ldots \otimes e_{i_n},$$

or after collecting all one dimensional projections $e_{i_1} \otimes \ldots \otimes e_{i_n}$ which correspond to the same eigenvalue of $D_{\varphi^{(n)}}$ as

$$D_{\varphi^{(n)}} = \sum_{n_1,\ldots n_d : n_1+\ldots+n_d=n} \left(\prod_{k=1}^{d} \lambda_k^{n_k} \right) p_{n_1,\ldots n_d}, \tag{A.1}$$

with

$$p_{n_1,\ldots n_d} := \sum_{(i_1,\ldots,i_n)\in I_{n_1,\ldots n_d}} e_{i_1} \otimes \ldots \otimes e_{i_n},$$

where

$$I_{n_1,\ldots n_d} := \{(i_1,\ldots,i_n) : \#\{j : i_j = k\} = n_k \text{ for } 1 \le k \le d\}.$$

We define the map

$$E_n : \mathcal{A}^{(n)} \to \bigoplus_{n_1,\ldots,n_d:\ n_1+\ldots+n_d=n} p_{n_1,\ldots,n_d} \mathcal{A}^{(n)} p_{n_1,\ldots,n_d},$$

by

$$E_n(a) := \sum_{n_1,\dots,n_d:\ n_1+\dots+n_d=n} p_{n_1,\dots,n_d} a\, p_{n_1,\dots,n_d}. \tag{A.2}$$

The next theorem is due to Hiai and Petz, Lemma 3.1 and Lemma 3.2 in [8]. The following simplified proof due to Ra. Siegmund-Schultze and this author and is from [5].

Theorem A.0.2 (Hiai and Petz) *If ψ is a stationary state on $\mathcal{A}^\infty$ and $\mathcal{D}_n$ is the abelian subalgebra of $\mathcal{A}^{(n)}$ generated by $\{p_{n_1,\dots n_d} D_{\psi^{(n)}} p_{n_1,\dots n_d}\}_{n_1,\dots,n_d} \cup \{p_{n_1,\dots n_d}\}_{n_1,\dots,n_d}$ then*

$$S(\psi^{(n)},\varphi^{(n)}) = S(\psi^{(n)} \restriction \mathcal{D}_n, \varphi^{(n)} \restriction \mathcal{D}_n) + S(\psi^{(n)} \circ E_n) - S(\psi^{(n)}), \tag{A.3}$$

and

$$S(\psi^{(n)} \circ E_n) - S(\psi^{(n)}) \le d \log(n+1), \tag{A.4}$$

and consequently we have

$$\lim_{n\to\infty} \frac{1}{n} S(\psi^{(n)} \restriction \mathcal{D}_n, \varphi^{(n)} \restriction \mathcal{D}_n) = s(\psi,\varphi). \tag{A.5}$$

Proof of Theorem A.0.2: We will prove the theorem only in the case where the density operator $D_{\varphi^{(1)}}$ is invertible, this extends trivially to the general situation. Since the algebra $\mathcal{D}_n$ is abelian it is representable as:

$$\mathcal{D}_n = \bigoplus_{i=1}^{a_n} \mathbb{C} \cdot q_{i,n}. \tag{A.6}$$

Using the definition of the algebra $\mathcal{D}_n$ and the representation (A.6) it can be easily seen that

$$D_{\psi^{(n)} \restriction \mathcal{D}_n} = \sum_{i=1}^{a_n} \mathrm{tr}(D_{\psi^{(n)}} q_{i,n}) q_{i,n}, \tag{A.7}$$

i.e. the density operator $D_{\psi^{(n)} \restriction \mathcal{D}_n} \in \mathcal{D}_n$ corresponding to $\psi^{(n)} \restriction \mathcal{D}_n$ with respect to the trace $\mathrm{tr}_{\mathcal{B}_n}$ of $\mathcal{D}_n$ is given by (A.7). Similarily we have also

$$D_{\varphi^{(n)} \restriction \mathcal{D}_n} = \sum_{i=1}^{a_n} \mathrm{tr}(D_{\varphi^{(n)}} q_{i,n}) q_{i,n}. \tag{A.8}$$

We have the following equation for the difference of the relative entropies:

$$\begin{aligned} S(D_{\psi^{(n)}}, D_{\varphi^{(n)}}) \;&-\; S(D_{\psi^{(n)} \restriction \mathcal{D}_n}, D_{\varphi^{(n)} \restriction \mathcal{D}_n}) = \\ &-\; S(D_{\psi^{(n)}}) - \mathrm{tr}(D_{\psi^{(n)}} \log D_{\varphi^{(n)}}) + S(D_{\psi^{(n)} \restriction \mathcal{D}_n}) \\ &+\; \mathrm{tr}_{\mathcal{D}_n}(D_{\psi^{(n)} \restriction \mathcal{D}_n} \log D_{\varphi^{(n)} \restriction \mathcal{D}_n}). \end{aligned}$$

Since $D_{\varphi^{(n)}}$ is invertible $\log D_{\varphi^{(n)}\upharpoonright\mathcal{D}_n}$ is well defined and contained in $\mathcal{D}_n$. All projections $q_{i,n}$ commute with $\log D_{\varphi^{(n)}}$. Thus we obtain

$$\begin{aligned}\mathrm{tr}_{\mathcal{D}_n}(D_{\psi^{(n)}\upharpoonright\mathcal{D}_n}\log D_{\varphi^{(n)}\upharpoonright\mathcal{D}_n}) &= \mathrm{tr}(D_{\psi^{(n)}}\log D_{\varphi^{(n)}}) \\ &+ \sum_{i=1}^{a_n}\mathrm{tr}(D_{\psi^{(n)}}q_{i,n})\log\mathrm{tr}(q_{i,n}), \quad \text{(A.9)}\end{aligned}$$

where we have used (A.7) and (A.8). And hence

$$\begin{aligned}S(D_{\psi^{(n)}},D_{\varphi^{(n)}}) - S(D_{\psi^{(n)}\upharpoonright\mathcal{D}_n},D_{\varphi^{(n)}\upharpoonright\mathcal{D}_n}) &= S(D_{\psi^{(n)}\upharpoonright\mathcal{D}_n}) - S(D_{\psi^{(n)}}) \\ &+ \sum_{i=1}^{a_n}\mathrm{tr}(D_{\psi^{(n)}}q_{i,n})\log\mathrm{tr}(q_{i,n}).\end{aligned} \tag{A.10}$$

On the other hand we can easily see that

$$E_n(D_{\psi^{(n)}}) = \sum_{i=1}^{a_n}\frac{\mathrm{tr}(D_{\psi^{(n)}}q_{i,n})}{\mathrm{tr}(q_{i,n})}q_{i,n},$$

and therefore

$$S(E_n(D_{\psi^{(n)}})) = S(D_{\psi^{(n)}\upharpoonright\mathcal{D}_n}) + \sum_{i=1}^{a_n}\mathrm{tr}(D_{\psi^{(n)}}q_{i,n})\log\mathrm{tr}(q_{i,n}).$$

Hence by (A.10) we have established the first equality of the theorem:

$$S(\psi^{(n)},\varphi^{(n)}) = S(\psi^{(n)}\upharpoonright\mathcal{D}_n,\varphi^{(n)}\upharpoonright\mathcal{D}_n) + S(\psi^{(n)}\circ E_n) - S(\psi^{(n)}).$$

In order to establish (A.4) we need the following entropy inequality:

$$S(\sum_{i=1}^{k}a_iD_i) \le \sum_{i=1}^{k}a_iS(D_i) + H(\{a_i\}_{i=1}^{k}),$$

where D_i are density operators, a_i are non negative real numbers summing up to 1 and $H(\{a_i\}_{i=1}^{k})$ denotes the Shannon entropy of the probability distribution $\{a_i\}_{i=1}^{k}$. An elementary proof of this inequality can be found in [17]. Consider a spectral resolution of the density operator $D_{\psi^{(n)}} = \sum_{i=1}^{d^n}\mu_{i,n}p_{i,n}$ into one-dimensional projections. Then we have

$$\begin{aligned}S(E_n(D_{\psi^{(n)}})) - S(D_{\psi^{(n)}}) &\le \sum_{i=1}^{d^n}\mu_{i,n}S(E_n(p_{i,n})) + H(\{\mu_{i,n}\}_{i=1}^{d^n}) \\ &- S(D_{\psi^{(n)}}) \\ &= \sum_{i=1}^{d^n}\mu_{i,n}S(E_n(p_{i,n})) \\ &\le d\log(n+1),\end{aligned}$$

since for each projection $p_{i,n}$ the operator $p_{n_1,\ldots,n_d}(p_{i,n})p_{n_1,\ldots,n_d}$ has at most one dimensional range and hence the density operator $E_n(p_{i,n})$ can have at most $(n+1)^d$ dimensional range. Here we used an obvious estimate for the cardinality of partitions of the number n into d numbers. $\square$

Appendix B

Proof of Theorem 3.1.3

We fix an $\mathbf{n} \in \mathbb{N}^\nu$ and consider the following expression for $\omega_\mathbf{n} \in A^{(\mathbf{n})}$:

$$d_\mathbf{n}(\omega_\mathbf{n}) := \frac{1}{\#\Lambda(\mathbf{n})} \log(P^{(\mathbf{n})}(\omega_\mathbf{n})) - \frac{1}{\#\Lambda(\mathbf{n})} \log(Q^{(n)}(\omega_\mathbf{n})). \tag{B.1}$$

Consider first the case $D_M(P,Q) < \infty$.
The Shannon-McMillan-Breiman Theorem shows that the first term on the right hand side of (B.1) converges $P-$ almost surely to the Kolmogorov-Sinai entropy of P. It is well known that for the ergodic measures the Kolmogorov-Sinai entropy and the Shannon mean entropy $h(P)$ coincide. Hence we have

$$\lim_{n \nearrow \mathbb{N}^\nu} \frac{1}{\#\Lambda(\mathbf{n})} \log P^{(\mathbf{n})}(\omega_\mathbf{n}) = h(P) \qquad P - \text{almost surely.} \tag{B.2}$$

On the other hand, the condition $D_M(P,Q) < \infty$ implies that the support of $Q^{(1)}$ contains the support of $P^{(1)}$. Indeed, this immediately follows from

$$D_M(P,Q) = \sup_\mathbf{n} \frac{1}{\#\Lambda(\mathbf{n})} D(P^{(\mathbf{n})}, Q^{(\mathbf{n})}).$$

Let x_1 denote the projection onto the coordinate corresponding to the site 0. According to the individual ergodic theorem we can conclude that

$$\begin{aligned} \frac{1}{\#\Lambda(\mathbf{n})} \log(Q^{(\mathbf{n})}(\omega_\mathbf{n})) &= \frac{1}{\#\Lambda(\mathbf{n})} \sum_{\mathbf{k} \in \Lambda(\mathbf{n})}^{n} \log Q^{(1)}(x_1(T(\mathbf{k})\omega)) \\ &\to E_P(\log(Q^{(1)} \circ x_1)) \end{aligned} \tag{B.3}$$

$P-$ almost surely as $\mathbf{n} \nearrow \mathbb{N}^\nu$, where we have used the assumption that Q is a stationary product measure. But obviously

$$E_P(\log(Q^{(1)} \circ x_1)) = \sum_{a \in A} P^{(1)}(a) \log Q^{(1)}(a)$$

holds. Thus (B.2) and (B.3) imply that the quantity $d_{\mathbf{n}}$ defined in (B.1) satisfies

$$\lim_{n \nearrow \mathbb{N}^\nu} d_{\mathbf{n}}(\omega_{\mathbf{n}}) = D_M(P, Q) \quad P - \text{almost surely},$$

which in turn proves our claim (3.2).
On the other hand, the condition $D_M(P, Q) = \infty$ implies the existence of an $a_0 \in A$ such that $Q^{(1)}(a_0) = 0$ and $P^{(1)}(a_0) > 0$. This shows that

$$\frac{1}{\#\Lambda(\mathbf{n})} \sum_{\mathbf{k} \in \Lambda(\mathbf{n})}^{n} -\log Q^{(1)}(x_1(T(\mathbf{k})\omega)) \to \infty \quad P - \text{almost surely}.$$

This in turn leads to

$$\lim_{n \nearrow \mathbb{N}^\nu} d_{\mathbf{n}}(\omega_{\mathbf{n}}) = \infty = D_M(P, Q) \quad P - \text{almost surely},$$

and we are done. $\square$

Bibliography

[1] R. Bhatia, Matrix Analysis, Graduate Texts in Mathematics 169, Springer, New York 1997

[2] I. Bjelaković, T. Krüger, Ra. Siegmund-Schultze, A. Szkoła, The Shannon-McMillan Theorem for Ergodic Quantum Lattice Systems, *Invent. Math.* 155, 203-222 (2003)

[3] I. Bjelaković, T. Krüger, Ra. Siegmund-Schultze, A. Szkoła, Chained Typical Subspaces - a Quantum Version of Breiman's Theorem, arXiv.org, quant-ph/0301177

[4] I. Bjelaković, Ra. Siegmund-Schultze, An Ergodic Theorem for the Quantum Relative Entropy, arXiv.org, quant-ph/0306094, to appear in *Commun. Math. Phys.*

[5] I. Bjelaković, Ra. Siegmund-Schultze, A New Proof of the Monotonicity of Quantum Relative Entropy for Finite Dimensional Systems, arXiv.org, quant-ph/0307170

[6] I. Bjelaković, A. Szkoła, The Data Compression Theorem for Ergodic Quantum Information Sources, to appear in *Quant. Inf. Proc.* (2003/2004)

[7] O. Bratteli, D. W. Robinson, Operator Algebras and Quantum Statistical Mechanics I, Springer, New York 1979

[8] F. Hiai, D. Petz, The Proper Formula for Relative Entropy and its Asymptotics in Quantum Probability, *Commun. Math. Phys.* 143, 99-114 (1991)

[9] F. Hiai, D. Petz, Entropy Densities for Algebraic States, *J. Functional Anal.* 125, 287-308 (1994)

[10] R.B. Israel, Convexity in the Theory of Lattice Gases, Princeton, New Jersey 1979

[11] R. Jozsa, B. Schumacher, A New Proof of the Quantum Noisless Coding theorem, *Journal of Modern Optics*, Vol.41, No.12, 2343-2349, (1994)

[12] A. Kaltchenko, E. H. Yang, Universal Compression of Ergodic Quantum Sources, *Quant. Inf. and Comput.* 3, 359-375 (2003)

[13] K. Kraus, States, Effects and Operations. Fundamental Notions of Quantum Theory, Lecture Notes in Physics 190, Springer-Verlag, Berlin 1983

[14] E. H. Lieb, M. B. Ruskai, Proof of the Strong Subadditivity of Quantum- Mechanical Entropy, *J. Math. Phys.* 14, 1938-1941 (1973)

[15] G. Lindblad, Completely Positive Maps and Entropy Inequalities, *Commun. Math. Phys.* 40, 147-151 (1975)

[16] E. Lindenstrauss, Pointwise Theorems for Amenable Groups, *Invent. Math.* 146, 259-295 (2001)

[17] M. A. Nielsen, I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge 2000

[18] M. Ohya, D. Petz, Quantum Entropy and its Use, Springer, Berlin 1993

[19] T. Ogawa, H. Nagaoka, Strong Converse and Stein's Lemma in Quantum Hypothesis Testing, *IEEE Trans. Inf. Th.*, vol. 46, No. 7, 2428-2433 (2000)

[20] D. Ornstein, B. Weiss, The Shannon-McMillan-Breiman Theorem for a Class of Amenable Groups, *Israel J. Math.* 44 (1), 53-60 (1983)

[21] D. Petz, M. Mosonyi, Stationary Quantum Source Coding, *J. Math. Phys.* 42, 4857-4864 (2001)

[22] D. Ruelle, Statistical Mechanics, W.A. Benjamin, New York 1969

[23] P. C. Shields, Two divergence-rate counterexamples, *J. Theor. Prob.* vol. 6 521-545 (1993)

[24] P.C. Shields, The Ergodic Theory of Discrete Sample Paths, American Mathematical Society 1996

[25] B. Simon, Trace Ideals and Their Application, Cambridge University Press, Cambridge 1979

[26] W. F. Stinespring, Positive Functions on C^*-Algebras, *Proc. Amer. Math. Soc*, 6, 211-216 (1955)

[27] A. Uhlmann, Relative Entropy and the Wigner-Yanase-Dyson-Lieb Concavity in an Interpolation Theory, *Commun. Math. Phys.* 54, 21-32 (1977)