The Shannon-McMillan Theorem and Related Results for Ergodic Quantum Spin Lattice Systems

and Applications in Quantum Information Theory

vorgelegt von

Diplom-Physikerin

Arleta Szkoła

Berlin

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. M. Scheutzow
Berichter:	Prof. Dr. R. Seiler
Berichter:	Prof. Dr. R. F. Werner

Tag der wissenschaftlichen Aussprache: 12. Februar 2004

Berlin 2004

D 83

Abstract

The aim of this thesis is to formulate and prove quantum extensions of the famous Shannon-McMillan theorem and its stronger version due to Breiman.

In ergodic theory the Shannon-McMillan-Breiman theorem is one of the fundamental limit theorems for classical discrete dynamical systems. It can be interpreted as a special case of the individual ergodic theorem. In this work, we consider spin lattice systems, which can be interpreted as dynamical systems under the action of the translation group. The Shannon-McMillan-Breiman theorem states that the Shannon entropy rate of an ergodic lattice system is the asymptotical rate of exponential decrease of probability of almost each individual spin configuration. In information theory, information sources are usually modeled by time-discrete stochastic processes or equivalently by 1-dimensional spin lattice systems. There, this theorem plays an important role, giving an interpretation to the Shannon entropy rate as the asymptotically mean information per signal. At the same time the entropy rate is an achievable lower bound for the compression rate of asymptotically error-free data compression algorithms.

It turns out, that there are analogues of the classical Shannon-McMillan theorem and Breiman's extension for quantum spin lattice systems, modeled as C*-dynamical systems with respect to the action of the translation group on a quasi-local C*-algebra. There, the concept of Shannon entropy for discrete probability distributions is generalized by the von Neumann entropy for density operators. A number of results, related to the quantum Shannon-McMillan-(Breiman) theorem are presented in this work. Similarly to classical information theory, the existence of asymptotically error-free data compression schemes for ergodic quantum sources is proven based on the quantum Shannon-McMillan theorem. There, the achievable lower bound on the compression rate is given by the von Neumann entropy rate of the quantum source. Furthermore, a structure theorem is proven, which describes the convex decomposition of ergodic states on quantum spin lattice systems into components which are ergodic with respect to some subgroup of the whole translation group.

Subsuming, we may say, that the presented results about quantum spin lattice systems establish the von Neumann entropy rate as the generalization of the Shannon entropy rate to the quantum case.

Zusammenfassung

Die vorliegende Dissertation hat zum Ziel gehabt, Quantenversionen des bekannten Shannon-McMillan Satzes und seiner auf Breiman zurückgehenden Verschärfung zu formulieren und zu beweisen.

In der Ergodentheorie ist der Shannon-McMillan-Breiman Satz einer der fundamentalen Grenzwertsätze für klassische diskrete dynamische Systeme und kann als ein Spezialfall des individuellen Ergodensatzes aufgefasst werden. Im Rahmen dieser Arbeit wurden Spingittersysteme behandelt, die mit der Wirkung der Translationsgruppe dynamische Systeme darstellen. Der Shannon-McMillan-Breiman Satz besagt, dass die Shannon-Entropierate ergodischer Gittersysteme die asymptotische Rate angibt, mit der die individuellen Wahrscheinlichkeiten fast jeder Spinkonfiguration exponentiell schnell abfallen. Der Satz spielt insbesondere in der klassischen Informationstheorie eine zentrale Rolle, wo Informationsquellen durch zeitdiskrete stochastische Prozesse bzw. äquivalent dazu durch eindimensionale Spingittersysteme modelliert werden. Ihm zufolge kann die Shannon-Entropierate ergodischer Informationsquellen als asymptotisch mittlere Information pro Signal interpretiert werden. Gleichzeitig gibt die Entropierate die erreichbare untere Schranke an die Komprimierungsrate asymptotisch fehlerfrei arbeitender Datenkomprimierungsalgorithmen an.

Es stellt sich heraus, dass für Quantenspingittersysteme, verstanden als C^*-dynamische Systeme bzgl. der Wirkung der Translationsgruppe auf einer quasilokalen C^*-Algebra, Analoga des klassischen Shannon-McMillan-Satzes und seiner auf Breiman zurückgehenden Verschärfung existieren. Dabei wird das Konzept der Shannon-Entropie für diskrete Wahrscheinlichkeitsverteilungen durch die für Dichteoperatoren definierte von Neumann-Entropie verallgemeinert. Eine Reihe von Resultaten, die mit dem Quanten-Shannon-McMillan-(Breiman-)Satz verwandt sind, werden in der Arbeit vorgestellt. Insbesondere konnte ähnlich wie in der klassischen Informationstheorie auf der Grundlage des Shannon-McMillan-Satzes die Existenz von asymptotisch fehlerfrei arbeitender Datenkomprimierungsschemen für ergodische Quanteninformationsquellen bewiesen werden. Dabei ist die erreichbare untere Schranke an die Komprimierungsrate durch die von Neumann-Entropierate der Quantenquelle gegeben. Des weiteren wurde ein Struktursatz über die Zerlegung ergodischer Quantenspingittersysteme in Komponenten, die bzgl. einer Translationsuntergruppe ergodisch sind, bewiesen.

Zusammenfassend lässt sich sagen, dass die in der Dissertation präsentierten Resultate für Quantenspingittersysteme die von Neumann-Entropierate als Verallgemeinerung der Shannon-Entropierate im informationstheoretischen Kontext bestätigen.

Contents

Chapter 1

Introduction

The aim of the present work is to formulate and prove quantum extensions of theorems for classical spin lattice systems, which had in the past a strong impact on classical information theory. Actually, applications to quantum information theory, in particular in the context of data compression, are one of our main motivations. We will concentrate on the Shannon-McMillan theorem (SM-theorem), which is usually referred to as the AEP (Asymptotic Equipartition Property) in classical information theory, and its stronger version, the Shannon-McMillan-Breiman theorem (SMB-theorem). For ergodic classical spin lattice systems both theorems are convergence theorems with the limit equal to the mean (per lattice site limit) Shannon entropy[1]. The SM-theorem is a convergence in probability statement. The SMB-theorem is a stronger almost sure convergence statement. We will show that for ergodic quantum spin lattice systems analogous convergence theorems exist with the mean (again, per site limit) von Neumann entropy as limit. Subsequently, with the quantum SM-theorem at our disposal we will prove a lossless quantum data compression theorem.

Classical spin systems on a ν-dimensional lattice corresponding to the group $\mathbb{Z}^\nu$ are mathematically modeled by a probability measure P on a measurable space (A^∞, Σ), where A^∞ is a set consisting of A-valued realizations (or configurations in the notation of statistical mechanics) over the lattice $\mathbb{Z}^\nu$, A is a finite set assigned to each lattice site $\mathbf{x} \in \mathbb{Z}^\nu$, and Σ is a σ-algebra generated by the collection of cylinder sets in A^∞. In the 1-dimensional case the elements of A^∞ are just doubly infinite A-valued sequences. Translations of the lattice canonically induce a dynamics on (A^∞, Σ) by shifts on A^∞. Thus if $T(\mathbb{Z}^\nu)$ is a representation of the translation group by shifts on A^∞ then the quadruple $(A^\infty, \Sigma, P, T(\mathbb{Z}^\nu))$ is a dynamical system corresponding to a classical ν-dimensional spin lattice system.
Roughly speaking, the SM-theorem states that under the ergodicity condition for the probability measure P the probabilities of the entropy-typical subsets on v-boxes, which are regions of finite volume v in the infinite ν-dimensional lattice $\mathbb{Z}^\nu$, converge to 1 in the thermodynamical limit. An entropy-typical subset on

[1]The SM(B)-theorem can be formulated for more general ergodic dynamical systems, not necessary corresponding to a spin lattice system. Then the limit is given by the Kolmogorov-Sinai (dynamical) entropy, which is identical to the mean Shannon entropy in the case of spin lattice systems.

a v-box consists of all the realizations on this box that have individual probabilities of order $\exp(-vh)$. Hereby h is the mean (base e) Shannon entropy of P. As a consequence the entropy-typical subsets are of approximate size equal to $\exp(vh)$, for v large enough. The stronger version of the theorem, the SMB-theorem, asserts that h is P-almost sure the rate (per lattice site) at which the individual probabilities of realizations on v-boxes exponentially decrease as the volume v tends to ∞, again in the case of an ergodic system.[2] Summarizing we can say that an ergodic probability measure concentrates on a typical subset $T \in \Sigma$ of A^∞ consisting of entropy-typical elements.

In classical information theory, 1-dimensional spin lattice systems, which can be equivalently considered as stochastic processes, model information sources. The success of the concept of Shannon entropy in classical information theory is to a large extent based on the AEP of ergodic systems. As the entropy-typical subsets become subsets of most probability on long enough blocks of the lattice $\mathbb{Z}$, the ability to approximate their size in terms of the mean Shannon entropy has direct consequences for data compression: The mean Shannon entropy h can be interpreted as the average information content of ergodic information sources. Indeed, a properly designed compression scheme can represent an ergodic information source asymptotically reliably using not more than $h/\log 2$ bits per lattice-site. Moreover, h determines the optimal rate of compression in the sense that any compression of rate lower than $h/\log 2$ fails to be reliable. This is due to the fact that for ergodic processes h is a lower bound on the asymptotic rate of the logarithmic size of subsets of most probability, which is a direct consequence of the AEP.

In order to treat both classical and quantum spin lattice systems in the same mathematical framework one uses the C*-algebraic formalism. In this formalism ν-dimensional spin lattice systems are modeled by C*-dynamical systems of the form $(\mathcal{A}^\infty, \Psi, T(\mathbb{Z}^\nu))$, where $\mathcal{A}^\infty$ is a quasi-local C*-algebra constructed from a finite dimensional unital C*-algebra $\mathcal{A}$, Ψ is a state, i.e. a normalized positive linear functional on $\mathcal{A}^\infty$ and $T(\mathbb{Z}^\nu)$ is the representation of the translation group $\mathbb{Z}^\nu$ by shifts on $\mathcal{A}^\infty$. The elements of the algebra $\mathcal{A}$ are the mathematical objects associated with the observables of single spins located at the lattice sites, while the algebra $\mathcal{A}^\infty$ is the algebra of observables of the whole lattice system. The quantum case corresponds to a non-commutative $\mathcal{A}$, while an abelian algebra $\mathcal{A}$ describes the classical situation. For quantum spin lattice systems the C*-algebraic description is a rather standard formalism, (cf. [42],[14]), as is the concept of quasi-local algebras in local quantum field theory. In contrast, the classical spin lattice systems are usually described as dynamical systems $(A^\infty, \Sigma, P, T(\mathbb{Z}^\nu))$, as introduced above, or especially in the 1-dimensional case as stochastic processes. We will present the classical results in this familiar setting and explain the correspondence to the C*-algebraic description, which we will use in the quantum case.

As mentioned, the convergence assertion in the SM(B)-theorem holds under the ergodicity condition. Of course, in the quantum version of the SM(B)-theorem

[2]In the 1-dimensional situation we would simply say that in the ergodic case the individual probabilities along P-almost every trajectory decrease exponentially fast at an asymptotical rate given by h.

formulated as an extension of the commutative to the non-commutative case, we need the (quantum version of) ergodicity assumption as well. Consequently, the class of ergodic quantum states on $\mathcal{A}^\infty$ shall be the main object in the present work. The shifts on the quasi-local C*-algebra $\mathcal{A}^\infty$ are canonically induced by lattice translations. It is a well known fact that the set of translation invariant states $\mathcal{T}(\mathcal{A}^\infty)$ on the observable algebra $\mathcal{A}^\infty$ forms a weak*-compact convex subset of the entire state space for the quantum lattice system. States which are extremal points of $\mathcal{T}(\mathcal{A}^\infty)$ (the existence of extremal points is ensured by the compactness of $\mathcal{T}(\mathcal{A}^\infty)$) are called shift-ergodic or, relating to the underlying translation group, $\mathbb{Z}^\nu$-ergodic states of the quantum lattice system.
We emphasize that we do not care about concrete realizations of ergodicity. In the present work ergodicity always appears just as a mathematical assumption on the quantum states required to obtain some asymptotical behavior of the lattice systems. Nevertheless, the ergodicity assumption is a physically motivated and physically reasonable assumption. Ergodic systems form the elementary components of translation invariant physical systems.

As mentioned at the beginning it turns out that the mean von Neumann entropy s of ergodic quantum spin lattice systems plays a role analogous to the Shannon entropy rate of ergodic classical spin lattice systems. Indeed, the quantum version of the SM-theorem, Theorem 3.3.1, states that in each local algebra $\mathcal{A}_\Lambda$ of observables corresponding to quantum spins located in the v-box Λ in the lattice $\mathbb{Z}^\nu$ there exists a projector p such that each minimal projector from $\mathcal{A}_\Lambda$ dominated by p has an expectation value of order $\exp(-vs)$ and the expectation value of p is close to 1, for v large enough. The expectations are calculated with respect to an ergodic quantum state on a quasi-local C*-algebra $\mathcal{A}^\infty$. The trace of the projector p, or alternatively the dimension of the corresponding Hilbert subspace, is approximately equal to $\exp(vs)$, for large v.

One of the main ideas behind the proof of the quantum SM-theorem is to apply the classical SM-theorem to commutative C*-dynamical subsystems of the considered ergodic non-commutative (quantum) spin lattice system, where the classical subsystems approximate the quantum system in mean entropy. The problem occurring in such an approach is that the approximating classical subsystems inherit their ergodic properties from the ergodic properties of the given quantum system with respect to a subgroup of $\mathbb{Z}^\nu$, typically of the form $l \cdot \mathbb{Z}^\nu$ with $l > 1$ an integer. However, in general a state fails to be ergodic with respect to arbitrary subgroups if the only assumption we make is the $\mathbb{Z}^\nu$-ergodicity. To deal with this problem we will prove an ergodic decomposition statement, Theorem 3.2.1, which shows that the $l \cdot \mathbb{Z}^\nu$-ergodic decompositions of a $\mathbb{Z}^\nu$-ergodic quantum state are of a rather good controllable structure.

As an application of the quantum SM-theorem (quantum AEP) to quantum information theory we formulate a lossless quantum data compression theorem, Theorem 3.4.1. It establishes the mean von Neumann entropy s as an information quantifying entity. Indeed, disposing of the quantum AEP we can prove the existence of asymptotically reliable compression schemes for ergodic quantum information sources. The compression will occur by block coding, which uses for the representation of source blocks asymptotically not more than $s/\log 2$ qubits per source signal. Moreover, s is an optimal rate of compression in the sense that using a lower number of representing qubits than $s/\log 2$ results in a lossy reconstruction of the original quantum information source, at least if measuring the success of reconstruction by entanglement fidelity or ensemble fidelity, as we will do. We emphasize that our construction of optimal compression schemes depends on the quantum state of the information source. Recently, the existence of universal quantum data compression schemes, which work optimally for a whole class of ergodic quantum information sources, have been proven, [30]. The class is characterized by the same mean von Neumann entropy of the quantum states.

The classical SMB-theorem is usually formulated as an almost sure i.e. pointwise convergence statement, where the notions of individual realizations or trajectories appear. However, in the non-commutative setting such notions do not exist. In order to avoid the trajectory notion we use an equivalent finite form of the classical SMB-theorem, which is based on the observation, that due to the SMB-theorem for an ergodic source the entropy-typical subsets on different v-boxes are not independent or isolated, but are nested in a sense specified later. Then, starting with the finite reformulation we succeed in expanding the (reformulated) SMB-theorem to the quantum case, Theorem 3.5.1. In fact, the quantum extension of the SM-theorem is also based on a finite formulation of the classical SM-theorem, which is a standard formulation in classical information theory. In ergodic theory or probability theory the preferred formulation of the SM-theorem uses the notion of trajectories.

Most of the asymptotically faithful compression algorithms used in practice are not block coding but sequential coding algorithms. They are working on forthcoming data strings using the source data sequences to be compressed simultaneously for generating a codebook. Properly designed sequential compression algorithms can achieve an optimal rate of compression for arbitrary (not even) stationary information sources without the prior knowledge of their statistics (universal coding). In general the reason why classical sequential compression schemes like for example the popular Lempel-Ziv algorithm can operate efficiently can be seen in the fact, that entropy-typical subsets are nested. A further crucial point is that in the classical situation one can investigate the unknown statistics of the information source by measurements without disturbing the information source. However, in the quantum setting any measurement has some influence on the (unknown) quantum state of the information source, which cannot be predicted but can be estimated in the best case. The existence of a quantum version of the SMB-theorem can be seen as an indication for the existence of quantum sequential data compression algorithms. However, it is an open question how to deal with the remaining problem connected with quantum measurements and their disturbance of quantum states.

The present work is organized as follows: In the second chapter we review some facts about commutative shift-invariant lattice systems. The focus is on the SM-theorem and its extension due to Breiman. We give an equivalent reformulation of the SMB-theorem especially suited for extension to the quantum case, Lemma 2.2.5. Moreover, we formulate and prove a convergence assertion for a non-consistent family of probability distributions on finite sets, Lemma 2.2.6. We will meet such non-consistent families in connection with a consistent family of density operators corresponding one-to-one to an ergodic quantum state on a quasi-local algebra $\mathcal{A}^\infty$: The probability distributions given by the eigen-value distributions of the density operators are usually not consistent due to quantum correlations between single quantum spins. Finally in the second section classical data compression is briefly discussed as application of the reviewed theorems. In general, the selection of presented classical results is due to their relevance for the third chapter. There we prove the quantum Shannon-McMillan theorem 3.3.1 and a quantum version of the Shannon-McMillan-Breiman theorem 3.5.1, which are our main results. They are formulated as extensions of commutative case theorems to the non-commutative case of quantum ergodic spin lattice systems. Moreover, we prove an ergodic decomposition theorem for ergodic quantum states, Theorem 3.2.1, which is essential in the proof of the quantum SM-theorem and present the asymptotically lossless quantum data compression theorem 3.4.1 as an application of the quantum SM-theorem to quantum information theory. In the last chapter the results are summarized and possible continuation of the work is discussed.

Chapter 2

Classical Shift-Invariant Spin Lattice Systems

This chapter is intended to give an overview of results characterizing the asymptotic behavior of classical shift-invariant spin lattice systems as far as these are useful to prepare the next chapter treating quantum ergodic spin lattice systems. The focus will be on the Shannon-McMillan theorem and its extension due to Breiman. As these are well known results in classical theories such as ergodic theory, information theory or probability theory we will only give references of the proofs. Beside the standard formulations of the theorems we will also present equivalent reformulations especially suited to be extended to the quantum context. There, the main point will be to avoid the notion of trajectories appearing in the standard formulations as there are no quantum analogues to these objects. For the equivalence statement of the standard formulation and the finite reformulation of the Shannon-McMillan-Breiman theorem we will present the proof in detail. Another result will be a convergence assertion for a non-consistent family of discrete probability spaces.

2.1 Main Concepts

As mentioned in the introduction, classical ν-dimensional spin lattice systems, $\nu \in \mathbb{N}$, can be described mathematically in the framework of dynamical systems on an infinite cartesian product space A^∞ over the lattice $\mathbb{Z}^\nu$, where the dynamics is introduced by the action of the translation group $\mathbb{Z}^\nu$. In the 1-dimensional case translation invariant spin lattice systems, i.e. spin chains, correspond to stationary discrete-time stochastic processes with a finite state space and model information sources in classical information theory. In this section we review the main mathematical concepts concerning classical spin lattice systems, thereby fixing notations. For a more detailed introduction we suggest for example [42].

Mathematical Model for Classical ν-dimensional Spin Lattice Systems

The ν-dimensional infinitely extended lattice corresponds to the group $\mathbb{Z}^\nu$. Associated to each $\mathbf{x} \in \mathbb{Z}^\nu$ there is a set $A_{\mathbf{x}}$. In the sequel we consider only the case that every $A_{\mathbf{x}}$ is equal to a fixed finite set A. We consider the infinite cartesian product space

$$A^\infty := \mathsf{X}_{\mathbf{x}\in\mathbb{Z}^\nu} A_{\mathbf{x}}.$$

In the context of statistical mechanics the elements $\mathbf{a} \in A^\infty$ are usually called (spin) configurations on the lattice $\mathbb{Z}^\nu$. In the 1-dimensional case one refers to the doubly infinite sequences $\mathbf{a} \in A^\infty$ as trajectories or especially in the context of information theory as data strings . In the sequel, abusing notions, we will refer to $\mathbf{a} \in A^\infty$ as trajectories also in the general ν-dimensional case.
Consider a finite subset $\Lambda \subset \mathbb{Z}^\nu$. For $\mathbf{a} \in A^\infty$

$$[\mathbf{a}]_\Lambda := \{\mathbf{b} \in A^\infty | \; b_{\mathbf{x}} = a_{\mathbf{x}}, \; \mathbf{x} \in \Lambda\}$$

defines a cylinder set in A^∞. Similarly, each finite sequence $(a_{\mathbf{i}})_{\mathbf{i}\in\Lambda} =: a_\Lambda$ from a finite cartesian product set $A_\Lambda := \mathsf{X}_{\mathbf{x}\in\Lambda} A$ defines a cylinder set

$$[a_\Lambda] := \{\mathbf{b} \in A^\infty | \; b_{\mathbf{x}} = a_{\mathbf{x}}, \; \mathbf{x} \in \Lambda\} \subset A^\infty.$$

In the 1-dimensional case, if $\Lambda \subset \mathbb{Z}$ is a discrete interval $[m, n] := \{m, \ldots, n\}$, $m, n \in \mathbb{Z}$ with $m \leq n$, we will use the simplifying notation a_m^n for a finite sequence $(a_i)_{i=m}^n \in A_{[m,n]}$. Analogously, for $\mathbf{a} \in A^\infty$ we will denote the cylinder set $[\mathbf{a}]_{[m,n]} \subset A^\infty$ by $[\mathbf{a}]_m^n$.

The product space A^∞ together with the σ-algebra Σ generated by the collection of cylinder sets form a measurable space (A^∞, Σ) . On this space we can define a probability measure P. It corresponds one-to-one to a consistent family of probability measures $\{P_\Lambda\}_{\Lambda\subset\mathbb{Z}^\nu}$ on the finite cartesian products $A_\Lambda \subset A^\infty$, respectively. Each P_Λ is determined by P through the relation:

$$P_\Lambda(a_\Lambda) := P([a_\Lambda]), \qquad \text{for all } a_\Lambda \in A_\Lambda, \tag{2.1}$$

i.e. P_Λ is the marginal distribution of P on A_Λ. By consistency of the family we mean that

$$P_\Lambda(a_\Lambda) = \sum_{b_{\Lambda'} \in A_{\Lambda'} : b_{\mathbf{x}} = a_{\mathbf{x}}, \forall \mathbf{x} \in \Lambda} P_{\Lambda'}(b_{\Lambda'}), \qquad \text{for all } a_\Lambda \in A_\Lambda,$$

is fulfilled for all finite subsets $\Lambda, \Lambda' \in \mathbb{Z}^\nu$ with $\Lambda \subseteq \Lambda'$.

Each $\mathbf{x} \in \mathbb{Z}^\nu$ defines a translation of the lattice $\mathbb{Z}^\nu$ and induces an automorphism $T_{\mathbf{x}}$ on (A^∞, Σ) given by:

$$(T_{\mathbf{x}}(\mathbf{a}))_{\mathbf{i}} = a_{\mathbf{i}+\mathbf{x}}, \qquad \mathbf{i} \in \mathbb{Z}^\nu, \; \mathbf{a} \in A^\infty$$

i.e. $\{T_{\mathbf{x}}\}_{\mathbf{x}\in\mathbb{Z}^\nu}$ is the action of the translation group $\mathbb{Z}^\nu$ by shifts on A^∞ and defines a discrete dynamics on (A^∞, Σ). The quadruple

$$(A^\infty, \Sigma, P, T(\mathbb{Z}^\nu)), \tag{2.2}$$

where $T(\mathbb{Z}^\nu)$ denotes the representation of the group $\mathbb{Z}^\nu$ by the shifts $T_{\mathbf{x}}$, describes a classical spin lattice system as a dynamical system. In the 1-dimensional case each shift T_x, $x \in \mathbb{Z}$, is obtained as the power T^x of the unit right shift $T := T_{x=1}$. Thus we will use the notation (A^∞, Σ, P, T) instead of (2.2).

For $\mathbf{n} = (n_1, \ldots, n_\nu) \in \mathbb{N}^\nu$ we denote by $\Lambda(\mathbf{n})$ the box in $\mathbb{Z}^\nu$ determined by

$$\Lambda(\mathbf{n}) := \{(x_1, \ldots, x_\nu) \in \mathbb{Z}^\nu |\ x_i \in \{0, \ldots, n_i - 1\},\ i \in \{1, \ldots, \nu\}\}, \tag{2.3}$$

and for $n \in \mathbb{N}$ we denote by $\Lambda(n)$ the hypercube

$$\Lambda(n) := \{\mathbf{x} \in \mathbb{Z}^\nu |\ \mathbf{x} \in \{0, \ldots, n-1\}^\nu\}.$$

In the following we simplify notations by defining for $\mathbf{n} \in \mathbb{N}^\nu$ (respectively for $n \in \mathbb{N}$)

$$A^{(\mathbf{n})} := A_{\Lambda(\mathbf{n})} = \mathsf{X}_{\mathbf{x} \in \Lambda(\mathbf{n})} A_{\mathbf{x}} \qquad \text{and} \qquad P^{(\mathbf{n})} := P_{\Lambda(\mathbf{n})}.$$

A probability measure P on (A^∞, Σ) is called translation invariant if it satisfies

$$P \circ T_{\mathbf{x}} = P, \quad \forall \mathbf{x} \in \mathbb{Z}^\nu. \tag{2.4}$$

Due to the invariance it is uniquely defined already by a sequence $\{P^{(\mathbf{n})}\}_{\mathbf{n} \in \mathbb{N}^\nu}$ of marginal distributions of the probability measure P on the finite cartesian products $A^{(\mathbf{n})}$, respectively, (cf. [43], Theorem I.1.2).
If G is a subgroup of $\mathbb{Z}^\nu$, typically $l \cdot \mathbb{Z}^\nu$ with $l \geq 1$ an integer, then we say that a probability measure P on (A^∞, Σ) is G-invariant , if it is invariant (in the sense of (2.4)) under the action of the translation subgroup G. Thus a translation invariant measure is also called $\mathbb{Z}^\nu$-invariant.

Spin Chains in Information Theory: Information Sources

In classical information theory discrete information sources (IS) are modeled by discrete-time stochastic processes, i.e. by sequences of random variables $\{X_i\}_{i \in \mathbb{Z}}$, each X_i taking values from a given set A, called alphabet. The group $\mathbb{Z}$ is associated to discrete times. A stationary stochastic process with alphabet A is completely determined by a sequence $\{P^{(n)}\}_{n \in \mathbb{N}}$, where each $P^{(n)}$, $n \in \mathbb{N}$, is a joint distribution of the finite sequences of random variables $(X_1, X_2, \ldots, X_n)$.

Now, consider the dynamical system (A^∞, Σ, P, T), which is associated to a shift-invariant classical spin chain, i.e. a 1-dimensional spin lattice system. Then the coordinate functions X_i, $i \in \mathbb{N}$:

$$X_i : A^\infty \longrightarrow A,\ \mathbf{a} \mapsto x_i(\mathbf{a}) := a_i,$$

are random variables forming a stationary stochastic process with joint distributions coinciding by construction with the probability measures $P^{(n)}$ on $A^{(n)}$, respectively.
Conversely, each stationary stochastic process with alphabet A has a representation as a dynamical system of the form (A^∞, Σ, P, T), where P is a translation invariant probability measure corresponding one-to-one to the consistent family of joint distributions of the stochastic process, (cf. Theorem I.1.2 in [43] for the uniqueness of P).

The representations of an IS by (A^∞, Σ, P, T) is known as the Kolmogorov representation or the Kolmogorov model for an IS. In the sequel we will think of classical IS interchangeably as spin chains or as stochastic processes. In analogy to the classical case quantum spin chains model quantum IS in quantum information theory.

Ergodicity

There are different equivalent definitions of the crucial notion of ergodicity. We recall one, which can be directly expanded in the next chapter to the case of non-commutative C*-dynamical systems corresponding to quantum spin lattice systems. Let G be a (discrete) group and $T(G)$ its representation by automorphisms on a measurable space $(\mathbf{A}, \Sigma)$. A G-invariant probability measure P on $(\mathbf{A}, \Sigma)$, i.e. $P \circ T(g) = P$ for all $g \in G$, is said to be *G-ergodic*, if any convex decomposition of P into G-invariant probability measures on $(\mathbf{A}, \Sigma)$ is trivial, i.e. a G-ergodic P is an extremal point in the convex set of G-invariant probability measures on $(\mathbf{A}, \Sigma)$. If the group G is evident we simply say that P is an ergodic probability measure.

Shannon Entropy Rate

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

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

where log denotes the natural logarithm. In the information theoretical setting it is more common to use $\log_2$, i.e. the base 2 logarithm, relating to the bit as the fundamental information unit. Thus, treating information theoretical topics we will switch to $\log_2$.
For any translation invariant probability measure P on the space (A^∞, Σ) over a ν-dimensional lattice system $\mathbb{Z}^\nu$ the limit

$$h(P) := \lim_{\mathbf{n} \to \infty} \frac{1}{|\mathbf{n}|} H(P^{(\mathbf{n})}), \tag{2.5}$$

exists, where for $\mathbf{n} = (n_1, \dots, n_\nu) \in \mathbb{N}^\nu$

$$|\mathbf{n}| := \prod_{i=1}^{\nu} n_i, \tag{2.6}$$

is the volume of the ν-dimensional box $\Lambda(\mathbf{n})$, $P^{(\mathbf{n})}$ denotes the marginal distribution of P on $\Lambda(\mathbf{n})$ and

$$\mathbf{n} \to \infty \ :\Longleftrightarrow \ \Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu, \tag{2.7}$$

i.e. $\mathbf{n} \to \infty$ implies that $n_i \to \infty$, for all $i \in \{1, \dots, \nu\}$. We refer to the limit (2.5) as *mean Shannon entropy* of P. It exists due to the subadditivity of Shannon entropy (cf. [43]).
In the 1-dimensional case $h(P) = \lim_{n \to \infty} \frac{1}{n} H(P^{(n)})$ is usually called the *Shannon entropy rate* of P.

Now, we remind briefly, how the celebrated Shannon entropy notion arises from the more general concept of Kolmogorov-Sinai (dynamical) entropy (KS-entropy) . For the sake of simplicity we restrict ourselves to the 1-dimensional case only.

The KS-entropy is introduced in ergodic theory as a function of pairs (P, T), where P is a probability measure and T is a transformation, both defined on a given measurable space $(\mathbf{A}, \Sigma)$ and satisfying $P \circ T = P$, (cf. [46]). In the case of spin chains we are in the situation that the measurable space (A^∞, Σ) and the transformation, the (right) shift T, are fixed and we are interested in calculating entropy as a function of shift-invariant probability measures P.

Consider a measurable space $(\mathbf{A}, \Sigma)$. Let $\mathcal{P}$ be a partition of $\mathbf{A}$, i.e. a finite collection of disjoint subsets $P_i \in \Sigma$, $i \in \{1, \ldots, m\}$, such that the union $\cup_{i=1}^m P_i$ is equal to $\mathbf{A}$. The entropy of $\mathcal{P}$ with respect to the probability measure P on $(\mathbf{A}, \Sigma)$ is defined by

$$H(\mathcal{P}) := -\sum_{i=1}^{m} P(P_i) \log P(P_i).$$

The join $\mathcal{P} \vee \mathcal{Q}$ of two partitions $\mathcal{P}, \mathcal{Q}$ of $\mathbf{A}$ is a further partition consisting of all intersections $P_i \wedge Q_j$, $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$. By $T\mathcal{P}$ we denote the partition obtained from $\mathcal{P}$ under the action of the transformation T, i.e. $T\mathcal{P} := \{T(P_1), \ldots, T(P_m)\}$. For a pair (P, T) defined on $(\mathbf{A}, \Sigma)$ and satisfying $P \circ T = P$ the limit

$$h(P, T, \mathcal{P}) := \lim_{n \to \infty} -\frac{1}{n} H(\vee_{i=1}^{n} T^{-i}\mathcal{P})$$

exists (mainly due to the subadditivity of the entropy H) and is called the KS-entropy of (P, T) with respect to $\mathcal{P}$. The KS-entropy of (P, T) is defined as

$$h(P, T) := \sup h(P, T, \mathcal{P}), \tag{2.8}$$

where the supremum is taken over all partitions $\mathcal{P}$ of $\mathbf{A}$.
By the Kolmogorov-Sinai theorem (cf. Theorem 4.17 in [46]) a Σ-generating partitions $\mathcal{P}$ achieves the supremum in (2.8). Here, Σ-generating means that the σ-algebra generated by $\vee_{i \in \mathbb{Z}} T^i \mathcal{P}$ is equal to the σ-algebra Σ of $\mathbf{A}$.

Let us return to spin chains now. For $a \in A$ consider the cylinder set $[a] = \{\mathbf{b} \in A^\infty | \ b_1 = a\} \subseteq A^\infty$. The collection $\mathcal{P} := \{[a]\}_{a \in A}$ is a partition of A^∞, which is Σ-generating as any cylinder set in A^∞ is contained in a finite partition $\vee_{i \in [m,n]} T^i \mathcal{P}$ for an appropriate discrete interval $[m, n]$, $m, n \in \mathbb{Z}$ with $m \leq n$. Consequently, by the Kolmogorov-Sinai theorem for each T-invariant probability measure P we have

$$\begin{aligned} h(p) &= \lim_{n \to \infty} -\frac{1}{n} \sum_{a_1^n \in A^{(n)}} P([a_1^n]) \log P([a_1^n]) \\ &= \lim_{n \to \infty} \frac{1}{n} H(P^{(n)}), \end{aligned} \tag{2.9}$$

where the second equality holds by (2.1). But the limit (2.9) is just the Shannon entropy rate of stationary stochastic processes or equivalently of translation invariant classical spin chains.

2.2 Convergence Theorems

2.2.1 The Shannon-McMillan-Breiman Theorem

Consider an ergodic probability measure P on the space (A^∞, Σ) over the lattice $\mathbb{Z}$ with the Shannon entropy rate $h(P)$. Then $-\frac{1}{n}\log P([\mathbf{a}]_1^n)$ converges to $h(P)$, as $n \to \infty$, in various senses. For us relevant convergence statements are convergence in probability and almost everywhere convergence. The first one means that for arbitrary $\varepsilon \in (0,1)$

$$\lim_{n\to\infty} P^{(n)}\left(\{a_1^n \in A^{(n)} \ : \ |-\frac{1}{n}\log P([a_1^n]) - h(P)| \le \varepsilon\}\right) = 1.$$

This is the assertion of the famous Shannon-McMillan theorem (SM-theorem), cf. [44] and [36]. The stronger almost everywhere convergence means that

$$P\left(\{\mathbf{a} \in A^\infty \ : \ \lim_{n\to\infty} -\frac{1}{n}\log P([\mathbf{a}]_1^n) = h(P)\}\right) = 1. \tag{2.10}$$

This extension of the SM-theorem is due to Breiman and is usually referred to as the Shannon-McMillan-Breiman theorem (SMB-theorem). For the proof we refer to the original paper [15] or alternatively to [9] (Theorem 13.1) or [31]. Consider an independent identically distributed (i.i.d.) probability measure P, i.e. P corresponds one-to-one to the family of marginal distributions given by $P^{(n)} = \times_{i=1}^n P^{(1)}$, where $P^{(1)}$ is a probability distribution on A. In this case

$$-\frac{1}{n}\log P^{(n)}([\mathbf{a}]_1^n) = -\frac{1}{n}\sum_{i=1}^n \log P^{(1)}(a_i)$$

and $h(P) = H(P^{(1)}) = -\sum_{i=1}^n P^{(1)}(a_i)\log P^{(1)}(a_i)$. Consequently (2.10) is not more than the strong (respectively weak in the case of convergence in probability) law of large numbers applied to the random variables $X_i = -\log P^{(1)}(A_i)$, $i \in \mathbb{N}$, (recall that functions of i.i.d. random variables are also i.i.d. random variables). Actually, in the general ergodic case the basic ingredient in the proof of (2.10) is the individual ergodic theorem (e.g. [43]), which can be seen as an extension of the law of large numbers to the class of stationary probability measures.

Now we formulate the ν-dimensional version of the SMB-theorem. It is proven in the more general setting of the action of an amenable group in the work [35] (see also [40] for the case of a class of amenable groups). The generalized SM-theorem (again for amenable group action) was proven by Kieffer in [34].

Theorem 2.2.1 (Shannon-McMillan-Breiman Theorem) *Let P be an $\mathbb{Z}^\nu$-ergodic probability measure on (A^∞, Σ) with the mean Shannon entropy $h(P)$. Then*

$$\lim_{\mathbf{n}\to\infty} -\frac{1}{|\mathbf{n}|}\log P([\mathbf{a}]_{\Lambda(\mathbf{n})}) = h(P), \qquad P-\textit{almost everywhere.} \tag{2.11}$$

The SM-theorem is the convergence in probability version of Theorem 2.2.1. Next, we give an equivalent formulation for the SM-theorem, where the notion of trajectories does not appear. It will be the starting point for formulating the extension of the SM-theorem to the quantum case. Its 1-dimensional version is the standard formulation of the SM-theorem in the context of classical information theory, where the theorem is known as AEP (asymptotic equipartition property) for ergodic information sources.

Proposition 2.2.2 (Shannon-McMillan Theorem) *Let P be an ergodic probability measure on (A^∞, Σ) with the mean Shannon entropy h. Then for all $\delta > 0$ there exists an $\mathbf{n}_\delta \in \mathbb{N}^\nu$ such that for all $\mathbf{n} \in \mathbb{N}^\nu$ with $\Lambda(\mathbf{n}) \supseteq \Lambda(\mathbf{n}_\delta)$ the entropy-typical subset $T_\delta^{(\mathbf{n})} \subseteq A^{(\mathbf{n})}$:*

$$T_\delta^{(\mathbf{n})} := \{\mathbf{a} \in A^{(\mathbf{n})} \mid P^{(\mathbf{n})}(\mathbf{a}) \in (e^{-|\mathbf{n}|(h+\delta)}, e^{-|\mathbf{n}|(h-\delta)})\}$$

satisfies

$$P^{(\mathbf{n})}(T_\delta^{(\mathbf{n})}) \geq 1 - \delta,$$

and

$$\# T_\delta^{(\mathbf{n})} \in (e^{|\mathbf{n}|(h-\delta)}, e^{|\mathbf{n}|(h+\delta)}).$$

We say that a subset $\Omega \subseteq A^{(\mathbf{n})}$ is ε-typical or relevant with respect to the probability measure P if $P^{(\mathbf{n})}(\Omega) \geq 1 - \varepsilon$, i.e. if the probability of Ω is close to 1 (ε is small, typically). Then the SM-theorem states that asymptotically the entropy-typical subsets become relevant with respect to the ergodic probability measure P on (A^∞, Σ). Furthermore, the probability measure is equidistributed to the first order in the exponent on the entropy-typical subsets. The name AEP is related to this point of view.

2.2.2 An Equivalent Finite Form of the SMB-Theorem

Next, we give an equivalent reformulation of the classical SMB-theorem. It avoids the notion of trajectory and allows an immediate translation into the C*-algebraic formalism, which is used to describe quantum spin lattice systems in the next chapter. We restrict ourselves to the 1-dimensional case, i.e. to the group $\mathbb{Z}$, as we will prove the quantum extension only of the 1-dimensional classical SMB-theorem.

According to Breiman's extension of the SM-theorem the entropy-typical subsets $T_\delta^{(n)} \subseteq A^{(n)}$ are not isolated objects, one for each n, but comprise bundles of trajectories and hence can be chosen being nested in the sense that the predecessor subset consists exactly of the same sequences, only shortened by one letter. (Notice that the product sets $A^{(\mathbf{n})}$, $\mathbf{n} \in \mathbb{N}^\nu$, are nested in this sense.) This simple observation leads to a finite, equivalent formulation of Breiman's theorem avoiding the notion of a trajectory. Consider an arbitrary (not necessarily translation invariant) probability measure P on the space (A^∞, Σ). Further, let h be a non-negative real number.

Definition 2.2.3 *We say that a probability measure P on (A^∞, Σ) satisfies the condition (B) with respect to h if for P-almost every sequence $\mathbf{a} \in A^\infty$ the limit of $-\frac{1}{n} \log P^{(n)}([\mathbf{a}]_1^n)$ exists and equals h, as $n \to \infty$.*

In this notation the SMB-theorem asserts that ergodic P satisfy (B) with respect to the Shannon entropy rate $h(P)$ of the process.

In the sequel we denote by $\mathbf{x}_\flat$ the sequence x_1^{n-1} obtained from the finite sequence $\mathbf{x} = x_1^n \in \bigcup_{k\in\mathbb{N}} A^{(k)}$ by dropping the last symbol.

Definition 2.2.4 *We say that a probability measure P on (A^∞, Σ) satisfies condition (B*) with respect to h if for each $\varepsilon > 0$ there exists a sequence $\{C_\varepsilon^{(n)}\}_{n\in\mathbb{N}}$ of subsets of $A^{(n)}$, respectively, and a number $N(\varepsilon)$ such that*

1. $C_\varepsilon^{(n)} = (C_\varepsilon^{(n+1)})_\flat$ *for* $n \geq 1$
   (subsets are nested)
2. $\#C_\varepsilon^{(n)} \in (e^{n(h-\varepsilon)}, e^{n(h+\varepsilon)})$ *for* $n \geq N(\varepsilon)$
   (exponential growth rate)
3. $P^{(n)}(\mathbf{x}) < e^{-n(h-\varepsilon)}$ *for* $n \geq N(\varepsilon)$ *and any* $\mathbf{x} \in C_\varepsilon^{(n)}$
   (upper semi-AEP)
4. $P^{(n)}(C_\varepsilon^{(n)}) > 1 - \varepsilon$
   (typical subsets).

Except for the concatenation condition (1), these are known properties of the entropy-typical sets in the case of an ergodic P. Condition (3) is a weakened version of the well known AEP, particularly it ensures that h is an asymptotic entropy rate in the general (possibly non-stationary) situation.
We obtain the equivalence assertion

Lemma 2.2.5 *A probability measure P on (A^∞, Σ) satisfies (B) if, and only if, it satisfies (B*).*

Combining the SMB-theorem 2.2.1 with the above lemma we derive that any ergodic probability measure P satisfies requirements (1)-(4) in Definition 2.2.4. Thus, Definition 2.2.4 supplies a reformulation of the SMB-theorem, which can be regarded as a finite form of the SMB-theorem.
Observe that (1) describes a tree graph of one-sided infinite trajectories.

Proof of Lemma 2.2.5:
1. Assume that (B) is fulfilled. Let

$$A_{(M,\varepsilon)} := \{\mathbf{x} \in A^\infty |\ P^{(n)}(x_1^n) \in (e^{-n(h+\varepsilon/2)}, e^{-n(h-\varepsilon/2)}),\ \forall n \geq M\}.$$

Obviously (B) implies $P(A_{(M,\varepsilon)}) \longrightarrow 1$, as $M \to \infty$, for fixed ε. So there is some $M(\varepsilon)$ with $P(A_{(M(\varepsilon),\varepsilon)}) > 1 - \varepsilon$. Define

$$C_\varepsilon^{(n)} := \{\mathbf{y} \in A^{(n)} |\ \exists \mathbf{x} \in A_{(M(\varepsilon),\varepsilon)} :\ y_i = x_i,\ 1 \leq i \leq n\},$$

let

$$k(\varepsilon) = \min\{k \in \mathbb{N} |\ (1-\varepsilon)e^{k(h-\varepsilon/2)} > e^{k(h-\varepsilon)}\} \tag{2.12}$$

and set $N(\varepsilon) = \max\{k(\varepsilon), M(\varepsilon)\}$. Then (1)-(4) are easily derived, taking into account that bounds on probabilities imply bounds on cardinality. So (B*) is a consequence of (B).

2. Assume (B*) to be satisfied. Then it is easy to see that we find a nested sequence $\{\overline{C}_\varepsilon^{(n)}\}$ and some $\overline{N}(\varepsilon)$ which fulfill (1)-(4) and even the full AEP condition

$$P^{(n)}(\mathbf{x}) \in (e^{-n(h+\varepsilon)}, e^{-n(h-\varepsilon)}), \quad \forall n \geq \overline{N}(\varepsilon),\ \forall \mathbf{x} \in \overline{C}_\varepsilon^{(n)}.$$

In fact, define

$$A_\varepsilon(n) := \{\mathbf{x} \in A^\infty |\ x_1^n \in C_\varepsilon^{(n)}\}$$

and observe that $A_\varepsilon(n) \searrow A_\varepsilon$, as $n \to \infty$, where A_ε is the tree of trajectories associated with $\{C_\varepsilon^{(n)}\}$. Condition (4) implies $P(A_\varepsilon) \geq 1 - \varepsilon$. Now let

$$\widetilde{A}_\varepsilon(n) := \{\mathbf{x} \in A_\varepsilon(n) |\ P^{(n)}(x_1^n) \leq e^{-n(h+2\varepsilon)}\}.$$

By (2) it follows that $P(\widetilde{A}_\varepsilon(n)) \leq \# C_\varepsilon^{(n)} \cdot e^{-n(h+2\varepsilon)} < e^{-n\varepsilon}$. Now, the Borel-Cantelli principle (cf. Lemma I.1.14 in [43]) implies that $P(\bigcup_{n\geq m} \widetilde{A}_\varepsilon(n)) \searrow 0$, as $m \to \infty$. So there is some $m(\varepsilon)$ with $P(\bigcup_{n \geq m(\varepsilon)} \widetilde{A}_\varepsilon(n)) < \varepsilon$. Let

$$\overline{N}(\varepsilon) := \max\{N(\varepsilon/2), m(\varepsilon/2), k(\varepsilon)\},$$

where $k(\varepsilon)$ is defined by (2.12), and

$$\begin{aligned} \overline{C}_\varepsilon^{(n)} := \{x_1^n \in C_{\varepsilon/2}^{(n)} \quad | \quad & \exists \mathbf{w} \in A_{\varepsilon/2} :\ w_1^n = x_1^n, \\ & P^{(k)}(w_1^k) > e^{-k(h+\varepsilon)},\ \forall k \geq N(\varepsilon)\}. \end{aligned}$$

We have, denoting by $\overline{A}_\varepsilon$ the trajectory tree associated with $\{\overline{C}_\varepsilon^{(n)}\}_n$,

$$\begin{aligned} P^{(n)}(\overline{C}_\varepsilon^{(n)}) \geq P(\overline{A}_\varepsilon) &= P(A_{\varepsilon/2} \setminus \bigcup_{n \geq \overline{N}(\varepsilon)} \widetilde{A}_{\varepsilon/2}(n)) \\ &> 1 - \varepsilon/2 - \varepsilon/2 = 1 - \varepsilon. \end{aligned}$$

Thus (4) is established. Item (1) and the ε-AEP are clearly fulfilled, and (2) follows from the AEP as in step 1. of the proof. Now (B) follows immediately. □

2.2.3 Convergence Assertion for a Non-Consistent Family of Probability Spaces

Concluding this section we formulate a convergence assertion for a family of probability distributions $\{P^{(\mathbf{n})}\}_{\mathbf{n} \in \mathbb{N}^\nu}$ defined on finite sets $A^{(\mathbf{n})}$, respectively. The interesting point is that we do not expect either the finite sets to be nested in the sense discussed in the last section nor the probability distributions to satisfy any consistency conditions. The $A^{(\mathbf{n})}$ are arbitrary finite sets, not required to be of the cartesian product form $\mathsf{X}_{\mathbf{i} \in \Lambda(\mathbf{n})} A$, for some finite set A. This implies that the family cannot be identified with a real stochastic process, in general. Indeed, we make a minimum of assumptions on the family of finite probability spaces $(A^{(\mathbf{n})}, P^{(\mathbf{n})})$ in order to make the following convergence statement useful in the setting of quantum spin lattice systems in the next chapter. There, such probability spaces will arise from eigen-decompositions of density operators, which are elements of a consistent family of density operators corresponding one-to-one to an ergodic quantum state of a quantum spin lattice system. The inconsistency will result from quantum correlations (entanglement) between single quantum spins in the lattice.

For $\mathbf{n} = (n_1, \ldots, n_\nu) \in \mathbb{N}^\nu$ we introduce the notation

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

Further we use the notations $|\mathbf{n}|$ and $\mathbf{n} \to \infty$ as introduced in Section 2.1.

Lemma 2.2.6 *Let $M > 0$ 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 $\frac{1}{|\mathbf{n}|}\log \# A^{(\mathbf{n})} \leq M$ for all $\mathbf{n} \in \mathbb{N}^\nu$ and $P^{(\mathbf{n})}$ is a probability distribution on $A^{(\mathbf{n})}$. Define*

$$\alpha_{\varepsilon,\mathbf{n}}(P^{(\mathbf{n})}) := \min\{\log \#\Omega |\ \Omega \subset A^{(\mathbf{n})}, P^{(\mathbf{n})}(\Omega) \geq 1 - \varepsilon\}. \tag{2.13}$$

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 $\varepsilon \in (0,1)$

$$\lim_{\mathbf{n}\to\infty} \frac{1}{|\mathbf{n}|}\alpha_{\varepsilon,\mathbf{n}}(P^{(\mathbf{n})}) = h. \tag{2.14}$$

We call subsets $\Omega \subseteq A^{(\mathbf{n})}$ achieving the minimum in (2.13) *high probability subsets* at level ε. They are the ε-typical subsets of minimal size. Notice that due to the SM-theorem any ergodic spin lattice system satisfies the conditions in the above Lemma with h being the mean Shannon entropy. It follows that h is a lower bound on the exponential growth rate of the size of relevant/typical subsets. In other words any sequence $\{\Omega^{(\mathbf{n})}\}_{\mathbf{n}\in\mathbb{N}^\nu}$ of subsets of $A^{(\mathbf{n})}$, respectively, with an asymptotically essential (non vanishing) probability with respect to an ergodic P must fulfill $\liminf_{\mathbf{n}\to\infty} \frac{1}{|\mathbf{n}|} \log \Omega^{(\mathbf{n})} \geq h$.

Proof of Lemma 2.2.6: Let $\delta > 0$ and define

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

We fix $\delta > 0$ and use the abbreviation $A_i^{(\mathbf{n})} = A_i^{(\mathbf{n})}(\delta)$, $i \in \{1,2,3\}$. To see that $\lim\limits_{\mathbf{n}\to\infty} P^{(\mathbf{n})}\left(A_3^{(\mathbf{n})}\right) = 0$ assume the contrary and observe that the upper bound on the probability of elements from $A_3^{(\mathbf{n})}$ implies a lower bound on the cardinality of elements from $A_3^{(\mathbf{n})}$ needed to cover an $\varepsilon-$fraction, $\varepsilon \in (0,1)$, of $A^{(\mathbf{n})}$ with respect to $P^{(\mathbf{n})}$. Namely one has

$$\min\left\{\#C \mid C \subset A_3^{(\mathbf{n})}, P^{(\mathbf{n})}(C) > \varepsilon\right\} > \varepsilon \cdot e^{|\mathbf{n}|(h+\delta)}$$

which contradicts condition 2 in the lemma. Furthermore the set $A_3^{(\mathbf{n})}$ cannot asymptotically contribute to the mean entropy h since

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

and

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

Here we used the fact that $\frac{\log \#A^{(\mathbf{n})}}{|\mathbf{n}|}$ stays bounded from above (by M) 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 $p_i, q_i \geq 0$, (cf. Lemma I.6.1 in [43]). Since $A_3^{(\mathbf{n})}$ does not contribute to the entropy one easily concludes that $\lim_{\mathbf{n}\to\infty} P^{(\mathbf{n})}\left(A_1^{(\mathbf{n})}\right) = 0$ because otherwise $\liminf_{\mathbf{n}\to\infty} \frac{1}{|\mathbf{n}|} H\left(P^{(\mathbf{n})}\right) < h$ would hold. Recall that $\delta > 0$ was chosen arbitrarily. Thus

$$\lim_{\mathbf{n}\to\infty} P^{(\mathbf{n})}\left(A_2^{(\mathbf{n})}(\delta)\right) = 1, \qquad \forall \delta > 0. \tag{2.15}$$

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 $\Omega \cap A_2^{(\mathbf{n})}(\delta)$. But δ can be chosen arbitrarily small. □

2.3 Applications to Information Theory

The presented classical results have interesting consequences if applied to information sources in classical information theory, in particular in the context of data compression.

Consider an ergodic information source $(A^{(n)}, \Sigma, P, T)$. Due to Lemma 2.2.6 for any sequence of high probability subsets at level ε, $\varepsilon \in (0,1)$, the limit (2.13) exists and is equal to the mean Shannon entropy h independent of ε. Now, let $\{\Omega_{\varepsilon_n}^{(n)}\}_{n\in\mathbb{N}}$ be a sequence of high probability subsets of $A^{(n)}$ at level ε_n, respectively. We can chose the sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ such that $\varepsilon_n \to 0$, as $n \to \infty$, and simultaneously

$$\lim_{n\to\infty} \frac{1}{n} \log_2 \#\Omega_{\varepsilon_n}^{(n)} = h, \tag{2.16}$$

if h is the base 2 Shannon entropy rate. In order to compress the ergodic IS we can encode finite sequences from $A^{(n)}$ by enumerating only the elements of $\Omega_{\varepsilon_n}^{(n)}$, and ignoring the remaining sequences from $A^{(n)}\backslash\Omega_{\varepsilon_n}^{(n)}$. It follows due to the asymptotically bounded size of $\Omega_{\varepsilon_n}^{(n)}$, eq. (2.16), that asymptotically we do not need more than h representing bits per source signal. Note that as $P^{(n)}(\Omega_{\varepsilon_n}^{(n)}) \to 1$, for $n \to \infty$, the sets of the ignored source sequences from $A^{(n)}\backslash\Omega_{\varepsilon_n}^{(n)}$ have an asymptotically vanishing probability. Consequently an asymptotically perfect reconstruction of the block encoded information source is possible.

Next, we demonstrate how the property of typical subsets to be nested (cf. finite-form of the SMB-theorem in Section 2.2.2) can be used to design a straight-forward sequential data compression scheme. We adopt notations used in Section 2.2.2 and define inductively a sequence of partitions $P_\varepsilon^{(n)}$ of the unit interval, where $\#P_\varepsilon^{(n)} = \#C_\varepsilon^{(n)}$ and each subinterval of $P_\varepsilon^{(n)}$ corresponds to a unique element in $C_\varepsilon^{(n)}$. For $n = 1$ we dissect the unit interval into subintervals of equal length, one for each member of $C_\varepsilon^{(1)}$. At the nth step, we dissect each interval in $P_\varepsilon^{(n-1)}$ into as many subintervals of equal length as the corresponding word in $C_\varepsilon^{(n-1)}$ has successors in $C_\varepsilon^{(n)}$. By this construction we derive from the given nested sequence of subsets a nested sequence of partitions of the unit interval. Each trajectory $\mathbf{x}$ of the bundle of trajectories given by $\{C_\varepsilon^{(n)}\}_{n\in\mathbb{N}}$ corresponds to a unique sequence of nested intervals. The upper endpoint of each such sequence converges to a real number $r(\mathbf{x})$. Now the sequential coding scheme can be described as follows: We take the binary representation $(r_1, r_2, ...)$ of $r(\mathbf{x})$ and encode the nth of the nested intervals by the shortest finite binary string $(r_1, r_2, ..., r_R(n))$ such that the rational number with this binary representation belongs to the given interval.

We emphasize that the discussed compression schemes are designed for ergodic information sources with known statistics. The famous Lempel-Ziv algorithm belongs to the class of universal compression schemes, operating on forth-coming data strings (sequential encoding). It represents arbitrary, not even stationary information sources with an asymptotically vanishing probability of error. Moreover, to represent an individual data string $\mathbf{x}$ the Lempel-Ziv algorithm uses asymptotically $h(p_\mathbf{x})$ bits per signal, whereby $p_\mathbf{x}$ denotes the limiting empirical probability measure on A^∞. It is determined by the limiting relative frequencies of finite sequences in the infinite sequence $\mathbf{x}$ (see e.g. [43] for exact definitions). For stationary information sources the empirical measures $p_\mathbf{x}$ exist almost surely. Furthermore, $p_\mathbf{x}$ are ergodic almost surely. These are essentially implications of the individual ergodic theorem. Recall that the Shannon entropy rate $h(p_\mathbf{x})$ exists for any ergodic $p_\mathbf{x}$. The Lempel-Ziv algorithm uses the method of string matching in order to generate code books for the individual data strings to be compressed. In quantum case string matching is problematic as any measuring process destroys the quantum states.

Chapter 3

Ergodic Quantum Spin Lattice Systems

This chapter contains our main results for ergodic quantum spin lattice systems. These are extensions of classical theorems presented in the last chapter. They show how the mean von Neumann entropy generalizes the mean Shannon entropy concept in the context of spin lattice systems. Indeed, quantum versions of the classical SM-theorem and Breiman's extension can be formulated with the von Neumann entropy replacing the Shannon entropy. In general, the mean von Neumann entropy determines the asymptotically minimal dimension of Hilbert subspaces (associated to finite boxes in $\mathbb{Z}^\nu$) which are relevant with respect to an ergodic quantum state. This is in complete analogy to the classical situation discussed in the last chapter, where the size of relevant subsets is given in terms of mean Shannon entropy.
Moreover, we will see in the present chapter that in the context of quantum information theory the mean von Neumann entropy can be interpreted as the average information content of (ergodic) quantum information sources. These are modeled by 1-dimensional quantum spin lattice systems, in analogy to classical information theory.
We start with a presentation of the standard mathematical formalism for the physical model of ν-dimensional, $\nu \in \mathbb{N}$, quantum spin lattice systems. A detailed introduction to the C*-algebraic formalism can be found e.g. in [14] or [42].

3.1 Mathematical Model

The ν-dimensional infinitely extended lattice corresponds to the group $\mathbb{Z}^\nu$. To each $\mathbf{x} \in \mathbb{Z}^\nu$ there is associated an algebra $\mathcal{A}_\mathbf{x}$ of observables for a spin located at site $\mathbf{x}$. It is given by

$$\mathcal{A}_\mathbf{x} := \tau(\mathbf{x})\mathcal{A},$$

where $\tau(\mathbf{x})$ is an isomorphism and $\mathcal{A}$ is a finite dimensional unital C^*-algebra. The local algebra $\mathcal{A}_\Lambda$ of observables for the finite subset $\Lambda \subset \mathbb{Z}^\nu$ is given by

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

The infinite lattice system is constructed from the finite subsets $\Lambda \subset \mathbb{Z}^\nu$. The algebra of observables corresponding to the whole lattice $\mathbb{Z}^\nu$ is the quasi-local C*-algebra $\mathcal{A}^\infty$. It is defined as the operator norm closure

$$\mathcal{A}^\infty := \overline{\bigcup_{\Lambda \subset \mathbb{Z}^\nu} \mathcal{A}_\Lambda}^{\|\cdot\|}.$$

Hereby $\bigcup_{\Lambda\subset\mathbb{Z}^\nu} \mathcal{A}_\Lambda$ is a normed *-algebra (sometimes called a local algebra), where the union is in the sense of inclusion maps: $\mathcal{A}_\Lambda \to \mathcal{A}_{\Lambda'},\ a \mapsto a \otimes \mathbf{1}_{\mathcal{A}_{\Lambda'\setminus\Lambda}}$, for all $\Lambda \subseteq \Lambda'$. A state of the infinite spin system is given by a normed positive functional Ψ on $\mathcal{A}^\infty$. It corresponds one-to-one to a consistent family of states $\{\Psi^{(\Lambda)}\}_{\Lambda\subset\mathbb{Z}^\nu}$, where each $\Psi^{(\Lambda)}$ is the restriction of Ψ to the finite dimensional subalgebra $\mathcal{A}_\Lambda$ of $\mathcal{A}^\infty$ and consistency means that

$$\Psi^{(\Lambda)} = \Psi^{(\Lambda')} \upharpoonright \mathcal{A}_\Lambda,$$

for $\Lambda \subset \Lambda'$. The one-to-one correspondence reflects the fact, that the state of the entire spin lattice system is assumed to be determined by the expectation values of all observables on finite subsystems Λ. Actually, it is sufficient to consider boxes only. For each $\Psi^{(\Lambda)}$ there exists a unique density operator $D_\Lambda \in \mathcal{A}_\Lambda$, such that

$$\Psi^{(\Lambda)}(a) = \mathrm{tr}_\Lambda D_\Lambda a, \qquad a \in \mathcal{A}_\Lambda$$

and tr_Λ is the trace on $\mathcal{A}_\Lambda$. By $\mathcal{S}(\mathcal{A}^\infty)$ we denote the state space of $\mathcal{A}^\infty$.
Every $\mathbf{x} \in \mathbb{Z}^\nu$ defines a translation of the lattice and induces an automorphism $T(\mathbf{x})$ on $\mathcal{A}^\infty$, which is a canonical extension of the isomorphisms for finite $\Lambda \subset \mathbb{Z}^\nu$:

$$\begin{aligned} T(\mathbf{x}) : \mathcal{A}_\Lambda &\longrightarrow \mathcal{A}_{\Lambda+\mathbf{x}} \\ a &\longmapsto \left(\bigotimes_{\mathbf{z}\in\Lambda} T_{\mathbf{z}}(\mathbf{x})\right) a, \end{aligned}$$

where $T_{\mathbf{z}}(\mathbf{x}) := \tau(\mathbf{x})\tau^{-1}(\mathbf{z})$. Then $\{T(\mathbf{x})\}_{\mathbf{x}\in\mathbb{Z}^\nu}$ is an action of the translation group $\mathbb{Z}^\nu$ by automorphisms (shifts) on $\mathcal{A}^\infty$. The triple

$$(\mathcal{A}^\infty, \Psi, T(\mathbb{Z}^\nu)) \tag{3.1}$$

is a C*-dynamical system describing a quantum spin lattice system, where $T(\mathbb{Z}^\nu)$ denotes the representation of the translation group $\mathbb{Z}^\nu$ by shifts $T(\mathbf{x})$ on $\mathcal{A}^\infty$. In the 1-dimensional case we have $\{T(x)\}_{x\in\mathbb{Z}} = \{T^x\}_{x\in\mathbb{Z}}$ with $T := T_{x=1}$ and we use the shorter notation $(\mathcal{A}^\infty, \Psi, T)$ instead of (3.1).

Let G be any subgroup of $\mathbb{Z}^\nu$ and denote by $\mathcal{T}(\mathcal{A}^\infty, G)$ the set of states, which are invariant under the translations associated with G, i.e.

$$\mathcal{T}(\mathcal{A}^\infty, G) := \{\Psi \in \mathcal{S}(\mathcal{A}^\infty) | \Psi \circ T(\mathbf{x}) = \Psi,\ \forall \mathbf{x} \in G\}.$$

We will be concerned mostly with the space of $\mathbb{Z}^\nu$-invariant states. We introduce the abbreviation $\mathcal{T}(\mathcal{A}^\infty) = \mathcal{T}(\mathcal{A}^\infty, \mathbb{Z}^\nu)$. Clearly, $\mathcal{T}(\mathcal{A}^\infty) \subset \mathcal{T}(\mathcal{A}^\infty, G)$ for any proper subgroup G of $\mathbb{Z}^\nu$.
For $\mathbf{n} \in \mathbb{N}^\nu$ we denote by $\Lambda(\mathbf{n})$ a box in $\mathbb{Z}^\nu$ as was defined by (2.3) in last chapter.

Further recall that for $n \in \mathbb{N}$ the hypercube $\Lambda(n)$ was given by $\Lambda(n) := \{\mathbf{x} \in \mathbb{Z}^\nu | \ \mathbf{x} \in \{0, \ldots, n-1\}^\nu\}$. In the following we simplify notations by defining

$$\mathcal{A}^{(\mathbf{n})} := \mathcal{A}_{\Lambda(\mathbf{n})}, \quad \Psi^{(\mathbf{n})} := \Psi^{(\Lambda(\mathbf{n}))}, \quad \text{and} \quad D_{\mathbf{n}} := D_{\Lambda(\mathbf{n})},$$

for $\mathbf{n} \in \mathbb{N}^\nu$ (respectively for $n \in \mathbb{N}$). Obviously, because of the shift-invariance any $\Psi \in \mathcal{T}(\mathcal{A}^\infty)$ is uniquely defined by the family of states $\{\Psi^{(\mathbf{n})}\}_{\mathbf{n}\in\mathbb{N}^\nu}$ or alternatively by the family of corresponding density operators $\{D_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}^\nu}$.

Commutative C*-dynamical Systems

For convenience, we briefly discuss the correspondence between a commutative C*-dynamical system $(\mathcal{A}^\infty, \Psi, T)$ and the quadruple (A^∞, Σ, P, T) modeling a classical spin lattice system, as presented in the section 2.1. Here, for simplicity of notations we restrict ourselves to 1-dimensional lattices.

We start with the triple $(\mathcal{A}^\infty, \Psi, T)$, where $\mathcal{A}^\infty$ is constructed from an abelian finite dimensional and unital algebra $\mathcal{A}$. Thus $\mathcal{A}^\infty$ is also abelian and unital. Consequently both $\mathcal{A}$ and $\mathcal{A}^\infty$ can be represented by the Gelfand isomorphism (e.g. Theorem VIII.2.1 in [17] or Theorem 2.1.11A in [14]) as algebras of (continuous) functions over the compact (in the weak* topology, which is inherited from the dual of $\mathcal{A}$ and $\mathcal{A}^\infty$, respectively) maximal ideal spaces of $\mathcal{A}$ and $\mathcal{A}^\infty$, respectively. Due to the finite dimensionality of the algebra $\mathcal{A}$ its maximal ideal space A is a set of finite cardinality with $\#A = \dim \mathcal{A}$. The quasi-local algebra $\mathcal{A}^\infty$ is $*$-isomorphic to $C(A^\infty)$, where $A^\infty := \times_{x\in\mathbb{Z}^\nu} A$. Further, the state Ψ on $\mathcal{A}^\infty$ induces a linear functional on the algebra of functions over A^∞. By the Riesz representation theorem there exists a probability measure P on the measurable space (A^∞, Σ), which is uniquely determined by

$$\Psi(a) = \sum_{\mathbf{a}\in A^\Lambda} f_a(\mathbf{a}) P_\Lambda(\mathbf{a}),$$

for all $a \in \mathcal{A}_\Lambda$ and arbitrary $\Lambda \subset \mathbb{Z}$. Here $f_a(\cdot) \in C(A_\Lambda)$ is the Gelfand transform of $a \in \mathcal{A}_\Lambda$. Again, translations of the lattice $\mathbb{Z}$ induce a dynamics on A^∞ given by the family of shifts $\{T^x\}_{x\in\mathbb{Z}}$ on A^∞. Thus we achieve from $(\mathcal{A}^\infty, \Psi, T)$ the quadruple (A^∞, Σ, P, T), which can be associated to a classical spin lattice system.

Contrary, given a quadruple (A^∞, Σ, P, T) associated to a classical translation invariant spin lattice system with the finite alphabet A, we can identify the elements a of A with a set of mutually orthogonal 1-dimensional projectors p_a on a Hilbert space $\mathcal{H}$. The algebra $\mathcal{A}$ generated by the projectors p_a, $a \in A$ is then finite dimensional and commutative. It can be taken to construct the commutative quasi-local algebra $\mathcal{A}^\infty$ over the lattice $\mathbb{Z}$. For each discrete interval $\Lambda = [m, n]$ in $\mathbb{Z}$ the probability measure P_Λ on A_Λ induces a density operator $D_\Lambda := \sum_{a_m^n \in A} P_\Lambda(a_m^n) \bigotimes_{i=m}^n p_{a_i}$ in $\mathcal{A}_\Lambda$. The family $\{D_\Lambda\}_{\Lambda\subset\mathbb{Z}}$ is consistent by the consistency of the probability measures $\{P_\Lambda\}_{\Lambda\subset\mathbb{Z}}$, and consequently corresponds to a state Ψ on $\mathcal{A}^\infty$. Thus this construction leads back to the C*-algebraic representation $(\mathcal{A}^\infty, \Psi, T)$. It follows the correspondence $(\mathcal{A}^\infty, \Psi, T) \longleftrightarrow (A^\infty, \Sigma, P, T)$ for commutative systems.

Ergodicity of Quantum States with Respect to Group Actions

The set $\mathcal{T}(\mathcal{A}^\infty, G)$ is a convex, weak*-compact subset of $\mathcal{S}(\mathcal{A}^\infty)$ for any subgroup G of $\mathbb{Z}^\nu$. We denote by $\partial_{\mathrm{ex}}\mathcal{T}(\mathcal{A}^\infty, G)$ the set of extremal points of $\mathcal{T}(\mathcal{A}^\infty, G)$ and refer to the elements of $\partial_{\mathrm{ex}}\mathcal{T}(\mathcal{A}^\infty, G)$ as *G-ergodic* states. The elements of $\partial_{\mathrm{ex}}\mathcal{T}(\mathcal{A}^\infty)$ are called *ergodic* states.

Example: The simplest example of an ergodic quantum state on $\mathcal{A}^\infty$ is the analogue of a classical i.i.d. stochastic process, which is uniquely characterized by the distribution of a single random variable. Similarly, an i.i.d. quantum state can be constructed from a single-site density operator $D \in \mathcal{A}$. Namely, it corresponds to the consistent family $\{D_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}^\nu}$, where each $D_{\mathbf{n}}$ is the tensor product density operator

$$D_{\mathbf{n}} := \bigotimes_{\mathbf{x}\in\Lambda(\mathbf{n})} D \in \mathcal{A}^{(\mathbf{n})}.$$

We remark that in general it is not easy to check by definition if a given quantum state Ψ on $\mathcal{A}^\infty$ is ergodic. Fortunately, for C*-dynamical systems of the form $(\mathcal{A}^\infty, \Psi, T(\mathbb{Z}^\nu))$ ergodicity is equivalent to the statement that

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \Psi\left(\left(\frac{1}{\#(\Lambda(\mathbf{n}))}\sum_{\mathbf{x}\in\Lambda(\mathbf{n})} T(\mathbf{x})(a)\right)^2\right) = \Psi(a)^2 \tag{3.2}$$

holds for all self-adjoint $a \in \mathcal{A}^\infty$, (cf. Proposition 6.3.5 in [42], see also [27]). The equivalence can be verified using the $\mathbb{Z}^\nu$-abelianness (a kind of asymptotic abelianness) of quasi-local algebras over $\mathbb{Z}^\nu$, (cf. Section 6.2 in [42]). Compared to the original definition of ergodicity, it should be easier to deal with condition (3.2).

Another useful characterization of ergodicity exists for the class of *C^*-finitely correlated* or alternatively called *algebraic* states , which form a weak*-dense subset in $\mathcal{T}(\mathcal{A}^\infty, \mathbb{Z})$, (cf. [20]). A state $\Psi \in \mathcal{T}(\mathcal{A}^\infty, \mathbb{Z})$ is called algebraic if it can be constructed from a triple $(\mathcal{B}, \mathbb{E}, \Omega)$ in the following way:

$$\Psi(a_1 \otimes a_2 \otimes \cdots \otimes a_m) = \Omega(\mathbb{E}(a_1 \otimes \mathbf{1}_{\mathcal{B}}) \circ \mathbb{E}(a_2 \otimes \mathbf{1}_{\mathcal{B}}) \circ \cdots \circ \mathbb{E}(a_m \otimes \mathbf{1}_{\mathcal{B}})),$$

where $\mathcal{B}$ is a finite dimensional C^*-algebra, $\mathbb{E} : \mathcal{A} \otimes \mathcal{B} \longrightarrow \mathcal{B}$ is a completely positive linear map and Ω is a (faithful) state on $\mathcal{B}$. The state Ψ is ergodic if and only if $\mathbf{1}_{\mathcal{B}}$ is the only eigenvector of $\mathbb{E}(\mathbf{1}_{\mathcal{A}} \otimes (\cdot))$ with respect to the eigenvalue 1, i.e. the characterization of ergodicity is in terms of a completely positive linear map between finite dimensional algebras.

Elements from the union $\cup_{\mathbf{l}\in\mathbb{N}^\nu}\partial_{\mathrm{ex}}\mathcal{T}(\mathcal{A}^\infty, \mathbf{l}\cdot\mathbb{Z}^\nu)$, i.e. states which are ergodic with respect to any shift $T(\mathbf{x})$, $\mathbf{x} \in \mathbb{N}^\nu$, are called *completely ergodic*. For the class of completely ergodic states of quantum spin chains Hiai and Petz obtained some results, which have strongly stimulated our work (we will explain this in more detail later), [23]. On the one hand the assumption of complete ergodicity is much less restrictive than the condition of i.i.d. states, where any kind of correlations is excluded. In the context of quantum information theory predominately i.i.d. quantum states have been studied in detail. On the other hand complete ergodicity is a rather strong restriction compared with the simple condition of ergodicity. In general, complete ergodicity does not allow any periodicity in the sense that the state (convexly) decomposes in states, which are ergodic with respect to some subgroups $\mathbf{l} \cdot \mathbb{Z}^\nu$, $\mathbf{l} \in \mathbb{N}^\nu$. The following chain of implications relates the completely ergodic property to the more familiar mixing properties of stochastic processes: strong mixing $\Longrightarrow$ weak mixing $\Longrightarrow$ completely ergodic $\Longrightarrow$ ergodic. Thus completely ergodicity is the next weakest mixing property after ergodicity. However, for example in the case of classical Markov processes complete ergodicity is equivalent to the weak and strong mixing property. The algebraic quantum states (mentioned in the examples above) can be regarded as quantum Markov states in some sense. For these states complete ergodicity is equivalent to weak and strong mixing property as well, as is shown in [24].

Mean Von Neumann Entropy

The von Neumann entropy of a state Ψ on an arbitrary finite dimensional C^*-algebra $\mathcal{A}$ (cf. [39]) is defined in terms of the corresponding density operator D_Ψ as

$$S(\Psi) := -\mathrm{tr}_{\mathcal{A}} D_\Psi \log D_\Psi,$$

where $\mathrm{tr}_{\mathcal{A}}$ is the trace on $\mathcal{A}$. We will interchangeably use a notation relating to the density operator, that is $S(D_\Psi)$ instead of $S(\Psi)$.
Notice that $S(\Psi)$ is equal to the Shannon entropy of the probability distribution P given by the eigenvalues λ of the corresponding density operator D_Ψ, namely

$$S(\Psi) = -\sum_\lambda \lambda \log \lambda = H(P).$$

It is well known that for every $\Psi \in \mathcal{T}(\mathcal{A}^\infty)$ the limit

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

exists due to the subadditvity of von Neumann entropy. We call $s(\Psi)$ the *mean von Neumann entropy* of Ψ. However, treating quantum spin chains in the context of quantum information theory we prefer to call $s(\Psi)$ the *von Neumann entropy rate* corresponding to the meaning of the $\mathbb{Z}$-translations as time shifts.

Let $l \in \mathbb{N}$ and consider the subgroup $G_l := l \cdot \mathbb{Z}^\nu$. For a G_l-invariant state Ψ we define the mean entropy with respect to G_l by

$$s(\Psi, G_l) := \lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} \frac{1}{\#(\Lambda(\mathbf{n}))} S(\Psi^{(l \cdot \mathbf{n})}). \tag{3.3}$$

Observe that if the state Ψ is $\mathbb{Z}^\nu$-invariant then we have the relation

$$s(\Psi, G_l) = l^\nu \cdot s(\Psi).$$

Further note that in the case of a commutative algebra $\mathcal{A}$ the mean von Neumann entropy s of a state $\Psi \in \mathcal{T}(\mathcal{A}^\infty)$ coincides with the mean Shannon entropy h. Therefore we can say that s is an extension of the mean Shannon entropy concept to non-commutative spin lattice systems, at least in the formal sense.

The quantum analogues of the SM- and SMB-theorems presented below establish the mean von Neumann entropy as an extension of h in a more general sense: It turns out that s, in complete analogy to h, appears as a limit in the quantum version of the SM-theorem 3.3.1, and in its stronger form Theorem 3.5.1, which can be understood as a quantum version of the SMB-theorem. Moreover, it can be interpreted, again in complete analogy to classical situation, from the information theoretical point of view as the average information content of the quantum information source. Operationally this means that the von Neumann entropy rate of an ergodic quantum information source is the necessary amount of quantum resources (usually measured in qubits, the quantum version of the classical bits) needed for asymptotically lossless compression of the source, (cf. Theorem 3.4.1 below).

Remark: There are several efforts to define an extension of the KS-entropy to arbitrary C^*-dynamical systems. Maybe the most famous such quantum dynamical entropy concept is the CNT-entropy introduced in [16]. It arises from a technically rather complicated construction involving variation of finite quantum state decompositions and is difficult to compute for concrete models. However it satisfies (by construction) some required properties like additivity.

Another quantum dynamical entropy concept for general C^*-dynamical systems (nonequivalent to the CNT-entropy) is the AF-entropy introduced in [3]. The construction starts with a finite operational partition of unity as a quantum generalization of finite partition of a measurable space in its interpretation as measurement. Correlation matrices generated by the initial unit partition and its 'refinements' due to the given (reversible) dynamics form a consistent family of density matrices corresponding to an imaginary one-sided quantum spin chain (unfortunately the spin chain is not necessary stationary), which can be regarded as a quantum symbolic dynamics of the initial system. The supremum over operational partitions of unity of the mean von Neumann entropy of the resulting imaginary quantum chains gives the AF-entropy of the given C^*-dynamical system. Compared with the CNT-entropy the AF-entropy has the advantage to be simpler technically.

Until now the CNT-entropy misses convincing interpretations, in particular in the context of information theory. Some attempts to give an information theoretical meaning to the CNT-entropy can be found in [5]. The work of Benatti and Knauf, [6], discusses in a comprehensible way the CNT-entropy, however especially as a possible quantity characterizing quantum chaos. Nevertheless, we mention it here as a useful reference helping to get some intuition on the CNT-entropy.

Concerning the AF-entropy we refer e.g. to the papers [1] and [2], where the meaning of the AF-entropy in quantum information theory is discussed in the context of classical-quantum coding and quantum communication channels.

3.2 An Ergodic Decomposition Theorem

Consider a $\mathbb{Z}^\nu$-ergodic quantum state Ψ on a quasi-local C*-algebra $\mathcal{A}^\infty$. Then, according to the definition of ergodicity with respect to a group, Ψ is invariant but in general not ergodic with respect to an arbitrary subgroup $l \cdot \mathbb{Z}^\nu$, where $l > 1$ integer. This would be the case under the condition of complete ergodicity.

However, it turns out that the (unique) convex decomposition of a $\mathbb{Z}^\nu$-ergodic Ψ in $l \cdot \mathbb{Z}^\nu$-ergodic states has a rather well controllable structure: There are not more than l^ν components in the decomposition, the convex weights are equidistributed and the $l \cdot \mathbb{Z}^\nu$-ergodic components are related to each other by suitable lattice translations. Moreover the mean von Neumann entropies of the components (computed with respect to the subgroup $l \cdot \mathbb{Z}^\nu$) are all identical to $s(\Psi, l \cdot \mathbb{Z}^\nu)$. A precise formulation of this situation is given in Theorem 3.2.1 below. The classical case is covered by this ergodic decomposition theorem if we apply it to a commutative quasi-local algebra $\mathcal{A}^\infty$. For the proof of the classical version in the framework of information theory (stochastic processes) we refer for example to [7] (Theorem 7.2.3).
In the present work we need the ergodic decomposition theorem as an intermediate result in the proof of the quantum SM-theorem. However we emphasize, that it is also an interesting result on its own.

Theorem 3.2.1 *Let Ψ be a $\mathbb{Z}^\nu$-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 $\Psi_{\mathbf{x}}$:*

$$\Psi = \frac{1}{\#(\Lambda(\mathbf{k}(l)))} \sum_{\mathbf{x} \in \Lambda(\mathbf{k}(l))} \Psi_{\mathbf{x}}. \tag{3.4}$$

The G_l-ergodic decomposition (3.4) has the following properties:

1. *$k_j(l) \leq l$ and $k_j(l) | l$ for all $j \in \{1, \dots, \nu\}$*
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 (3.4) 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.*

$$s(\Psi_{\mathbf{x}}, G_l) = s(\Psi, G_l) \tag{3.5}$$

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

The basic tool for proving the above theorem will be the GNS representation of the C*-dynamical system $(\mathcal{A}^\infty, \Psi, T(\mathbb{Z}^\nu))$. The concept of the GNS construction is introduced in detail e.g. in [14].
The proof relies on the fact, that quasi-local C*-algebras are asymptotically abelian with respect to the underlying group $\mathbb{Z}^\nu$ and its subgroups G_l, (cf. [42], Section 6.2). This property implies e.g. a simplex structure of the compact convex set $\mathcal{T}(\mathcal{A}^\infty)$: An extremal convex decomposition of a $\mathbb{Z}^\nu$-invariant state within $\mathcal{T}(\mathcal{A}^\infty)$ is unique. A variety of geometrical properties, which are given in the case of asymptotical abelianness (with respect to a group), are involved in the proof.

Proof of Theorem 3.2.1: Let $(\mathcal{H}_\Psi, \pi_\Psi, \Omega_\Psi, U_\Psi)$ be the GNS representation of the C^*-dynamical system $(\mathcal{A}^\infty, \Psi, T(\mathbb{Z}^\nu))$. U_Ψ is the unitary representation of $\mathbb{Z}^\nu$ on $\mathcal{H}_\Psi$. It satisfies for every $\mathbf{x} \in \mathbb{Z}^\nu$:

$$U_\Psi(\mathbf{x})\Omega_\Psi = \Omega_\Psi, \tag{3.6}$$

$$U_\Psi(\mathbf{x})\pi_\Psi(a)U_\Psi^*(\mathbf{x}) = \pi_\Psi(T(\mathbf{x})a), \qquad \forall a \in \mathcal{A}^\infty. \tag{3.7}$$

Define

$$\begin{aligned}\mathcal{N}_{\Psi,G_l} &:= \pi_\Psi(\mathcal{A}^\infty)\cup U_\Psi(G_l),\\ \mathcal{P}_{\Psi,G_l} &:= \{P\in\mathcal{N}'_{\Psi,G_l}|\ P=P^*=P^2\}.\end{aligned}$$

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 [14] or Lemma IV.3.4 in [27]). This is, essentially, due to the fact that the quasi-local algebra is by construction asymptotically abelian with respect to G_l. The details can be found in the references cited above.

Consider some $l>1$ such that $\Psi\notin\partial_{\mathrm{ex}}\mathcal{T}(\mathcal{A}^\infty,G_l)$. (If there is no such l the statement of the theorem is trivial.) Then

$$\mathcal{P}_{\Psi,G_l}\backslash\{0,\mathbf{1}\}\neq\emptyset. \tag{3.8}$$

In fact, $\Psi\notin\partial_{\mathrm{ex}}\mathcal{T}(\mathcal{A}^\infty,G_l)$ is equivalent to the reducibility of $\mathcal{N}_{\Psi,G_l}$, (cf. Theorem 4.3.17 in [14]). This means that there is a non-trivial closed subspace of $\mathcal{H}_\Psi$ invariant under the action of $\pi_\Psi(\mathcal{A}^\infty)$ and $U_\Psi(G_l)$. Let P be the projector on this subspace and $P^\perp=\mathbf{1}-P$. Then $P,P^\perp\notin\{0,\mathbf{1}\}$ and of course P and $P^\perp$ are contained in $\mathcal{N}'_{\Psi,G_l}$. Thus (3.8) is clear.

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

$$\{Q_i\}_{i\in I}\text{ orthogonal partition of unity in }\mathcal{N}'_{\Psi,G_l}\Longrightarrow|I|\leq l^\nu. \tag{3.9}$$

To see (3.9) 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

$$\begin{aligned}\pi_\Psi(a)U_\Psi(\mathbf{x})QU_\Psi^*(\mathbf{x}) &= U_\Psi(\mathbf{x})\pi_\Psi(T(-\mathbf{x})a)QU_\Psi^*(\mathbf{x}) \quad \text{(by (3.7))}\\ &= U_\Psi(\mathbf{x})Q\pi_\Psi(T(-\mathbf{x})a)U_\Psi^*(\mathbf{x})\\ &= U_\Psi(\mathbf{x})QU_\Psi^*(\mathbf{x})\pi_\Psi(a) \quad \text{(by (3.7))}\end{aligned}$$

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. 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. Define

$$\bar{Q}:=\bigvee_{\mathbf{x}\in\Lambda(l)}U_\Psi(\mathbf{x})QU_\Psi^*(\mathbf{x}),$$

which has the representation as $\bigvee_{\mathbf{x}\in\Lambda(l)}1_{Q_\mathbf{x}}=1_{\bigcup_{\mathbf{x}\in\Lambda(l)}Q_\mathbf{x}}$. Note that if $Q\in\mathcal{P}_{\Psi,G_l}\backslash\{0\}$ then for any $\mathbf{y}\in\mathbb{Z}^\nu$ we have

$$U_\Psi(\mathbf{y})QU_\Psi^*(\mathbf{y})=U_\Psi(\mathbf{y}(\mathrm{mod}\ \mathbf{l}))QU_\Psi^*(\mathbf{y}(\mathrm{mod}\ \mathbf{l})),$$

where $\mathbf{l}=(l,\dots,l)\in\mathbb{Z}^\nu$. 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 $\bar{Q} = \mathbf{1}$. 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:

$$\begin{aligned} 1 = \langle \Omega_\Psi, \bar{Q}\Omega_\Psi\rangle &\leq \sum_{\mathbf{x}\in\Lambda(l)} \langle \Omega_\Psi, U_\Psi(\mathbf{x})QU_\Psi^*(\mathbf{x})\Omega_\Psi\rangle \\ &= l^\nu \cdot \langle \Omega_\Psi, Q\Omega_\Psi\rangle \quad \text{(by (3.6))}. \end{aligned}$$

Thus (3.9) is clear.
Combining the results (3.8) and (3.9) 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 $P_i \in \mathcal{P}_{\Psi,G_l}\backslash\{0,\mathbf{1}\}$. Here we use the standard definition of minimality:

$$\begin{aligned} P \text{ minimal projection in } \mathcal{N}'_{\Psi,G_l} \quad :\Longleftrightarrow \quad & 0 \neq P \in \mathcal{P}_{\Psi,G_l} \text{ and } Q \leq P \\ & \Longrightarrow Q = P, \quad \forall Q \in \mathcal{P}_{\Psi,G_l}\backslash\{0\} \end{aligned}$$

The abelianness 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:

$$Q \in \mathcal{P}_{\Psi,G_l} \implies \exists\{P_{i_j}\}_{j=0}^{s\leq n_l-1} \subset \mathcal{P}_{\Psi,G_l} \text{ such that } Q = \sum_{j=0}^{s} P_{i_j}. \tag{3.10}$$

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

$$p_0 \leq p_i, \qquad \forall i \in \{1,\ldots,n_l-1\}. \tag{3.11}$$

Let

$$G(P_0) := \{\mathbf{x} \in \mathbb{Z}^\nu |\ U(\mathbf{x})P_0U^*(\mathbf{x}) = P_0\}.$$

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

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

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 $\mathcal{P}_{\Psi,G_l}$, $\{U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x})\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \subseteq \mathcal{P}_{\Psi,G_l}$ for $\mathbf{k}(l) = (k_1(l),\ldots,k_\nu(l))$. Thus by (3.10) each $U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x})$, $\mathbf{x} \in \Lambda(\mathbf{k}(l))$, can be represented as a sum of minimal projections. But then by linearity of the expectation values and the assumed ordering (3.11) each $U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x})$ must be a minimal projection for $\mathbf{x} \in \Lambda(\mathbf{k}(l))$. Otherwise there would be a contradiction to $\langle \Omega_\Psi, U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x})\Omega_\Psi\rangle = p_0$. Consequently $\{U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x})\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} \subseteq \{P_i\}_{i=0}^{n_l-1}$. Consider the projection

$$\bar{P}_0 = \sum_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x}).$$

Using the same argument as for $\bar{Q}$, defined 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

$$\{U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x})\}_{\mathbf{x}\in\Lambda(\mathbf{k}(l))} = \{P_j\}_{j=0}^{n_l-1}.$$

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

$$P_i = U_\Psi(\mathbf{x})P_0U_\Psi^*(\mathbf{x}) =: P_{\mathbf{x}}. \tag{3.12}$$

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

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

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

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

From (3.12), (3.6) and (3.7) 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 $\Psi_{\mathbf{x}}$.

Finally, we prove item 3. saying that the G_l-ergodic components of Ψ have all a mean Shannon entropy equal to $s(\Psi, G_l)$. It is a well known result that the mean von Neumann 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 [42]). Thus to prove (3.5) 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

$$|S(\Psi_{\mathbf{x}}^{(l\mathbf{n})}) - S(\Psi_0^{(l\mathbf{n})})| = o(|\mathbf{n}|) \qquad \text{as } \mathbf{n}\to\infty. \tag{3.13}$$

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 as follows

$$S(\Psi_0^{(\widetilde{\Lambda})}) - \log \mathrm{tr}_{\widetilde{\Lambda}\setminus\Lambda(l\mathbf{n})}\mathbf{1} \leq S(\Psi_\gamma^{(l\mathbf{n})}) \leq S(\Psi_0^{(\widehat{\Lambda})}) + \log \mathrm{tr}_{\Lambda(l\mathbf{n})\setminus\widehat{\Lambda}}\mathbf{1},$$

where $\gamma \in \{\mathbf{x}, 0\}$. Thus (3.13) follows immediately. $\square$

3.3 Quantum Shannon-McMillan Theorem

The main result in this work is a generalization of the classical Shannon-McMillan theorem to the quantum case.

Theorem 3.3.1 (Quantum Shannon-McMillan Theorem) *Let Ψ be an ergodic state on $\mathcal{A}^\infty$ with mean von Neumann 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 $0 \neq p \in \mathcal{A}^{(\mathbf{n})}$ 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. $e^{\#(\Lambda(\mathbf{n}))(s(\Psi)-\delta)} < \mathrm{tr}_\mathbf{n}(p_\mathbf{n}(\delta)) < e^{\#(\Lambda(\mathbf{n}))(s(\Psi)+\delta)}$.

As will be seen in the proof of the above theorem the entropy-typical Hilbert subspace corresponding to the projector $p_\mathbf{n}(\delta)$ can be chosen as the linear hull of the eigenvectors of the density operator $D_{\Psi^{(\mathbf{n})}}$ whose eigenvalues are of order $e^{-\#(\Lambda(\mathbf{n}))s(\Psi)}$.

The quantum SM-theorem reduces to the classical SM-theorem (Proposition 2.2.2) in its C*-algebraic formulation if we choose $\mathcal{A}^\infty$ to be constructed from an abelian finite dimensional unital C^*-algebra $\mathcal{A}$.

The subsequent proposition can be considered as a reformulation of the above quantum SM-theorem in terms of high probability subspaces and is especially suited for possible applications in (quantum) data compression. We first fix notations.

Definition 3.3.2 *For $\varepsilon \in (0,1)$ and $\mathbf{n} \in \mathbb{N}^\nu$ we define*

$$\beta_{\varepsilon,\mathbf{n}}(\Psi) := \min\{\log(tr_{\mathbf{n}}\ q) |\ q \in \mathcal{A}^{(\mathbf{n})} \text{ projection},\ \Psi^{(\mathbf{n})}(q) \geq 1-\varepsilon\}. \tag{3.14}$$

A projector $p \in \mathcal{A}^{(\mathbf{n})}$ which achieves the minimum (3.14) is called high probability projector *at level ε and the corresponding Hilbert subspace is referred to as* high probability subspace *at level ε.*

Roughly speaking, the high probability subspaces are the relevant subspaces (with respect to Ψ) of minimal dimension. The proposition states that the mean von Neumann entropy of an ergodic quantum state Ψ gives the first order exponent for the increase of per site dimension of high probability subspaces, independent of the level ε.

Proposition 3.3.3 *Let Ψ be an ergodic state on $\mathcal{A}^\infty$ with mean von Neumann entropy $s(\Psi)$. Then for every $\varepsilon \in (0,1)$*

$$\lim_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#(\Lambda(\mathbf{n}))}\beta_{\varepsilon,\mathbf{n}}(\Psi) = s(\Psi). \tag{3.15}$$

As we will see in the next section, in the setting of quantum information theory high probability subspaces of quantum spin lattice systems modeling quantum information sources are the crucial objects for formulating and proving data coding/compression theorems. As subspaces of probability close to 1 they are the relevant subspaces in the sense that the expectation values of any observables restricted to these subspaces are almost equal to the corresponding ones on the entire space. On the other hand the minimal dimension allows an economical use of resources (qubits) needed for quantum information storage and transmission.

We remark that Proposition 3.15 already appeared in the papers [24] and [23] of Petz and Hiai as a special case of a conjecture, which was formulated in a more general setting of quantum relative entropy. For any two states ψ, ϕ on a finite dimensional algebra $\mathcal{A}$ the relative entropy is defined through

$$S(\psi,\phi) := \begin{cases} \mathrm{tr}_{\mathcal{A}} D_\psi \log(D_\psi - D_\phi), & \mathrm{supp}(\psi) \leq \mathrm{supp}(\phi) \\ \infty, & \text{otherwise,} \end{cases}$$

where $\mathrm{supp}(\psi)$ denotes the support projector of the state ψ. The conjecture says that the mean (per site limit) quantum relative entropy between an ergodic quantum state Ψ and an i.i.d. reference state Φ on a quasi-local algebra $\mathcal{A}^\infty$ gives the first order in the exponential decrease of the expectation value with respect to Φ of the so-called minimal separating projectors. Choosing as a reference state the special i.i.d. quantum state Φ with the single-site density operator $D = \frac{1}{d^\nu}\mathbf{1} \in \mathcal{A}$, the minimal separating projectors at level ε coincide with the high probability projectors at level ε, where $\varepsilon \in (0,1)$. The conjecture in its general form can be interpreted in the context of (quantum) hypothesis testing. Recently, in continuation of [11] the conjecture has been proven, [13].

Now, we sketch the main steps of the proof of Theorem 3.3.3 and Proposition 3.3.1, respectively. One basic tool for the proof of the statements under the general assumption of ergodicity is the structural assertion in Theorem 3.2.1. It is used to circumvent the complete ergodicity assumption, as already mentioned in the last section. Theorem 3.2.1 combined with the subsequent lemma allow to control not only the mean (per site limit) entropies of the ergodic components (with respect to the sublattice obtained by a coarsening of the lattice $\mathbb{Z}^\nu$ into larger boxes), but also to cope with the obstacle that some of these components might have an atypical entropy on these large but finite boxes. Using these prerequisites, we prove Lemma 3.3.5 which is the extension of the Hiai/Petz upper bound result for completely ergodic states to the case of ergodic states. Finally, from the probabilistic argument expressed in Lemma 2.2.6 we derive 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 $\Psi \in \partial_{ex}\mathcal{T}(\mathcal{A}^\infty)$ 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 ergodic decomposition theorem 3.2.1 we know that

$$s(\Psi_{\mathbf{x}}, G_l) = s(\Psi, G_l) = l^\nu \cdot s(\Psi), \qquad \forall \mathbf{x} \in \Lambda(\mathbf{k}(l)). \tag{3.16}$$

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

$$A_{l,\eta} := \{\mathbf{x} \in \Lambda(\mathbf{k}(l)) |\ s_{\mathbf{x}}^{(l)} \geq s + \eta\}. \tag{3.17}$$

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 3.3.4 *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 3.3.4: 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 (3.17) 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:

$$\begin{aligned} \#A^c_{l_j,\eta_0} \cdot \min_{\mathbf{x}\in A^c_{l_j,\eta_0}} s_{\mathbf{x}}^{(l_j)} &\geq \#A^c_{l_j,\eta_0} \cdot \min_{\mathbf{x}\in A^c_{l_j,\eta_0}} \frac{1}{l_j^\nu} s(\Psi_{\mathbf{x}}, G_{l_j}) \\ &= \#A^c_{l_j,\eta_0} \cdot s(\Psi) \qquad (\text{by}(3.16)). \end{aligned}$$

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, $a = 0$. $\square$

The subsequent lemma states that in the case of any ergodic system $(\mathcal{A}^\infty, \Psi, T(\mathbb{Z}^\nu))$ for large $\Lambda \subset \mathbb{Z}^\nu$ there exist orthogonal projectors in $\mathcal{A}_\Lambda$ which have an expectation value (with respect to Ψ) close to 1 and simultaneously a not too large dimension of the corresponding Hilbert subspace. Precisely, it is required that the exponent of the dimension is bounded from above in terms of the mean von Neumann entropy $s(\Psi)$, namely by a value of order $\#(\Lambda) \cdot s(\Psi)$.

Lemma 3.3.5 *Let Ψ be an ergodic state on $\mathcal{A}^\infty$. Then for every $\varepsilon \in (0,1)$*

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

The strategy to verify the existence of projectors $q \in \mathcal{A}^{(\mathbf{n})}$ with $\Psi(q) \geq 1-\varepsilon$ and $\log \mathrm{tr} q < \#(\Lambda(\mathbf{n}))s(\Psi) + \varepsilon$ is based on approximating non-commutative spin lattice systems by abelian subsystems, to which one can apply the classical results (SM(B)-theorem). The approximation is in the sense of mean entropy. This method was successfully applied by Hiai and Petz in [23] to obtain the statement of Lemma 3.3.5 in a more general setting of relative entropy, as shortly discussed above, however under the restrictive assumption of complete ergodicity of Ψ.

An abelian subsystem of a non-commutative system $(\mathcal{A}^\infty, \Psi, T(\mathbb{Z}^\nu))$ can be obtained in the simplest case as follows. First we choose an abelian subalgebra $\mathcal{B} \subseteq \mathcal{A}$. Next we construct over the lattice $\mathbb{Z}^\nu$ a quasi-local algebra $\mathcal{B}^\infty$ from $\mathcal{B}$, as described in the previous section. Then $\mathcal{B}^\infty$ is a commutative subalgebra of $\mathcal{A}^\infty$ and the system $(\mathcal{B}^\infty, m, T(\mathbb{Z}^\nu))$, where $m := \Psi \upharpoonright \mathcal{B}^\infty$ is the reduction of the state Ψ onto the algebra $\mathcal{B}^\infty$, is an abelian subsystem. However, to ensure that an abelian subsystem approximates the original non-commutative system in mean entropy it is necessary to consider a sublattice of $\mathbb{Z}^\nu$ typically of the form $l\cdot\mathbb{Z}^\nu$, $l > 1$ integer, and subsequently to construct from a suitable maximal abelian subalgebra $\mathcal{B} \subset \mathcal{A}_{\Lambda(l)}$ an abelian subsystem $(\mathcal{B}^\infty, m_l := \Psi \upharpoonright_{\mathcal{B}^\infty}, T(l\cdot\mathbb{Z}^\nu))$ of $(\mathcal{A}^\infty_{\Lambda(l)}, \Psi \upharpoonright_{\mathcal{A}^\infty_{\Lambda(l)}}, T(l\cdot\mathbb{Z}^\nu))$, as described above. Next, if we want to apply the classical SM-theorem to the abelian subsystem we need the ergodicity of m_l

(with respect to the subgroup $l \cdot \mathbb{Z}^\nu$). However, the abelian subsystem inherits the ergodic properties from $(\mathcal{A}^\infty_{\Lambda(l)}, \Psi \upharpoonright_{\mathcal{A}^\infty_{\Lambda(l)}}, T(l \cdot \mathbb{Z}^\nu))$. This can be derived from Theorem 4.3.17 in [14]. Consequently it fails to be $l \cdot \mathbb{Z}^\nu$-ergodic if we do not suppose complete ergodicity of Ψ on $\mathcal{A}^\infty$. The ergodic decomposition theorem (Theorem 3.2.1) allows to cope with this problem.

Proof of Lemma 3.3.5: 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 3.3.4 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 (3.17). This inequality implies

$$\frac{1}{\#\Lambda(\mathbf{k}(l))} \#A^c_{l,\eta} \cdot (1 - \frac{\varepsilon}{2}) \geq 1 - \varepsilon. \tag{3.18}$$

On the other hand by

$$S(\Psi^{(\mathbf{n})}) = \inf\{S(\Psi^{(\mathbf{n})} \upharpoonright \mathcal{B}) |\ \mathcal{B} \text{ maximal abelian } C^* - \text{subalgebra of } \mathcal{A}^{(\mathbf{n})}\}$$

(cf. Theorem 11.9 in [38] 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

$$\frac{1}{\#\Lambda(l)} S(\Psi^{(\Lambda(l))}_\mathbf{x} \upharpoonright \mathcal{B}_\mathbf{x}) < s(\Psi) + \eta, \qquad \forall\ \mathbf{x} \in A^c_{l,\eta}. \tag{3.19}$$

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

$$m_\mathbf{x} := \Psi_\mathbf{x} \upharpoonright \mathcal{B}^\infty_\mathbf{x} \text{ and } m^{(\mathbf{n})}_\mathbf{x} := \Psi_\mathbf{x} \upharpoonright \mathcal{B}^{(\mathbf{n})}_\mathbf{x}$$

for $\mathbf{x} \in A^c_{l,\eta}$ 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 quasi-local algebra. This easily follows from Theorem 4.3.17. in [14]. Moreover, by the Gelfand isomorphism and Riesz representation theorem, we identify the states $m_\mathbf{x}$ with probability measures on corresponding (compact) maximal ideal spaces of $\mathcal{B}^\infty_\mathbf{x}$. By commutativity and finite dimensionality of the algebras $\mathcal{B}_\mathbf{x}$ these compact spaces can be represented as $B^\infty_\mathbf{x}$ with finite sets $B_\mathbf{x}$ for all $\mathbf{x} \in A^c_{l,\eta}$. By the classical SMB-theorem (cf. Theorem 2.11)

$$\lim_{\Lambda(\mathbf{n}) \nearrow \mathbb{N}^\nu} -\frac{1}{\#\Lambda(\mathbf{n})} \log m^{(\mathbf{n})}_\mathbf{x}(\omega_\mathbf{n}) = h_\mathbf{x} \tag{3.20}$$

$m_\mathbf{x}$-almost surely for all $\mathbf{x} \in A^c_{l,\eta}$, where $h_\mathbf{x}$ denotes the mean Shannon entropy of $m_\mathbf{x}$, and $\omega_\mathbf{n} \in {B_\mathbf{x}}^{\Lambda(\mathbf{n})}$ are the components of $\omega \in B^\infty_\mathbf{x}$ corresponding to the box $\Lambda(\mathbf{n})$. Actually, as we shall see, we need the theorem cited above only in its weaker form as convergence in probability, i.e. we need only the SM-theorem. For each n and $\mathbf{x} \in A^c_{l,\eta}$ let

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

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

$$\#C^{(\mathbf{n})}_{\mathbf{x}} = \mathrm{tr}_{\mathbf{n}}\left(p^{(\mathbf{n})}_{\mathbf{x}}\right) \leq e^{\#\Lambda(\mathbf{n})\cdot(h_{\mathbf{x}}+\delta)} \leq e^{\#\Lambda(\mathbf{n})\cdot(l^\nu(s(\Psi)+\eta)+\delta)} \tag{3.21}$$

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

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

(cf. [42]), and by (3.19 for the second equality). Recall that H denotes the Shannon entropy.
From (3.20) it follows that there is an $N \in \mathbb{N}$ (depending on l) such that for all $\mathbf{n} \in \mathbb{N}^\nu$ with $\Lambda(\mathbf{n}) \supset \Lambda(N)$

$$m^{(\mathbf{n})}_{\mathbf{x}}(C^{(\mathbf{n})}_{\mathbf{x}}) \geq 1 - \frac{\varepsilon}{2}, \qquad \forall\, \mathbf{x} \in A^c_{l,\eta}. \tag{3.22}$$

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

$$q^{(l\mathbf{n})} := \bigvee_{\mathbf{x}\in A^c_{l,\eta}} p^{(\mathbf{n})}_{\mathbf{x}}.$$

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 (3.22) and (3.18) we obtain

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

Thus the condition in the definition of $\beta_{\varepsilon,\mathbf{y}}(\Psi)$ is satisfied. Moreover by the definition of $q_{\mathbf{y}}$ and (3.21)

$$\begin{aligned} \beta_{\varepsilon,\mathbf{y}}(\Psi) &\leq \log \mathrm{tr}_{\mathbf{y}} q_{\mathbf{y}} = \log \mathrm{tr}_{l\mathbf{n}} q^{(l\mathbf{n})} + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})} \mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}) \\ &\leq \log(\sum_{\mathbf{x}\in A^c_{l,\eta}} e^{\#\Lambda(\mathbf{n})\cdot(h_{\mathbf{x}}+\delta)}) + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})} \mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}) \\ &\leq \log(\#A^c_{l,\eta} \cdot e^{\#\Lambda(\mathbf{n})(l^\nu(s(\Psi)+\eta)+\delta)}) \qquad \text{(by (3.21))} \\ &\quad + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})} \mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}) \\ &\leq \log(\#A^c_{l,\eta}) + \#\Lambda(\mathbf{n})(l^\nu(s(\Psi)+\eta)+\delta) \\ &\quad + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})} \mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}) \\ &\leq \log(\#A^c_{l,\eta}) + \#\Lambda(l\mathbf{n})(s(\Psi)+\eta+\frac{\delta}{l^\nu}) \\ &\quad + \log(\mathrm{tr}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})} \mathbf{1}_{\Lambda(\mathbf{y})\setminus\Lambda(l\mathbf{n})}). \end{aligned}$$

We can conclude from this that

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

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

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

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

Proof of Proposition 3.3.3: $\mathcal{A}^{(\mathbf{n})}$ 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 1-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 $\bigoplus_{j=1}^M \mathcal{B}(\mathcal{H}_j^{(\mathbf{n})}) \subseteq \mathcal{B}(\mathcal{H}^{(\mathbf{n})})$. Consider the spectral representation of the density operator $D_\mathbf{n}$ of $\Psi^{(\mathbf{n})}$ in $\mathcal{B}(\mathcal{H}^{(\mathbf{n})})$:

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

i.e. $\lambda_i^{(\mathbf{n})} \in [0,1]$ are eigen-values and $q_i^{(\mathbf{n})} \in \mathcal{B}(\mathcal{H}^{(\mathbf{n})})$ are (minimal) eigen-projectors.
For $\mathbf{n} = (n_1, \dots, n_\nu) \in \mathbb{N}^\nu$ let $A^{(\mathbf{n})}$ be the finite set consisting of the eigen-projectors $q_i^{(\mathbf{n})}$ of $\Psi^{(\mathbf{n})}$, i.e.

$$A^{(\mathbf{n})} := \{q_i^{(\mathbf{n})}\}_{i=1}^{d_\mathbf{n}}. \tag{3.23}$$

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

$$P^{(\mathbf{n})}(q_i^{(\mathbf{n})}) := \Psi^{(\mathbf{n})}(q_i^{(\mathbf{n})}) = \lambda_i^{(\mathbf{n})}. \tag{3.24}$$

Recall that $|\mathbf{n}| = \prod_{i=1}^\nu n_i$. Let $M := \log(\dim\mathcal{H}^{(0)})$, then $\frac{1}{|\mathbf{n}|}\log\#A^{(\mathbf{n})} \le M$ 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 2.2.6 and consequently

$$\lim_{\mathbf{n}\to\infty}\frac{1}{|\mathbf{n}|}\alpha_{\varepsilon,\mathbf{n}}(P^{(\mathbf{n})}) = \lim_{\mathbf{n}\to\infty}\frac{1}{|\mathbf{n}|}H(P^{(\mathbf{n})}), \qquad \forall\varepsilon\in(0,1). \tag{3.25}$$

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

$$h := \lim_{\mathbf{n}\to\infty}\frac{1}{|\mathbf{n}|}H(P^{(\mathbf{n})}) = s(\Psi). \tag{3.26}$$

Next assume the following ordering:

$$i < j \implies \lambda_i^{(\mathbf{n})} \ge \lambda_j^{(\mathbf{n})}$$

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

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

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

$$\alpha_{\varepsilon,\mathbf{n}}(P^{(\mathbf{n})}) = \beta_{\varepsilon,\mathbf{n}}(\Psi^{(\mathbf{n})}), \qquad \forall \varepsilon \in (0,1). \tag{3.27}$$

From $\Psi^{(\mathbf{n})}(\sum_{i=1}^{n_{\varepsilon,\mathbf{n}}} q_i^{(\mathbf{n})}) \geq 1-\varepsilon$ and $\mathrm{tr}_{\mathbf{n}} \sum_{i=1}^{n_{\varepsilon,\mathbf{n}}} q_i^{(\mathbf{n})} = n_{\varepsilon,\mathbf{n}}$ 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 projector $q \in \mathcal{A}^{(\mathbf{n})}$ with $\Psi^{(\mathbf{n})}(q) \geq 1-\varepsilon$ such that $m := \mathrm{tr}_{\mathbf{n}}\, q < n_{\varepsilon,\mathbf{n}}$. Let $\sum_{i=1}^{m} q_i$, where $q_i \in \mathcal{B}(\mathcal{H}^{(\mathbf{n})})$, be the spectral representation of q. For $D_{\mathbf{n}}$ as density matrix on $\mathcal{H}^{(\mathbf{n})}$ we use Ky Fan's maximum principle, [8], and obtain the contradiction

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

Ψ is ergodic. Thus we can apply Lemma 3.3.5:

$$\limsup_{\Lambda(\mathbf{n})\nearrow\mathbb{N}^\nu} \frac{1}{\#(\Lambda(\mathbf{n}))} \beta_{\varepsilon,\mathbf{n}}(\Psi^{(\mathbf{n})}) \leq s(\Psi), \qquad \forall \varepsilon \in (0,1). \tag{3.28}$$

Setting (3.27) and (3.26) in (3.28) and using that $\#(\Lambda(\mathbf{n})) = |\mathbf{n}|$ we obtain

$$\limsup_{\mathbf{n}\to\infty} \frac{1}{|\mathbf{n}|} \alpha_{\varepsilon,\mathbf{n}}(P^{(\mathbf{n})}) \leq h, \qquad \forall \varepsilon \in (0,1). \tag{3.29}$$

With (3.26) and (3.29) both conditions in Lemma 2.2.6 are satisfied. It follows (3.25). Now we set back (3.27) and (3.26) in (3.25) and arrive at

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

Proof of the Quantum Shannon-McMillan Theorem:
Fix $\delta > 0$. Adopt the family $\{(A^{(\mathbf{n})}, P^{(\mathbf{n})})\}_{\mathbf{n}\in\mathbb{N}^\nu}$ and further notations from the proof of Proposition 3.3.3. Choose some $\delta' < \delta$. Let $A_2^{(\mathbf{n})}(\delta')$ be the subset of $A^{(\mathbf{n})}$ defined in the proof of Lemma 2.2.6 with $h = s(\Psi)$, appropriate to (3.26). 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 (2.15) 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 projector $0 \neq p \in \mathcal{A}^{(\mathbf{n})}$ with $p \leq p_{\mathbf{n}}(\delta)$ is represented as a projector onto a 1-dimensional subspace of $\mathcal{H}^{(\mathbf{n})}$ spanned by a (normalized) vector $v_p \in \mathcal{H}^{(\mathbf{n})}$ such that $v_p = \sum_{i\in I_{\mathbf{n}}(\delta')} \gamma_i v_{q_i^{\mathbf{n}}}$, where $v_{q_i^{\mathbf{n}}} \in \mathcal{H}^{(\mathbf{n})}$ are corresponding to the eigen-projectors $q_i^{\mathbf{n}}$ and $\sum_{i\in I_{\mathbf{n}}(\delta')} |\gamma|_i^2 = 1$. 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

$$e^{-\#\Lambda(\mathbf{n})(s(\Psi)+\delta)} < \Psi^{(\mathbf{n})}(p) < e^{-\#\Lambda(\mathbf{n})(s(\Psi)-\delta)}. \tag{3.30}$$

Using the linearity of $\Psi^{(\mathbf{n})}$ and applying (3.30) to the projections $q_i^{(\mathbf{n})}$ we arrive at the following estimation

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

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

Concluding this section we remark that for an i.i.d. quantum state Ψ on $\mathcal{A}^\infty$ the quantum SM-theorem (and even the finite reformulation of Breiman's extension) can be directly derived from the classical SM(B)-theorem. This is due to the fact that the family $\{(A^{\mathbf{n}}, P^{\mathbf{n}})\}_{\mathbf{n}\in\mathbb{N}^\nu}$ associated to the spectral decompositions of the density operators $D_{\mathbf{n}}$, respectively, as introduced in the proof of Proposition 3.3.3 (cf. definitions (3.23) and (3.24)), is consistent and consequently forms an (ergodic) stochastic process, to which we can directly apply the classical SM(B)-theorem. In general, the whole class of quantum states, which can be obtained from ergodic stochastic processes by mapping the alphabet of the process onto a set of mutually commuting 1-dimensional projectors on a finite dimensional Hilbert space (classical-quantum coding) allows such a simple derivation of the quantum SM(B)-theorems. These states however do not possess any quantum correlations.

3.4 Quantum Data Compression Theorem

In this section we demonstrate how the quantum results from the previous section can be applied to quantum information theory. In analogy to classical information theory 1-dimensional quantum spin lattice systems $(\mathcal{A}^\infty, \Psi, T)$, as presented in Section 3.1, model (discrete-time) quantum information sources. The goal in the present section is to formulate and prove an asymptotically lossless data compression theorem for ergodic quantum information sources.

One of the interests in quantum information theory is an economical and error-free storage or transmission of quantum information. In other words the question is: What is the minimal amount of resources measured in units of qubits or equivalently in Hilbert space dimensions needed to store quantum states faithfully? This question has been solved in the case of i.i.d. sources using the entanglement fidelity F_e as a criterion for reliability, [38]: Any compression scheme possessing a rate smaller than the (base 2) von Neumann entropy S of the single-site density operator D cannot be reliable in the sense that the entanglement fidelity tends to 0. Here by rate we mean the asymptotic number of qubits per signal used to represent the quantum information source (QIS). It has been shown in [4] that for encodings of classical memoryless sources into some fixed set of pure quantum states, as will be described below, an analogous assertion holds. In this case the reliability is measured by the ensemble fidelity $\bar{F}$. In both cases compression schemes have been constructed with rates, that can be made arbitrary close to the von Neumann entropy S of the single-site density operator determining the i.i.d. source. An essential ingredient is the quantum AEP for i.i.d. information sources. With the quantum AEP 3.3.1 or especially with the related Proposition 3.3.3 for general ergodic quantum lattice systems we dispose of a necessary tool to extend the asymptotically lossless data compression theorem to the more general case of ergodic quantum sources. It turns out that again the von Neumann entropy rate s is the optimal rate in the sense that a rate $R \geq s$ is a necessary condition for asymptotical reliability of compression schemes and even more the fidelity of any compression scheme with rate $R < s$ vanishes asymptotically. Of course, this result depends on the underlying fidelity notion for the compression/decompression operations on the QIS. Our results hold if reliability is measured by $\bar{F}$ or by F_e. Moreover, we provide an asymptotically *optimal* compression scheme based on the concept of high probability subspaces. We show by construction (cf. Theorem 3.4.1 below) that for any ergodic QIS there exists an asymptotically reliable compression scheme with rate R equal to the von Neumann entropy rate s. The main tool to construct compression schemes achieving the optimal rate s are high probability subspaces. Compression maps that are essentially projections onto high probability subspaces provide the optimal data compression. A basic result concerning high probability subspaces was proven in the previous section. Due to this result the mean (per site) minimal logarithmic dimension of these subspaces converges to the von Neumann entropy rate s in the case of ergodic QIS.

The concept of high probability subspaces is also crucial in the work of Petz and Mosonyi [41], where they prove a coding theorem for the class of completely ergodic QIS. Using projections onto high probability subspaces they show that completely ergodic QIS can be compressed with any rate $R \geq s$ in such a way that the ensemble fidelity $\bar{F}$ is asymptotically equal to 1. On the other hand $\bar{F}$ cannot achieve 1 asymptotically if the rate satisfies $R < s$. The reason why they cannot conclude that for $R < s$ the asymptotical fidelity $\bar{F}$ is equal to 0 is that they use the result of Hiai/Petz [23] which provides bounds on limit superior and limit inferior and not the limit of the minimal logarithmic dimension rate of the high probability subspaces. The result of Petz/Mosonyi represents an extension of the coding theorem formulated in [29] by Jozsa and Schumacher for the smaller class of independent identically distributed (i.i.d.) QIS and proved in [29] and [4]. An analogous result for i.i.d. QIS using the entanglement fidelity F_e as a criterion for the reliability of compression schemes is presented by Nielsen and Chuang in [38].

Preliminary Remarks

Compared to the previous sections the notations in the present section are much simpler due to the fact that we deal with the 1-dimensional lattice $\mathbb{Z}$ only. To avoid further technical complications we propose to restrict our considerations to the case of finite dimensional C*-algebras of the form $\mathcal{A} = \mathcal{B}(\mathcal{H})$, where $\mathcal{B}(\mathcal{H})$ is the algebra of linear operators on the Hilbert space $\mathcal{H}$ of dimension $\dim \mathcal{H} = d < \infty$. Recall that in the most general case, including abelian algebras corresponding to the classical situation, the finite dimensional algebra $\mathcal{A}$ is *-isomorphic to a finite direct sum $\bigoplus_{j=1}^{M} \mathcal{B}(\mathcal{H}_j)$, where each $\mathcal{H}_j$ is a Hilbert space with $\dim \mathcal{H}_j = d_j < \infty$. Under the proposed assumption for any $n \in \mathbb{N}$

the local algebra $\mathcal{A}^{(n)}$ is *-isomorphic to $\mathcal{B}(\mathcal{H}^{\otimes n})$.
Moreover, we use in this section the Dirac notation $|\phi\rangle$ for vectors in $\mathcal{H}$ and $\langle\phi|$ for the corresponding dual vectors.
We will compute the von Neumann entropy with respect to the base 2 logarithm $\log_2$, relating to the qubit as the standard quantum information unit.
All notational renewals are intended on the one hand for simplifications, on the other hand to fit in the standard (notational) framework of quantum information theory (see for example [38]).

3.4.1 Data Compression Schemes

In order to define lossless data compression schemes for encoding quantum signals we need the concept of *trace preserving quantum operations.* A physical approach to trace preserving quantum operations can be obtained as follows. Consider a quantum system S prepared in some state ρ corresponding to the density operator D_ρ acting on the Hilbert space $\mathcal{H}$. We imagine that this system interacts with its environment, a quantum system S_{env} in a state ρ_{env} on the finite dimensional Hilbert space $\mathcal{H}_{\mathrm{env}}$. The system $S \times S_{\mathrm{env}}$ is closed and we make the assumption that it is initially in the product state with the density operator $D_\rho \otimes D_{\rho_{\mathrm{env}}}$ on $\mathcal{H} \otimes \mathcal{H}_{\mathrm{env}}$. As a state of a closed system it undergoes a unitary evolution represented by a unitary operator U on $\mathcal{H} \otimes \mathcal{H}_{\mathrm{env}}$. The corresponding evolution of the state ρ of S is usually not unitary, i.e. irreversible. It is given by a trace preserving quantum operation $\mathcal{E}$ on $\mathcal{B}(\mathcal{H})$:

$$\mathcal{E}(D_\rho) := \mathrm{tr}_{\mathcal{H}_{\mathrm{env}}}(U(D_\rho \otimes D_{\rho_{\mathrm{env}}})U^*). \tag{3.31}$$

It can be shown that each trace preserving quantum operation $\mathcal{E}$ possesses the following representation known as Kraus or sum representation (cf. [21], [22], [38])

$$\mathcal{E}(D_\rho) = \sum_i E_i D_\rho E_i^*,$$

where $E_i \in \mathcal{B}(\mathcal{H})$ and $\sum_i E_i^* E_i = \mathbf{1}$. This description contains, for example, the cases of the unitary time evolution and general measurements. For a treatment of quantum operations within the framework of completely positive linear maps between C^*-algebras we suggest [18] or [33] .

A (block) *compression scheme* $(\mathcal{C}, \mathcal{D})$ for stationary QIS is a sequence $\{(\mathcal{C}^{(n)}, \mathcal{D}^{(n)})\}_{n\in\mathbb{N}}$ of pairs of trace preserving quantum operations

$$\begin{aligned} \mathcal{C}^{(n)} : \mathcal{S}(\mathcal{H}^{\otimes n}) &\longrightarrow \mathcal{S}(\mathcal{H}^{(n)}), \\ \mathcal{D}^{(n)} : \mathcal{S}(\mathcal{H}^{(n)}) &\longrightarrow \mathcal{S}(\mathcal{H}^{\otimes n}) \end{aligned}$$

where $\mathcal{H}^{(n)}$ is a subspace of $\mathcal{H}^{\otimes n}$ for all $n \in \mathbb{N}$ and $\mathcal{S}(\mathcal{H}^{\otimes n})$, $\mathcal{S}(\mathcal{H}^{(n)})$ denote in this section the sets of density operators on $\mathcal{H}^{\otimes n}$ and $\mathcal{H}^{(n)}$, respectively. We refer to $\mathcal{C}^{(n)}$ and $\mathcal{D}^{(n)}$ as compression and decompression map, respectively.
The *rate* $R(\mathcal{C})$ of a compression scheme $(\mathcal{C}, \mathcal{D})$ is defined by

$$R(\mathcal{C}) := \limsup_{n\to\infty} \frac{1}{n} \log_2 \dim \mathcal{H}^{(n)}. \tag{3.32}$$

3.4.2 Fidelities

We need a quantitative criterion to decide if a compression scheme operates reliably. For this purpose we review in this section different relevant notions of fidelity and their basic properties.
The fidelity F between two density operators D_ρ and D_σ acting on some finite dimensional Hilbert space $\mathcal{H}$ is defined by

$$F(D_\rho, D_\sigma) := \mathrm{tr}\sqrt{\sqrt{D_\rho}D_\sigma\sqrt{D_\rho}}. \tag{3.33}$$

The fidelity is symmetric in its entries and takes values between 0 and 1 with $F(D_\rho, D_\sigma) = 0$ if, and only if, D_ρ and D_σ are supported on orthogonal subspaces. The value 1 is achieved only in the case $D_\rho = D_\sigma$. In view of these properties it is reasonable to interpret the fidelity as a measure of distinguishability between two density operators which reduces to the well known overlap $|\langle\psi|\phi\rangle|$ in the case of pure states $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$ on $\mathcal{H}$. Moreover F is jointly concave and increasing under trace preserving quantum operations. The proofs of these facts may be found in [38]. The fidelity F is related to the familiar trace distance of two density operators by:

$$1 - F(D_\rho, D_\sigma) \leq \frac{1}{2}\mathrm{tr}|D_\rho - D_\sigma| \leq \sqrt{1 - (F(D_\rho, D_\sigma))^2}, \tag{3.34}$$

(cf. [38]). The trace distance can be represented as (cf. [38])

$$\frac{1}{2}\mathrm{tr}|D_\rho - D_\sigma| = \max\{\mathrm{tr}(P(D_\rho - D_\sigma)) : P = P^* = P^2\}.$$

This equality has the following meaning: the orthogonal projections appearing in the above equation are usually interpreted as ideal "yes-no" measurements. The outcome "yes" (respectively "no") is represented by P (respectively $\mathbf{1} - P$). The trace distance quantifies the largest difference of probabilities for obtaining outcome "yes" if we perform measurements on quantum systems in the states ρ and σ, respectively. This relation between the fidelity and the trace distance gives us an idea about the operational interpretation of the fidelity.

The question how well the state of an open quantum system is preserved by a time evolution, a measurement or more generally an arbitrary quantum operation leads to several fidelity concepts.

Entanglement Fidelity

We begin with the concept of *entanglement fidelity* F_e. It is a function of a density operator D_ρ and a quantum operation $\mathcal{E}$. It is defined by

$$F_e(D_\rho, \mathcal{E}) := [F(|\psi\rangle\langle\psi|, (\mathbf{1} \otimes \mathcal{E})(|\psi\rangle\langle\psi|))]^2, \tag{3.35}$$

where $|\psi\rangle \in \mathcal{H}' \otimes \mathcal{H}$ is an arbitrary purification of D_ρ, i.e. $\mathrm{tr}_{\mathcal{H}'}|\psi\rangle\langle\psi| = D_\rho$. It can be shown that this definition does not depend on the particular choice of the purification of D_ρ, cf. [38]. Entanglement fidelity measures how well the purifications of a given state are preserved under quantum operations $\mathcal{E}$. If the state is mixed then all purifications are entangled pure states.

Now, we present a further formula for the entanglement fidelity, which turns out to be very useful in the proof of the quantum data compression theorem 3.4.1. Let $\mathcal{E}(D_\rho) = \sum_i E_i D_\rho E_i^*$ be the sum representation of $\mathcal{E}$ where $E_i \in \mathcal{B}(\mathcal{H})$ and $\sum_i E_i^* E_i = \mathbf{1}$. Then

$$F_e(D_\rho, \mathcal{E}) = \sum_i |\mathrm{tr} D_\rho E_i|^2. \tag{3.36}$$

This formula implies that the entanglement fidelity is a convex function of the density operator. Indeed, the last expression is merely the squared norm of a complex vector with the components $\mathrm{tr}(D_\rho E_i)$, which depend affinely on D_ρ. Moreover, every norm is a convex function, so we obtain the claimed convexity of F_e.

Ensemble fidelity

In order to define the *ensemble fidelity* $\bar{F}$ we start with a finite set of symbols $\{1, \ldots, n\}$ (a classical alphabet) which are drawn according to a probability distribution $(p_1, \ldots, p_n)$. We associate to this set of symbols a fixed set of density operators $\{D_1, \ldots D_n\}$ on $\mathcal{H}$ and define the ensemble fidelity by

$$\bar{F}(\{(p_i, D_i)\}_{i=1}^n, \mathcal{E}) := \sum_{i=1}^n p_i \left(F\left(D_i, \mathcal{E}(D_i)\right)\right)^2, \tag{3.37}$$

where $\mathcal{E}$ is a quantum operation. The weighted ensemble of n quantum states $\{(p_i, D_i)\}_{i=1}^n$ represents a convex decomposition of the density operator $D_\rho = \sum_{i=1}^n p_i D_i$. If the D_i are all pure states then we call the ensemble or the convex decomposition a pure one. We will denote by $F_s(D_\rho, \mathcal{E})$ the supremum over pure convex decompositions of the ensemble fidelities for a density operator D_ρ and a quantum operation $\mathcal{E}$:

$$F_s(D_\rho, \mathcal{E}) := \sup\{\bar{F}(\{(p_i, P_i)\}_{i=1}^n, \mathcal{E}) \quad : \quad \{(p_i, P_i)\}_{i=1}^n \text{ pure convex decomposition of } D_\rho\}. \tag{3.38}$$

The idea behind the definition (3.37) is that the classical alphabet is represented by quantum systems prepared in the states from some fixed set. For example we can encode the alphabet $\{0, 1\}$ into two different polarization directions of photons. The probability of occurrence of each polarization direction is determined by the probability distribution on the classical alphabet. The ensemble fidelity $\bar{F}$ appears mainly in problems concerning classical information to be e.g. stored on or transmitted via quantum states.

We conclude this section with a relation among the fidelity notions introduced here. For a fixed quantum state ρ with corresponding density operator D_ρ we define

$$\bar{F}_{\rho,\mathcal{E}} := \{\bar{F}(\{(p_i, D_i)\}_i, \mathcal{E}) | \sum_i p_i D_i = D_\rho\}.$$

It holds

$$0 \le F_e(D_\rho, \mathcal{E}) \le \bar{F} \le F(D_\rho, \mathcal{E}(D_\rho)) \le 1, \quad \bar{F} \in \bar{F}_{\rho,\mathcal{E}}. \tag{3.39}$$

The second inequality follows immediately from the convexity of the entanglement fidelity. The third inequality holds because the fidelity F is jointly concave. Observe that according to (3.39) we can give upper and lower bounds for $\bar{F}$ which depend exclusively on the density operator D_ρ corresponding to the convex decomposition in consideration. The inequality (3.39) will play a crucial role in our derivation of data compression theorem.

3.4.3 Lossless Data Compression Theorem for Ergodic Quantum Information Sources

Now we are in a position to formulate and prove the asymptotically lossless data compression theorem for ergodic QIS.
Recall that according to the assumption $\mathcal{A} = \mathcal{B}(\mathcal{H})$ any ergodic quantum state Ψ on $\mathcal{A}^\infty$ corresponds one-to-one to a consistent family of density operators $\{D_n\}_{n\in\mathbb{N}}$ on $\mathcal{H}^{\otimes n}$, respectively.

Theorem 3.4.1 (Data Compression Theorem) *Let $(\mathcal{A}^\infty, \Psi, T)$ be an ergodic quantum information source with entropy rate $s(\Psi)$ and $\{D_n\}_{n\in\mathbb{N}}$ the family of density operators corresponding to Ψ.*

1) Each compression scheme $(\mathcal{C}, \mathcal{D})$ satisfying

$$\lim_{n\to\infty} \bar{F}(\{(\lambda_i^{(n)}, P_i^{(n)})\}_{i=1}^{k_n}, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}) = 1 \tag{3.40}$$

for some sequence $\{\{(\lambda_i^{(n)}, P_i^{(n)})\}_{i=1}^{k_n}\}_{n\in\mathbb{N}}$ of pure convex decompositions of D_n, respectively, fulfills

$$R(\mathcal{C}) \geq s(\Psi).$$

2) There exists a compression scheme $(\mathcal{C}, \mathcal{D})$ with $R(\mathcal{C}) = s(\Psi)$ such that

$$\lim_{n\to\infty} F_e(D_n, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}) = 1.$$

3) Any compression scheme $(\mathcal{C}, \mathcal{D})$ with $R(\mathcal{C}) < s(\Psi)$ satisfies

$$\lim_{n\to\infty} F_s(D_n, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}) = 0, \tag{3.41}$$

where F_s is defined by (3.38).

Remark: Taking into account the relation (3.39) we use worst case fidelities in the separate parts of the above theorem. In this way we obtain that the von Neumann entropy rate is the optimal compression rate using the fidelity $\bar{F}$ as well as F_e.

Before we present the proof of the above theorem we sketch the main ideas. The first statement in the theorem is essentially a consequence of the monotonicity of the relative entropy (cf. [45]) and the Fannes inequality (cf. [19]). The second statement is derived from Proposition 3.3.3, which states that the asymptotic rate of $\beta_{\varepsilon,n}$ is given by the von Neumann entropy rate and does not depend on the level ε. Compression schemes $(\mathcal{C}, \mathcal{D})$ consisting of compression maps which are essentially projections onto the high probability subspaces and the canonical embedding as decompression maps posses a rate equal to the von Neumann entropy rate. So, if we combine appropriately high probability subspaces such that their corresponding levels tend to 0, we can achieve that the entanglement fidelity becomes arbitrary close to 1. This strategy leads directly to the proof of the second part of the theorem. Finally, the third statement in the above theorem can be proven using the fact that F_s is bounded from above by the maximal expectation value of projectors $P \in \mathcal{H}^{\otimes n}$ satisfying the dimension condition $\mathrm{tr}P = \dim \mathcal{H}^{(n)}$. But if the rate of a data compression scheme is asymptotically smaller than the von Neumann entropy rate then according to the Proposition 3.3.3 the expectation values of projectors providing the upper bounds for F_s must vanish asymptotically.

Proof of Theorem 3.4.1: *Proof of 1)* Fix a convex decomposition of D_n into 1-dimensional projectors $\{P_i^{(n)}\}_{i=1}^{k_n}$ corresponding to the set of weights $\{\lambda_i^{(n)}\}_{i=1}^{k_n}$. Following an idea of M. Horodecki in [26] we arrive at the following elementary inequalities using the relative entropy and its decreasing behavior with respect to the trace preserving operations (cf. [45]):

$$\begin{aligned}\log_2 \dim \mathcal{H}^{(n)} &\geq S(\mathcal{C}^{(n)}(D_n)) \\ &\geq S(\mathcal{C}^{(n)}(D_n)) - \sum_{i=1}^{k_n} \lambda_i^{(n)} S(\mathcal{C}^{(n)}(P_i^{(n)})) \\ &= \sum_{i=1}^{k_n} \lambda_i^{(n)} S(\mathcal{C}^{(n)}(P_i^{(n)}), \mathcal{C}^{(n)}(D_n)) \\ &\geq \sum_{i=1}^{k_n} \lambda_i^{(n)} S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)}), \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(D_n)) \\ &= S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(D_n)) - \sum_{i=1}^{k_n} \lambda_i^{(n)} S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)}))\end{aligned}$$

In the next step we will show that

$$\lim_{n\to\infty} \frac{1}{n} S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(D_n)) = s(\Psi), \tag{3.42}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{k_n} \lambda_i^{(n)} S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)})) = 0, \tag{3.43}$$

holds, which implies the first part of the theorem. By the Fannes inequality for density operators D_ρ and D_σ acting on a Hilbert space of dimension d (cf. [19])

$$|S(D_\rho) - S(D_\sigma)| \leq (\log_2 d)\mathrm{tr}|D_\rho - D_\sigma| + 1, \tag{3.44}$$

and by (3.34) we have

$$\frac{1}{n}|S(D_n) - S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(D_n))| \leq 2\log_2 d \sqrt{1 - \left(F(D_n, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(D_n))\right)^2} + \frac{1}{n}.$$

Employing the limit assertion (3.40) and joint concavity of the fidelity we obtain (3.42).
Fix $\varepsilon \in (0,1)$. We consider the set

$$A_\varepsilon^{(n)} := \{i \in \{1,\ldots,k_n\} | \left(F\left(P_i^{(n)}, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)})\right)\right)^2 < 1-\varepsilon\}$$

and estimate

$$\sum_{i=1}^{k_n} \lambda_i^{(n)} \left(F\left(P_i^{(n)}, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)})\right)\right)^2 \leq (1-\varepsilon) \sum_{i \in A_\varepsilon^{(n)}} \lambda_i^{(n)} + \sum_{i \in A_\varepsilon^{(n)c}} \lambda_i^{(n)}, \tag{3.45}$$

where $A_\varepsilon^{(n)c}$ denotes the complement of $A_\varepsilon^{(n)}$. We claim that for all $\varepsilon \in (0,1)$

$$\lim_{n\to\infty} \sum_{i \in A_\varepsilon^{(n)}} \lambda_i^{(n)} = 0. \tag{3.46}$$

In fact, suppose that for some $\varepsilon \in (0,1)$

$$\limsup_{n\to\infty} \sum_{i \in A_\varepsilon^{(n)}} \lambda_i^{(n)} = a > 0.$$

Then there would exist a subsequence, which we denote again by $\{A_\varepsilon^{(n)}\}_{n\in\mathbb{N}}$ for simplicity, with

$$\lim_{n\to\infty} \sum_{i \in A_\varepsilon^{(n)}} \lambda_i^{(n)} = a.$$

After taking limits in (3.45) this would imply the following contradictory inequality

$$1 \leq (1-\varepsilon)a + (1-a).$$

By (3.46), it suffices to show that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i \in A_\varepsilon^{(n)c}} \lambda_i^{(n)} S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)})) = 0.$$

For small $\varepsilon \in (0,1)$ and for n large enough we have

$$\begin{aligned}
\frac{1}{n} \sum_{i \in A_\varepsilon^{(n)c}} \lambda_i^{(n)} S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)})) &\leq \frac{1}{n} \sum_{i \in A_\varepsilon^{(n)c}} \lambda_i^{(n)} (2n \log_2(d)\sqrt{\varepsilon} + 1) \\
&\leq 2\log_2(d)\sqrt{\varepsilon} + \frac{1}{n},
\end{aligned}$$

where in the first inequality we have applied Fannes inequality (3.44) to the expressions $S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)})) = |S(P_i^{(n)}) - S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)}))|$, respectively. Since ε can be made arbitrarily small, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i \in A_\varepsilon^{(n)c}} \lambda_i^{(n)} S(\mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}(P_i^{(n)})) = 0.$$

Proof of 2) Denote by $\mathcal{P}_\varepsilon^{(n)}$ the high probability subspace of $\mathcal{H}^{\otimes n}$ with respect to Ψ at level ε. Then by Proposition 3.3.3 we have

$$\lim_{n\to\infty} \frac{1}{n} \log_2 \dim \mathcal{P}_\varepsilon^{(n)} = \lim_{n\to\infty} \frac{1}{n} \beta_{\varepsilon,n} = s(\Psi)$$

for each $\varepsilon \in (0,1)$. A simple argument shows that there exists a sequence $\varepsilon_n \searrow 0$, as $n \to \infty$, such that

$$\lim_{n\to\infty} \frac{1}{n} \beta_{\varepsilon_n,n} = s(\Psi). \tag{3.47}$$

We consider the compression scheme $(\mathcal{C}, \mathcal{D})$, where for each $n \in \mathbb{N}$ the compression map $\mathcal{C}^{(n)}$ is given by

$$\mathcal{C}^{(n)}(D_n) = P_{\varepsilon_n}^{(n)} D_n P_{\varepsilon_n}^{(n)} + \sum_{e\in S^{(n)}} |0\rangle\langle e|D_n|e\rangle\langle 0|,$$

whereby $P_{\varepsilon_n}^{(n)}$ is a high probability projector at level ε_n corresponding to the subspace $\mathcal{P}_{\varepsilon_n}^{(n)}$, $|0\rangle \in \mathcal{P}_{\varepsilon_n}^{(n)}$ and $S^{(n)}$ is an orthonormal system in $(\mathcal{P}_{\varepsilon_n}^{(n)})^\perp$. Then the compression map $\mathcal{C}^{(n)}$ is onto $\mathcal{P}_{\varepsilon_n}^{(n)}$. The decompression map $\mathcal{D}^{(n)}$ is just the canonical embedding of $\mathcal{S}(\mathcal{P}_{\varepsilon_n}^{(n)})$ into $\mathcal{S}(\mathcal{H}^{\otimes n})$. Then by (3.47) and by definition (3.32) we have

$$R(\mathcal{C}) = s(\Psi).$$

Using the formula (3.36) for F_e we obtain

$$F_e(D_n, \mathcal{C}^{(n)}) = |\mathrm{tr} D_n P_{\varepsilon_n}^{(n)}|^2 + \sum_{e\in S^{(n)}} |\mathrm{tr} D_n |0\rangle\langle e||^2 \geq |\mathrm{tr} D_n P_{\varepsilon_n}^{(n)}|^2.$$

By definition of high probability projectors $\mathrm{tr} D_n P_{\varepsilon_n}^{(n)} \geq 1 - \varepsilon_n$ for all $n \in \mathbb{N}$. Thus

$$|\mathrm{tr} D_n P_{\varepsilon_n}^{(n)}|^2 \geq (1-\varepsilon_n)^2 \geq 1 - 2\varepsilon_n.$$

Recall that $\varepsilon_n \searrow 0$, as $n \to \infty$, and thus assertion *2)* follows.

Proof of 3) Let us define for a density operator D_ρ on $\mathcal{H}$ and some integer $d \leq \dim \mathcal{H}$

$$\eta_d(D_\rho) := \max\{\mathrm{tr} D_\rho P | \ P \text{ projector on } \mathcal{H}, \mathrm{tr} P = d\}.$$

As was proven in [4], for any compression scheme $(\mathcal{C}, \mathcal{D})$ we have

$$F_s(D_n, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}) < 6 \cdot \eta_{d^{(n)}}(D_n), \qquad \forall n \in \mathbb{N},$$

where $d^{(n)} := \dim \mathcal{H}^{(n)}$. Let $\limsup_{n\to\infty} \frac{1}{n} \log_2 d^{(n)} = R(\mathcal{C}) < s(\Psi)$. Then $\lim_{n\to\infty} \eta_{d^{(n)}}(D_n) = 0$. Otherwise there would exist a sequence $\{P^{(n)}\}_{n\in\mathbb{N}}$ of projectors in $\mathcal{H}^{\otimes n}$, respectively, with asymptotically not vanishing expectation values $\mathrm{tr} P^{(n)} D_n = \eta_{d^{(n)}}(D_n)$ and $\lim_{n\to\infty} \frac{1}{n} \log_2 \mathrm{tr} P^{(n)} = R(\mathcal{C}) < s(\Psi)$. This would be a contradiction to Proposition 3.3.3. □

3.5 A Quantum Version of Breiman's Theorem

Our starting point for a quantum SMB-theorem shall be the finite reformulation of the classical SMB-theorem, Lemma 2.2.5. Recall that the reformulation is in terms of nested typical subsets. Substituting typical subsets by typical projectors corresponding to typical subspaces of Hilbert spaces, the property of them to be nested can be expressed easily in terms of partial traces. It turns out that the proof of the quantum Shannon-McMillan theorem 3.3.1 yields the necessary tools to prove that typical subspaces can be nested.
We restrict ourselves to the case of 1-dimensional lattices $\mathbb{Z}$ in the present section. This is the relevant case in the setting of quantum information theory, where we see possible applications of the theorem. Moreover, this simplifies notations.
In the sequel we denote by $R(a)$ the support projector of a self-adjoint element a of the local algebra $\mathcal{A}_{[m,n]}$. By $\mathrm{tr}_{[k,l]}a$ we denote the partial trace of a over the local algebra $\mathcal{A}_{[k,l]} \subset \mathcal{A}_{[m,n]}$.

Theorem 3.5.1 (Quantum Shannon-McMillan-Breiman Theorem) *Let Ψ be an ergodic state on the quasi-local C^*-algebra $\mathcal{A}^\infty$ with mean entropy s. Then to each $\varepsilon > 0$ there is a sequence of orthogonal projectors $\{p_\varepsilon^{(n)}\}_{n=1}^\infty$ in $\mathcal{A}^{(n)}$, respectively, and some $N(\varepsilon)$, such that*

(q1) $p_\varepsilon^{(n)} = R(\mathrm{tr}_{n+1}\, p_\varepsilon^{(n+1)})$,

(q2) $\mathrm{tr}\, p_\varepsilon^{(n)} \in (e^{n(s-\varepsilon)}, e^{n(s+\varepsilon)})$, *for* $n \geq N(\varepsilon)$,

(q3) there exist minimal projectors $p_i \in \mathcal{A}^{(n)}$ *fulfilling* $p_\varepsilon^{(n)} = \sum_{i=1}^{\mathrm{tr}\, p_\varepsilon^{(n)}} p_i$ *and* $\Psi^{(n)}(p_i) < e^{-n(s-\varepsilon)}$, $\forall n \geq N(\varepsilon)$,

(q4) $\Psi^{(n)}(p_\varepsilon^{(n)}) > 1 - \varepsilon$.

According to the fact that each finite dimensional unital *-algebra $\mathcal{A}$ is isomorphic to $\bigoplus_{i=1}^s \mathcal{B}(\mathcal{H}_i)$, where $\mathcal{H}_i$ are finite dimensional Hilbert spaces, we may associate to each nested projector $p_\varepsilon^{(n)}$ in the theorem above a typical subspace of $\mathcal{H}^{\otimes n}$ with $\mathcal{H} := \bigoplus_{i=1}^s \mathcal{H}_i$.

Proof of Theorem 3.5.1:
1. Let $\varepsilon > 0$ be given. Choose an integer $l > 0$ sufficiently large such that the entropy of $\Psi^{(l)}$ satisfies

$$s \leq \frac{1}{l} S(\Psi^{(l)}) < s + \varepsilon^2.$$

Take a complete set V_l of mutually orthogonal eigen- projectors for $\Psi^{(l)}$. Let $\mathcal{B}$ denote the abelian subalgebra of $\mathcal{A}^{(l)}$ generated by these projectors. The completeness of V_l implies that $\mathcal{B}$ is maximal abelian. Furthermore the entropy of $\Psi^{(l)} \upharpoonright_{\mathcal{B}}$, the restriction of $\Psi^{(l)}$ to the subalgebra $\mathcal{B}$, is identical to $S(\Psi^{(l)})$. Generally we have the relation

$$S(\Psi^{(n)}) = \min\{S(\Psi^{(n)} \upharpoonright_{\mathcal{C}})|\ \mathcal{C} \subset \mathcal{A}^{(n)} \text{ max. abelian subalgebra}\}. \qquad (3.48)$$

The quasi-local algebra $\mathcal{B}^\infty$ constructed from $\mathcal{B}$ is an abelian subalgebra of $\mathcal{A}^\infty$ and Ψ acts on this algebra as a stochastic process P_l with alphabet V_l. The Shannon mean entropy h_l of this process can be estimated by

$$s \leq \frac{1}{l} h_l \leq \frac{1}{l} S(\Psi^{(l)} \upharpoonright_{\mathcal{B}}) < s + \varepsilon^2.$$

The first inequality is a consequence of (3.48). P_l is a stationary, but not necessarily ergodic process. We apply the corresponding version of the classical SMB-theorem for stationary processes, cf. [31], to this process. We obtain that there is a set of trajectories $V_l^* \subset V_l^{\mathbb{Z}}$ of measure one such that for each $(v_i)_{i\in\mathbb{Z}} \in V_l^*$ the limit (individual mean Shannon entropy) $h((v_i)_{i\in\mathbb{Z}}) := \lim_{n\to\infty} -\frac{1}{n} \log P_l^{(n)}((v_i)_{i=1}^n)$ exists and the expectation value $\mathbb{E}_{P_l} h((v_i)_{i\in\mathbb{Z}})$ is equal to h_l.

2. Let $V_l^{\varepsilon,-} \subset V_l^*$ be the subset of those trajectories, for which the relation $\frac{1}{l} h((v_i)_{i\in\mathbb{Z}}) < s - \varepsilon^2$ holds. We have $P_l(V_l^{\varepsilon,-}) = 0$. In fact, consider the sets

$$W_l^{(n),-} := \{(w_i)_{i=1}^n \in V_l^{(n)} \mid P_l^{(n)}((w_i)_{i=1}^n) > e^{-nl(s-\varepsilon^2)}\}$$

obviously containing the sets

$$\begin{aligned} V_l^{(n),-} := \{(v_i)_{i=1}^n \mid \quad & \exists (w_i)_{i\in\mathbb{Z}} \in V_l^{\varepsilon,-} \text{ with } (w_i)_{i=1}^n = (v_i)_{i=1}^n \text{ and} \\ & P_l^{(m)}((v_i)_{i=1}^m) > e^{-ml(s-\varepsilon^2)} \text{ for all } m \geq n\}, \end{aligned}$$

respectively. This means that $P_l^{(n)}(W_l^{(n),-}) \geq P_l^{(n)}(V_l^{(n),-})$. The cardinality of each $W_l^{(n),-}$ is bounded from above by $e^{nl(s-\varepsilon^2)}$. Now suppose $P_l(V_l^{\varepsilon,-}) > 0$. Then for n sufficiently large we would have $P_l^{(n)}(V_l^{(n),-}) > c$ for some $c > 0$ implying $P_l^{(n)}(W_l^{(n),-}) > c$. For the quantum state this would have the consequence that for large n there are projectors

$$p^{(nl)} := \sum_{(w_i)_{i=1}^n \in W_l^{(n),-}} \otimes_{i=1}^n w_i$$

with $\mathrm{tr}\, p^{(nl)} < e^{nl(s-\varepsilon^2)}$ and $\Psi^{(nl)}(p^{(nl)}) > c$. This contradicts Proposition 3.3.3, saying that no sequence of projectors in $\mathcal{A}^{(nl)}$, respectively, of non-zero expectation can have asymptotically a smaller trace than $e^{nl(s-\varepsilon^2)}$.

3. Let $V_l^{\varepsilon,+} \subset V_l^*$ be the subset of those trajectories, for which the relation $\frac{1}{l} h((v_i)_{i\in\mathbb{Z}}) > s + \varepsilon$ holds. By 2. and by the relation $\mathbb{E}_{P_l} h((v_i)_{i\in\mathbb{Z}}) = h_l$ we obtain

$$h_l > l(s - \varepsilon^2)(1 - P_l(V_l^{\varepsilon,+})) + l(s + \varepsilon) P_l(V_l^{\varepsilon,+})$$

resulting in $P_l(V_l^{\varepsilon,+}) < 2\varepsilon$.

4. Combining the preceeding results we can easily derive for each $\varepsilon > 0$ the existence of an l and of some $N(\varepsilon)$ such that there is a subset $\widetilde{V}_l^* \subset V_l^*$ with the properties

(a) $P_l(\widetilde{V}_l^*) > 1 - \varepsilon$,

(b) $e^{-nl(s+\varepsilon)} < P_l^{(n)}((v_i)_{i=1}^n) < e^{-nl(s-\varepsilon)}$, $\forall (v_i)_{i\in\mathbb{Z}} \in \widetilde{V}_l^*$ and $n > N(\varepsilon)$.

Indeed, assume $\varepsilon < 1$ (otherwise we would obtain the result above with ε^2 instead of ε) and set $A_{l,\varepsilon} := V_l^* \setminus (V_l^{\frac{\varepsilon}{2},-} \cup V_l^{\frac{\varepsilon}{2},+})$. We have $P_l(A_{l,\varepsilon}) > 1-\varepsilon$ and $A_{l,\varepsilon} \subseteq \bigcup_{n\geq 0} \bigcap_{k\geq n} A_{l,\varepsilon}^{(k)}$, where

$$A_{l,\varepsilon}^{(k)} := \left\{ (v_i)_{i\in\mathbb{Z}} | \ -\frac{1}{k} \log P_l^{(k)}((v_i)_{i=1}^k) \in (l(s-\varepsilon), l(s+\varepsilon)) \right\}.$$

Then there exists $N(\varepsilon) \in \mathbb{N}$ such that $P_l(\bigcap_{k\geq N(\varepsilon)} A_{l,\varepsilon}^{(k)}) > 1-\varepsilon$. The set $\widetilde{V}_l^* := \bigcap_{k\geq N(\varepsilon)} A_{l,\varepsilon}^{(k)}$ fulfills both conditions above.
Obviously $\widetilde{V}_l^*$ generates a sequence of nested sets $\{C_\varepsilon^{(n)}\}_{n=1}^\infty$ fulfilling 1-4 given in Definition 2.2.4. In the given situation, we may reformulate these properties as follows:

(a) $C_\varepsilon^{(n)} = (C_\varepsilon^{(n+1)})_\flat$ for $n \geq 1$

(b) $\#C_\varepsilon^{(n)} \in (e^{nl(s-\varepsilon)}, e^{nl(s+\varepsilon)})$ for $n \geq N(\varepsilon)$

(c) $\Psi^{(nl)}(p) < e^{-nl(s-\varepsilon)}$ for $n \geq N(\varepsilon)$ and any $p = \otimes_{k=1}^n v_k$, where $(v_k)_{k=1}^n \in C_\varepsilon^{(n)}$

(d) $\Psi^{(nl)}(\sum_{(v_k)_{k=1}^n \in C_\varepsilon^{(n)}} \otimes_{k=1}^n v_k) > 1-\varepsilon$.

Now it is easy to define nested projectors , first for multiples of l:

$$p_\varepsilon^{(nl)} := \sum_{(v_k)_{k=1}^n \in C_\varepsilon^{(n)}} \otimes_{k=1}^n v_k,$$

and then for general $n = ml + r, r < l$ by the set-up

$$p_\varepsilon^{(n)} := R\left(\mathrm{tr}_{[ml+r+1,(m+1)l]} p_\varepsilon^{((m+1)l)} \right).$$

Observe that both definitions are compatible. Obviously, by definition the property (q1) is fulfilled by the defined system of projectors. Next, we have with $n = ml + r, r < l$

$$\begin{aligned} e^{n(s-2\varepsilon)} < \#C_\varepsilon^{(m)} \leq \mathrm{tr}\ p_\varepsilon^{(n)} &\leq \#C_\varepsilon^{(m)} \mathrm{tr}\ \mathbf{1}_{\mathcal{A}^{(l)}} \\ &< e^{n(s+\varepsilon)} \mathrm{tr}\ \mathbf{1}_{\mathcal{A}^{(l)}} < e^{n(s+2\varepsilon)} \end{aligned} \tag{3.49}$$

for n sufficiently large. In fact, the first inequality in this chain is obvious. By definition we have

$$p_\varepsilon^{(ml)} = \sum_{i=1}^{\mathrm{tr}\ p_\varepsilon^{(ml)}} q_i^{(m)}$$

for certain minimal projectors $q_i^{(m)}$ from $(\mathcal{B}_{(l)})^{(m)}$, and

$$p_\varepsilon^{((m+1)l)} = \sum_{i=1}^{\mathrm{tr}\ p_\varepsilon^{(ml)}} \sum_{j=1}^{k_i} q_i^{(m)} \otimes q_{i,j}$$

for some minimal projectors $q_{i,j}$ from $\mathcal{B}_{[m+1]}$. In order to simplify our notation let

$$I(m,r) := [ml+r+1,(m+1)l].$$

We obtain

$$\begin{aligned}
\mathrm{tr}\; p_\varepsilon^{(ml+r)} &= \mathrm{tr}\; R\left(\sum_{i=1}^{\mathrm{tr}\; p_\varepsilon^{(ml)}} \sum_{j=1}^{k_i} q_i^{(m)} \otimes \mathrm{tr}_{I(m,r)} q_{i,j}\right) \\
&= \mathrm{tr} \sum_{i=1}^{\mathrm{tr}\; p_\varepsilon^{(ml)}} R\left(q_i^{(m)} \otimes \sum_{j=1}^{k_i} \mathrm{tr}\;_{I(m,r)} q_{i,j}\right) \\
&= \sum_{i=1}^{\mathrm{tr}\; p_\varepsilon^{(ml)}} \mathrm{tr}\; q_i^{(m)} \otimes R\left(\sum_{j=1}^{k_i} \mathrm{tr}\;_{I(m,r)} q_{i,j}\right) \\
&= \sum_{i=1}^{\mathrm{tr}\; p_\varepsilon^{(ml)}} \mathrm{tr}\;\; R\left(\sum_{j=1}^{k_i} \mathrm{tr}_{I(m,r)} q_{i,j}\right) \qquad (3.50) \\
&\geq \mathrm{tr}\; p_\varepsilon^{(ml)} = \# C_\varepsilon^{(m)}.
\end{aligned}$$

Here in the second step we have made use of the mutual orthogonality of the $q_i^{(m)}$. This proves the second inequality in (3.49). The third inequality also immediately follows from the formula (3.50). So (q2) is fulfilled, too (with 2ε instead of ε).

By (c) we see that (q3) is fulfilled if n is a multiple of l. In the general case $n = ml + r$ observe that in the representation

$$p_\varepsilon^{(ml+r)} = \sum_{i=1}^{\mathrm{tr}\; p_\varepsilon^{(ml)}} q_i^{(m)} \otimes R\left(\sum_{j=1}^{k_i} \mathrm{tr}_{I(m,r)} q_{i,j}\right)$$

we sum over mutually orthogonal projectors each of them fulfilling

$$\begin{aligned}
& \Psi^{(ml+r)}\left(q_i^{(m)} \otimes R\left(\sum_{j=1}^{k_i} \mathrm{tr}_{I(m,r)} q_{i,j}\right)\right) \\
\leq\; & \Psi^{(ml+r)}\left(q_i^{(m)} \otimes \mathbf{1}_{I(m,r)}\right) \\
=\; & \Psi^{(ml)}(q_i^{(m)}) < e^{-ml(s-\varepsilon)} < e^{-n(s-2\varepsilon)}
\end{aligned}$$

if n is sufficiently large. Now (q3) follows easily, again with 2ε instead of ε.

Finally, we have

$$
\begin{aligned}
& \Psi^{(ml+r)}(p_\varepsilon^{(ml+r)}) \\
= \quad & \sum_{i=1}^{\mathrm{tr}\ p_\varepsilon^{(ml)}} \Psi^{(ml+r)} \left(q_i^{(m)} \otimes R \left(\sum_{j=1}^{k_i} \mathrm{tr}_{I(m,r)} q_{i,j} \right) \right) \\
= \quad & \sum_{i=1}^{\mathrm{tr}\ p_\varepsilon^{(ml)}} \Psi^{((m+1)l)} \left(q_i^{(m)} \otimes R \left(\sum_{j=1}^{k_i} \mathrm{tr}_{I(m,r)} q_{i,j} \right) \otimes \mathbf{1}_{\mathcal{A}_{I(m,r)}} \right) \\
\geq \quad & \sum_{i=1}^{\mathrm{tr}\ p_\varepsilon^{(ml)}} \Psi^{((m+1)l)} \left(q_i^{(m)} \otimes \sum_{j=1}^{k_i} q_{i,j} \right) = \Psi^{((m+1)l)}(p_\varepsilon^{((m+1)l)}) > 1 - \varepsilon.
\end{aligned}
$$

Hereby, the first inequality can be verified using the Schmidt decomposition for 1-dimesional projectors from tensor product algebras, (e.g. [38]). The last inequality follows from (d). This proves (q4). $\square$

Chapter 4

Conclusions and Open Problems

We have shown that there exists an extension of the famous Shannon-McMillan theorem and its stronger version due to Breiman to the case of quantum spin lattice systems. These systems are mathematically modeled as C*-dynamical systems, where the dynamics is given by the action of the translation group $\mathbb{Z}^\nu$ by shifts on a quasi-local C*-algebra $\mathcal{A}^\infty$. The case of a commutative $\mathcal{A}^\infty$ corresponds to a classical spin lattice system. For such systems the dynamical entropy is given by the mean Shannon entropy. In view of our results its non-commutative generalization turns out to be the mean von Neumann entropy. In complete analogy to the classical case it appears in the quantum Shannon-McMillan theorem as the asymptotically logarithmic dimension of typical Hilbert subspaces with respect to an ergodic quantum state on $\mathcal{A}^\infty$. These typical subspaces replace the concept of typical subsets in the classical situation.

As an interesting intermediate result we have proven an ergodic decomposition theorem for quantum spin lattice systems. It gives the structure of the convex decomposition of an $\mathbb{Z}^\nu$-ergodic state into G-ergodic components, where $G = l \cdot \mathbb{Z}^\nu$, $l > 1$ integer, is a subgroup of $\mathbb{Z}^\nu$. The G-ergodic states have all the same mean von Neumann entropy, similar to the situation for classical spin lattice systems.

In the case of more general classical dynamical systems the crucial entropy notion is the Kolmogorov-Sinai dynamical entropy, which reduces to the mean Shannon entropy in the case of spin lattice systems, considered as dynamical systems. The question arises what the generalizations of the Kolmogorov-Sinai entropy and the Shannon-McMillan theorem are in the case of more general classes of C*-dynamical systems. There is no satisfying answer known to this problem. Some attempts exist to solve this problem, however only partial results could be achieved until now. For the class of asymptotically abelian C*-dynamical systems with locality (containing the here discussed class of quantum spin lattice systems) there is an analogue of the Shannon-McMillan theorem using the CNT-entropy as an extension of the classical Kolmogorov-Sinai entropy, cf. [37]. However, it is proven for a rather small class of tracial quantum states. For a quantum spin lattice system the class of tracial quantum states consists of just one i.i.d. quantum state determined by the equidistributed single-site density operator. For such an i.i.d. quantum state the mean von Neumann entropy coincides with the CNT-entropy.

In the context of (quantum) information theory the relevant dynamical systems are 1-dimensional (quantum) spin lattice systems modeling (quantum) information sources. Based on the quantum Shannon-McMillan theorem we have formulated and proven a data compression theorem for quantum information sources. The theorem states that the mean von Neumann entropy of an ergodic quantum information source gives the achievable lower bound on the compression rate for asymptotically reliably operating block compression schemes. We have quantified the reliability of compression by both, the entanglement fidelity and the ensemble fidelity. These are well established fidelity concepts in quantum information theory. Until now, except for the completely ergodic case, similar quantum data compression theorems could be proven only for i.i.d. information sources, where any kind of correlations between the quantum spins is excluded.

One aspect of Breiman's extension of the Shannon-McMillan theorem is that the family of entropy-typical subsets is nested. We have proven that this is also true in the quantum setting for typical Hilbert subspaces. Recall that the classical Shannon-McMillan-Breiman theorem is originally formulated as an almost everywhere convergence statement involving the notion of trajectory, which does not exists in the algebraic formalism. We have shown that the property of the typical subsets to be nested is equivalent to the Shannon-McMillan-Breiman theorem. In view of this equivalence it is justified to say that we have proven a quantum version of the Shannon-McMillan-Breiman theorem.

In the classical situation, the property of entropy-typical subsets to be nested can be used to design sequential data compression algorithms. The question arises if the quantum version of the Shannon-McMillan-Breiman theorem can be used as a starting point to derive quantum sequential data compression schemes. However, classical sequential coding is to a large extent also based on simultaneous measurements of the data strings to be compressed. In the quantum situation the disturbing effect of measurements of quantum states can not be excluded. Consequently, essentially adopting the ideas behind the classical algorithms will not be sufficient in the quantum situation.

Index

Symbols and Notations

Bibliography

[1] R. Alicki, Quantum Ergodic Theory and Communication Channels, Open Sys. and Information Dyn. 4, 53 (1997)

[2] R. Alicki, Information Theoretical Meaninig of Quantum Dynamical Entropy, arXiv: quant-ph/0201012

[3] R. Alicki, M. Fannes, Defining Quantum Dynamical Entropy, Lett. Math. Phys. 32, 75-82 (1994)

[4] H. Barnum, C. Fuchs, R. Jozsa, B. Schumacher, General Fidelity Limit for Quantum Channels, Phys. Rev. A 54, No 6, 4707-4711 (1996)

[5] F. Benatti, Entropy of a Subalgebra and Quantum Estimation, J.Math. Phys. 37, 5244 (1996)

[6] F. Benatti, T. Hudetz, A. Knauf, Quantum Chaos and Dynamical Entropy, Commun. Math. Phys. 198, 607-688 (1998)

[7] T. Berger, Rate Distortion Theory, Prentice-Hall (1971)

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

[9] P. Billingsley, Ergodic Theory and Information, John Wiley & Sons (1965)

[10] I. Bjelaković, A. Szkoła, The Data Compression Theorem for Ergodic Quantum Information Sources, akzeptiert für Publikation in *Quant. Inf. Proc.* (2003/2004)

[11] I. Bjelaković, T. Krüger, Ra. Siegmund-Schultze, A. Szkoła, The Shannon-McMillan Theorem for Ergodic Quantum Lattice Systems, Invent. math. 155, 203-222 (2004)

[12] I. Bjelaković, T. Krüger, Ra. Siegmund-Schultze, A. Szkoła, Chained Typical Subspaces - a Quantum Version of Breiman's Theorem, arXiv.org, quant-ph/0 301177

[13] I. Bjelaković, Ra. Siegmund-Schultze, An Ergodic Theorem for the Quantum Relative Entropy, arXiv:quant-ph/0306094

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

[15] L. Breiman, The Individual Ergodic Theorem of Information Theory, Ann. Math. Stat. 28, Nr. 3, 809-811 (1957)

[16] A. Connes, H. Narnhofer, W.Thirring, Dynamical Entropy of C^*-algebras and von Neumann Algebras, Comm. Math. Phys. 112, 691-719 (1987)

[17] J.R. Conway, A Course in Functional Analysis, Springer, New York (1985)

[18] E. B. Davies, Quantum Theory of Open Systems, Academic Press, London 1976

[19] M. Fannes, A Continuity Property of the Entropy Density for Spin Lattice Systems, Commun. Math. Phys. 31, 291-294 (1973)

[20] M. Fannes, B. Nachtergaele, R.F. Werner, Finitely Correlated States on Quantum Spin Chains, Commun. Math. Phys. 144, 443-490 (1992)

[21] K.-E. Hellwig, K.Kraus, Pure Operations and Measurements, Commun. Math. Phys. 11, 214-220 (1969)

[22] K.-E. Hellwig, K.Kraus, Operations and Measurements II, Commun. Math. Phys. 16, 142-147 (1970)

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

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

[25] F. Hiai, D. Petz, Entropy Densities for Gibbs States of Quantum Spin Systems, Rev. Math. Phys. 5 No.4, 693-712 (1993)

[26] M. Horodecki, Limits for Compression of Quantum Information by Ensembles of Mixed States, Phys. Rev. A 57, 3364-3369 (1998)

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

[28] O. Johnson, Y. Suhov, The von Neumann Entropy and Information Rate for Ideal Quantum Gibbs Ensembles, math-phys/0109023

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

[30] A. Kaltchenko, E.-H. Yang, Universal Compression of Ergodic Quantum Sources, arXiv: quant-ph/0302174

[31] G. Keller, Equilibrium States in Ergodic Theory, London Mathematical Society Student Texts 42, Cambridge (1998)

[32] C. King, A. Leśniewski, Quantum Sources and a Quantum Coding Theorem, J. Math. Phys. 39 (1), 88-101 (1998)

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

[34] J.C. Kieffer, A generalized Shannon-McMillan Theorem for the Action of an Amenable Group on a probability Space.

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

[36] B. McMillan, The Basic Theorems of Information Theory, Ann. Math. Stat. 24, 196-219 (1953)

[37] S.Neshveyev, E.Størmer, The McMillan Theorem for a Class of Asymptotically Abelian C^*-Algebras, Ergod. Th. & Dynam. Sys. 22, 889-897 (2002)

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

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

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

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

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

[43] P. C. Shields, The Ergodic Theory of Discrete Sample Paths, Graduate Studies in Mathematics Vol. 13, American Mathematical Society 1996

[44] C.E. Shannon, A Mathematical Theory of Communication, Bell System Technical Journal 27, 379-423, (1948)

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

[46] P. Walters, An Introduction to Ergodic Theory, Springer, New York 1982

Danksagung

Mein besonders herzlicher Dank gilt meinem wissenschaftlichen Betreuer Prof. Dr. Ruedi Seiler, der das Problem der Quantenversion des Shannon-McMillan-Breiman Satzes aufgezeigt, das entsprechende Interesse an dem Problem erweckt hat und mich im Verlauf meiner Doktorarbeit immer unterstützt hat.
Weiterhin möchte ich mich bei meinen wissenschaftlichen Betreuern Dr. Tyll Krüger und Dr. Rainer Siegmund-Schultze als auch bei Igor Bjelaković bedanken, mit denen in einer fruchtbaren Zusammenarbeit eine Reihe von Publikationen entstanden sind, die die Grundlage dieser Dissertation darstellen.
Schließlich gilt mein ganz besonderer Dank meinem lieben Freund Tobias Prager, der diese Arbeit nicht nur Korrektur gelesen hat, sondern mich während der gesamten Entstehungsphase der Dissertation in jeglicher Hinsicht unterstützt hat.