Commun. Math. Phys. 143, 99–114 (1991)

Communications in
Mathematical
Physics

The Proper Formula for Relative Entropy and its Asymptotics in Quantum Probability

Fumio Hiai[1] **and Dénes Petz**[2]

[1] Department of Mathematics, Ibaraki University, Mito, Ibaraki 310, Japan
[2] Mathematical Institute, Hungarian Academy of Sciences, P.O. Box 127, H-1364 Budapest, Hungary

Received February 22, 1991; in revised form June 3, 1991

Abstract. Umegaki's relative entropy $S(\omega,\varphi)=\operatorname{Tr} D_\omega(\log D_\omega-\log D_\varphi)$ (of states ω and φ with density operators D_ω and D_φ, respectively) is shown to be an asymptotic exponent considered from the quantum hypothesis testing viewpoint. It is also proved that some other versions of the relative entropy give rise to the same asymptotics as Umegaki's one. As a byproduct, the inequality $\operatorname{Tr} A\log AB \geqq \operatorname{Tr} A(\log A+\log B)$ is obtained for positive definite matrices A and B.

1. Introduction and Motivation

The relative entropy is an information quantity attached to two states of a system. In commutative (or classical) probability theory the states correspond to measures on a measurable space. When $\nu=(\nu_1,\nu_2,\ldots,\nu_n)$ and $\mu=(\mu_1,\mu_2,\ldots,\mu_n)$ are probability distributions, for the sake of simplicity, on an n-point space, the relative entropy (called also information divergence) introduced by Kullback and Leibler [17] is defined by

$$S(\nu,\mu)=\sum_i \nu_i\log\frac{\nu_i}{\mu_i}. \tag{1.1}$$

In noncommutative (or quantum) probability theory the relative entropy of normal positive functionals was first studied by Umegaki [33] in the case of semifinite von Neumann algebras as the noncommutative extension of information divergence. Later on Araki [1, 2] extended it to the case of general von Neumann algebras by means of the notion of a relative modular operator. On the other hand Uhlmann [32] introduced the relative entropy of positive functionals of arbitrary *-algebras by a quadratic interpolation method. The importance of relative entropy has been justified by the fact that one encounters this quantity in dealing with a number of different problems.

In quantum theory the states of a system correspond to positive operators of trace one. (These operators are called densities.) In particular, in the setting of matrix algebras Umegaki's relative entropy of a state ω with respect to another state φ is defined by

$$S(\omega,\varphi)=\operatorname{Tr} D_\omega(\log D_\omega-\log D_\varphi), \tag{1.2}$$

where Tr denotes the usual trace on matrices and D_ω the density of ω with respect to Tr. [Note that $S(\omega,\varphi)$ sometimes is written as $S(\varphi,\omega)$.] This definition does not seem to be canonical. Compared with (1.1) one could suppose that other expressions like

$$S_{\mathrm{co}}(\omega,\varphi)=\sup\left\{\sum_i \omega(p_i)\log\frac{\omega(p_i)}{\varphi(p_i)}: p_1,\ldots,p_n \text{ are projections}, \sum_i p_i=1\right\}, \tag{1.3}$$

$$S_{\mathrm{cp}}(\omega,\varphi)=\sup\left\{\sum_i \omega(a_i)\log\frac{\omega(a_i)}{\varphi(a_i)}: a_1,\ldots,a_n\geqq 0, \sum_i a_i=1\right\} \tag{1.4}$$

or

$$S_{\mathrm{BS}}(\omega,\varphi)=\operatorname{Tr} D_\omega \log D_\omega^{1/2}D_\varphi^{-1}D_\omega^{1/2} \tag{1.5}$$

are as good as (1.2). The quantities S_{co} and S_{cp} appeared in [8] and they may be related to observations which are projection-valued or positive operator-valued measures. The definitions (1.4) and (1.3) are of the form $\sup\{S(\omega\circ\alpha,\ \varphi\circ\alpha):\alpha\}$, where α runs over all positive (and multiplicative, respectively) unital maps of finite dimensional commutative C^*-algebras into the given matrix algebra. Given $a_1,\ldots,a_n\geqq 0$, $\sum_i a_i=1$, the inequality $\sum_i \omega(a_i)\log\omega(a_i)/\varphi(a_i)\leqq S(\omega,\varphi)$ follows by applying the monotonicity of relative entropy [16, 32] to a positive unital map $\alpha((\xi_i))=\sum_i \xi_i a_i$, $(\xi_i)\in l_n^\infty$. See also [8, 9] for this inequality. Thus the inequality $S_{\mathrm{cp}}(\omega,\varphi)\leqq S(\omega,\varphi)$ holds, while $S_{\mathrm{co}}(\omega,\varphi)\leqq S_{\mathrm{cp}}(\omega,\varphi)$ is trivial. But the equality here is very restrictive. In fact, it is known [24] that if ω and φ are faithful normal states of a von Neumann algebra $\mathscr{M}$, then ω must commute with φ whenever $S(\omega|\mathscr{N},\varphi|\mathscr{N})=S(\omega,\varphi)<+\infty$ holds for some commutative von Neumann subalgebra $\mathscr{N}$ of $\mathscr{M}$. The quantity S_{BS} was introduced in [5] (in a more general setting) and it appeared in [12] in operator form. In this paper it will be shown that the entropy quantities S_{co}, S_{cp}, and S give rise to the same asymptotic mean in the infinite tensor product system.

We want to deal with the question of the proper definition of information divergence of two states in noncommutative probability theory. This question can be approached from two essentially different points of view. One can search for plausible postulates which should be satisfied by a good notion of relative entropy and one can try to show that $S(\omega,\varphi)$ is the only functional which meets all the desiderata. In this point it was proved in [27] that up to a constant factor only Umegaki's relative entropy satisfies a reasonable set of postulates. Our approach in the present paper is more pragmatic. We consider the asymptotics of certain probabilities and observe that Umegaki's relative entropy naturally shows up.

In Sect. 2 of this paper we state the main results in the framework of finite dimensional C^*-algebras. Let $\mathscr{A}$ be a finite dimensional C^*-algebra (i.e. a finite direct sum of matrix algebras) with a fixed state φ. As the reference state we take the product state $\varphi_\infty=\bigotimes_{-\infty}^{\infty}\varphi$ on the infinite C^*-tensor product $\mathscr{A}_\infty=\bigotimes_{-\infty}^{\infty}\mathscr{A}$. Let ψ be a stationary (i.e. invariant for the right shift) state of $\mathscr{A}_\infty$. Then one has the mean relative entropy $S_{\mathrm{M}}(\psi,\varphi_\infty)=\lim_{n\to\infty}\frac{1}{n}S(\psi_n,\varphi_n)$, where $\psi_n=\psi\Big|\bigotimes_1^n\mathscr{A}$ and

$\varphi_n = \varphi_\infty \Big| \bigotimes_1^n \mathscr{A}$. The mean relative entropy plays an important role in classical as well as quantum statistical mechanics and behaves as a rate function in limit theorems of large deviation type (cf. [11, 25, 28, 29]). Our first theorem says that $\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{co}}(\psi_n, \varphi_n) = S_{\mathrm{M}}(\psi, \varphi_\infty)$ for every stationary state ψ of $\mathscr{A}_\infty$. In particular we have $\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{co}}(\omega_n, \varphi_n) = S(\omega, \varphi)$ for every state ω of $\mathscr{A}$. This has an interesting corollary that $S_{\mathrm{BS}}(\omega, \varphi) \geqq S(\omega, \varphi)$ for all states ω and φ of $\mathscr{A}$.

Moreover for $n \geqq 1$ and $0 < \varepsilon < 1$ let us introduce the following quantities:

$$\beta_\varepsilon(\psi_n, \varphi_n) = \inf \left\{ \log \varphi_n(q) : q \text{ is a projection in } \bigotimes_1^n \mathscr{A} \text{ with } \psi_n(q) \geqq 1 - \varepsilon \right\} \tag{1.6}$$

and

$$S_{\mathrm{pr}}(\psi_n, \varphi_n) = \sup \left\{ \psi_n(q) \log \frac{\psi_n(q)}{\varphi_n(q)} + (1 - \psi_n(q)) \log \frac{1 - \psi_n(q)}{1 - \varphi_n(q)} : q \text{ is a projection in } \bigotimes_1^n \mathscr{A} \right\}. \tag{1.7}$$

The quantity $\beta_\varepsilon(\psi_n, \varphi_n)$ has a natural meaning from the viewpoint of quantum hypothesis testing (cf. [4, 7, 13]). More precisely, let us suppose two hypotheses H_0 and H_1 so that the system $\mathscr{A}_\infty$ has states ψ and φ_∞ under H_0 and H_1, respectively. A projection q in $\bigotimes_1^n \mathscr{A}$ means a "quantum question" of size n, whose outcomes are the eigenvalues 1 or 0. We decide that H_0 (respectively H_1) is true if the outcome of q is 1 (respectively 0). Then $\varphi_n(q)$ (respectively $\psi_n(1-q)$) gives the "probability" of the error of accepting H_0 (respectively H_1) when H_1 (respectively H_0) actually is true. In this way we can consider the quantity $\exp\{\beta_\varepsilon(\psi_n, \varphi_n)\}$ as the bound of the first error probability over all decision rules of size n such that the second error probability does not exceed ε.

Now assume that ψ is completely ergodic in the sense that it is ergodic for any power of the right shift. (This is the case when ψ is weakly mixing.) Then our second theorem says that $\limsup_{n\to\infty} \frac{1}{n} \beta_\varepsilon(\psi_n, \varphi_n) \leqq -S_{\mathrm{M}}(\psi, \varphi_\infty)$ and $\liminf_{n\to\infty} \frac{1}{n} \beta_\varepsilon(\psi_n, \varphi_n) \geqq -\frac{1}{1-\varepsilon} S_{\mathrm{M}}(\psi, \varphi_\infty)$. Thus we can relate the mean relative entropy to a certain kind of asymptotic error bound in the quantum hypothesis testing. It could be mentioned that a desire for the visualization of noncommutative relative entropy as the logarithm of certain probabilities was formulated in [8] in connection with an interpretation of quantum theory. We finally establish that $\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{pr}}(\psi_n, \varphi_n) = S_{\mathrm{M}}(\psi, \varphi_\infty)$.

The proofs of these theorems are given in Sect. 3. The first theorem can be proved by a direct combinatorial computation. The proof of the second theorem is based on the Shannon-McMillan-Breiman theorem and the mean ergodic theorem together with the first theorem.

In Sect. 4 we note that our theorems hold in AF C^*-algebras or hyperfinite von Neumann algebras as well. Furthermore we show that if ψ is a tracial ergodic state of $\mathscr{A}_\infty$, then $\lim_{n\to\infty} \frac{1}{n} \beta_\varepsilon(\psi_n, \varphi_n) = -S_{\mathrm{M}}(\psi, \varphi_\infty)$ holds for every $0 < \varepsilon < 1$.

2. Main Results in Finite Dimensional C^*-Algebras

In this section we state our main theorems in the setting of finite dimensional C^*-algebras. Their proofs will be presented in Sect. 3. Although the theorems hold true in the framework of AF C^*-algebras or hyperfinite von Neumann algebras as will be noted in Sect. 4, these extensions are very easy and our essential ideas consist in the finite dimensional case; so we restrict our detailed discussions to this case.

First let us fix the notations. Let $\mathscr{A}$ be a finite dimensional C^*-algebra. Then $\mathscr{A}$ is identified with $\bigoplus_{l=1}^{L} M_{d_l}(\mathbf{C})$ which is the direct sum of $d_l \times d_l$ matrix algebras $M_{d_l}(\mathbf{C})$, $1 \leqq l \leqq L$, canonically represented on the Hilbert space $\bigoplus_{l=1}^{L} \mathbf{C}^{d_l}$. We denote by $\mathscr{A}_n$ the n-fold C^*-tensor product $\bigotimes_{1}^{n} \mathscr{A}$ for $n \geqq 1$, and by $\mathscr{A}_\infty$ the two-sided infinite C^*-tensor product $\bigotimes_{-\infty}^{\infty} \mathscr{A}$. Then $\{\mathscr{A}_n\}$ is considered as an increasing sequence of (finite dimensional) C^*-subalgebras of $\mathscr{A}_\infty$ by the natural inclusions. Let γ denote the right shift automorphism of $\mathscr{A}_\infty$. A state ψ of $\mathscr{A}_\infty$ is said to be stationary if ψ is γ-invariant (i.e. $\psi \circ \gamma = \psi$). For a state φ of $\mathscr{A}$ we denote by φ_∞ the infinite product state $\bigotimes_{-\infty}^{\infty} \varphi$ of $\mathscr{A}_\infty$, and let $\varphi_n = \varphi_\infty \Big| \mathscr{A}_n \left(= \bigotimes_{1}^{n} \varphi \right)$. We fix a state φ of $\mathscr{A}$ in the following discussions.

Let ψ be a stationary state of $\mathscr{A}_\infty$ and $\psi_n = \psi | \mathscr{A}_n$, $n \geqq 1$. The relative entropies $S(\psi_n, \varphi_n)$ are defined as (1.2). Then in view of the superadditivity of relative entropy [23] and the stationarity of ψ we get

$$S(\psi_{m+n}, \varphi_{m+n}) \geqq S(\psi_m, \varphi_m) + S(\psi_n, \varphi_n), \qquad m, n \geqq 1,$$

so that $\lim_{n\to\infty} \frac{1}{n} S(\psi_n, \varphi_n)$ exists; in fact (cf. [10, p. 274])

$$\lim_{n\to\infty} \frac{1}{n} S(\psi_n, \varphi_n) = \sup_{n \geqq 1} \frac{1}{n} S(\psi_n, \varphi_n). \tag{2.1}$$

We denote this limit by $S_{\mathrm{M}}(\psi, \varphi_\infty)$, which is called the mean relative entropy of ψ with respect to φ_∞. Note that if ω is a state of $\mathscr{A}$ then $S_{\mathrm{M}}(\omega_\infty, \varphi_\infty) = S(\omega, \varphi)$ because $S(\omega_n, \varphi_n) = nS(\omega, \varphi)$, $n \geqq 1$. In Sect. 1 besides Umegaki's relative entropy we referred to some other entropy quantities (1.3)–(1.5) and (1.7). The quantities $S_{\mathrm{co}}(\psi_n, \varphi_n)$ as in (1.3) are equivalently defined by

$$S_{\mathrm{co}}(\psi_n, \varphi_n) = \sup \{ S(\psi_n | \mathscr{B}, \varphi_n | \mathscr{B}) : \mathscr{B} \text{ is a commutative } C^*\text{-subalgebra of } \mathscr{A}_n \}.$$

The next theorem shows that the asymptotic limit of $S_{\mathrm{co}}(\psi_n, \varphi_n)$ exists and coincides with the mean relative entropy.

Theorem 2.1. *For every stationary state ψ of $\mathscr{A}_\infty$,*

$$\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{co}}(\psi_n, \varphi_n) = S_{\mathrm{M}}(\psi, \varphi_\infty). \tag{2.2}$$

A stationary state ψ of $\mathscr{A}$ is said to be ergodic if it is extremal in the set of stationary states. For ergodicity in general C^*-dynamical systems, see [10, 30] for instance. We say that ψ is weakly mixing if for every $a, b \in \mathscr{A}_\infty$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\psi(\gamma^i(a)b) - \psi(a)\psi(b)| = 0.$$

Obviously this is the case when ψ is strongly mixing (or strongly clustering), that is, $\lim_{n\to\infty} \psi(\gamma^n(a)b) = \psi(a)\psi(b)$ for every $a, b \in \mathscr{A}_\infty$. As in classical ergodic theory it is known [10] that if ψ is weakly mixing then it is ergodic. Note that the product state ω_∞ defined by a state ω of $\mathscr{A}$ is strongly mixing. In the following we say that ψ is completely ergodic if it is ergodic for all γ^n, $n \geqq 1$. We know that a weakly mixing state ψ is completely ergodic because it is weakly mixing for all γ^n.

Let $\beta_\varepsilon(\psi_n, \varphi_n)$ be defined by (1.6) for $n \geqq 1$ and $0 < \varepsilon < 1$. The next theorem shows that we have for large n

$$\exp\left\{\frac{1}{n}\beta_\varepsilon(\psi_n, \varphi_n)\right\} \approx \exp\{-S_{\mathrm{M}}(\psi, \varphi_\infty)\}$$

when ε is sufficiently small and ψ is completely ergodic. Thus $\exp\{-S_{\mathrm{M}}(\psi, \varphi_\infty)\}$ can be considered as the asymptotic error bound in the quantum hypothesis test for $\{\psi, \varphi_\infty\}$.

Theorem 2.2. *If ψ is a completely ergodic state of $\mathscr{A}_\infty$, then for every $0 < \varepsilon < 1$,*

$$\limsup_{n\to\infty} \frac{1}{n} \beta_\varepsilon(\psi_n, \varphi_n) \leqq -S_{\mathrm{M}}(\psi, \varphi_\infty), \tag{2.3}$$

$$\liminf_{n\to\infty} \frac{1}{n} \beta_\varepsilon(\psi_n, \varphi_n) \geqq -\frac{1}{1-\varepsilon} S_{\mathrm{M}}(\psi, \varphi_\infty). \tag{2.4}$$

It may be possible that $\lim_{n\to\infty} \frac{1}{n} \beta_\varepsilon(\psi_n, \varphi_n) = -S_{\mathrm{M}}(\psi, \varphi_\infty)$ for every $0 < \varepsilon < 1$ and every ergodic state ψ. In particular when ψ is a tracial ergodic state, this will be shown in Sect. 4.

The quantities $S_{\mathrm{pr}}(\psi_n, \varphi_n)$ in (1.7) are also defined by

$$S_{\mathrm{pr}}(\psi_n, \varphi_n) = \sup\{S(\psi_n|\mathscr{B}, \varphi_n|\mathscr{B}) : \mathscr{B} \text{ is a two dimensional subalgebra of } \mathscr{A}_n\}.$$

Note that

$$S_{\mathrm{pr}}(\psi_n, \varphi_n) \leqq S_{\mathrm{co}}(\psi_n, \varphi_n) \leqq S_{\mathrm{cp}}(\psi_n, \varphi_n) \leqq S(\psi_n, \varphi_n) \tag{2.5}$$

by the monotonicity of relative entropy as mentioned in Sect. 1.

As for completely ergodic states we can make Theorem 2.1 extremely sharp as follows. Indeed the method in proving (2.4) will work for Theorem 2.3 as well.

Theorem 2.3. *If ψ is a completely ergodic state of $\mathscr{A}_\infty$, then*

$$\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{pr}}(\psi_n, \varphi_n) = S_{\mathrm{M}}(\psi, \varphi_\infty).$$

As a special case we have for every state ω of $\mathscr{A}$,

$$\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{pr}}(\omega_n, \varphi_n) = \lim_{n\to\infty} \frac{1}{n} S_{\mathrm{co}}(\omega_n, \varphi_n) = \lim_{n\to\infty} \frac{1}{n} S_{\mathrm{cp}}(\omega_n, \varphi_n) = S(\omega, \varphi). \tag{2.6}$$

This means that Umegaki's relative entropy comes out when we first adopt any of the quantities S_{pr}, S_{co} or S_{cp} and then take asymptotics.

The following examples show that the ergodicity assumption of ψ is essential in Theorems 2.2 and 2.3. We are indebted to the referee for the first example.

Examples 2.4. (1) Let $0<\varepsilon<1/2$ and ψ be a stationary state of $\mathscr{A}_\infty$ with $\psi \neq \varphi_\infty$. Then $S_{\mathrm{M}}(2\varepsilon\varphi_\infty + (1-2\varepsilon)\psi, \varphi_\infty) > 0$ because otherwise by (2.1) we get $2\varepsilon\varphi_\infty + (1-2\varepsilon)\psi = \varphi_\infty$ so that $\psi = \varphi_\infty$. On the other hand, since $2\varepsilon\varphi_n(q) + (1-2\varepsilon)\psi_n(q) \geqq 1-\varepsilon$ implies $\varphi_n(q) \geqq 1/2$ for a projection q in $\mathscr{A}_n$, we have

$$\beta_\varepsilon(2\varepsilon\varphi_n + (1-2\varepsilon)\psi_n, \varphi_n) \geqq -\log 2 .$$

Hence

$$\lim_{n\to\infty} \frac{1}{n} \beta_\varepsilon(2\varepsilon\varphi_n + (1-2\varepsilon)\psi_n, \varphi_n) = 0 .$$

(2) Let $\mathscr{A} = \mathbf{C}\oplus\mathbf{C}$ and ω^1, ω^2, φ be given with the densities $(1,0)$, $(0,1)$, $(\alpha, 1-\alpha)$, respectively, where $0<\alpha<1/2$. Let $\psi = \frac{1}{2}(\omega^1_\infty + \omega^2_\infty)$. By the affinity of $S_{\mathrm{M}}(\cdot\,, \varphi_\infty)$ (see [28]) we get

$$S_{\mathrm{M}}(\psi, \varphi_\infty) = \tfrac{1}{2}\{S(\omega^1, \varphi) + S(\omega^2, \varphi)\} = -\tfrac{1}{2}\{\log\alpha + \log(1-\alpha)\} .$$

But we easily see that $S_{\mathrm{pr}}(\psi_n, \varphi_n)$ is the maximum of

$$\log \frac{1}{\alpha^n + (1-\alpha)^n} = -n \log\{\alpha^n + (1-\alpha)^n\}^{1/n}$$

and

$$\frac{1}{2}\log\frac{1}{2\alpha^n} + \frac{1}{2}\log\frac{1}{2(1-\alpha^n)} = -\frac{n}{2}\log\alpha - \frac{1}{2}\log(1-\alpha^n) - \log 2 .$$

Hence

$$\lim_{n\to\infty} \frac{1}{n} S_{\mathrm{pr}}(\psi_n, \varphi_n) = \max\left\{-\log(1-\alpha), -\frac{1}{2}\log\alpha\right\} < S_{\mathrm{M}}(\psi, \varphi_\infty) .$$

In the rest of this section, using Theorem 2.1 let us establish the relation between the relative entropy $S(\omega, \varphi)$ and the entropy quantity $S_{\mathrm{BS}}(\omega, \varphi)$ in (1.5). The expression (1.5) implicitly means that $S_{\mathrm{BS}}(\omega, \varphi) = +\infty$ if the support projection of ω is not dominated by that of φ. The main properties of S_{BS} follow from [12] devoted to an operator-valued relative entropy. We here state the additivity and the monotonicity of S_{BS}.

Proposition 2.5. (1) *$S_{\mathrm{BS}}(\omega_1\otimes\omega_2, \varphi_1\otimes\varphi_2) = S_{\mathrm{BS}}(\omega_1, \varphi_1) + S_{\mathrm{BS}}(\omega_2, \varphi_2)$ when ω_i and φ_i are states of (finite dimensional) C^*-algebras $\mathscr{A}_i$, $i=1,2$.*
(2) *$S_{\mathrm{BS}}(\omega|\mathscr{B}, \varphi|\mathscr{B}) \leqq S_{\mathrm{BS}}(\omega, \varphi)$ for any C^*-subalgebra $\mathscr{B}$ of $\mathscr{A}$.*

Indeed (1) is obvious from the definition. Although (2) follows from the operator-valued version of [12], we briefly recall the proof for the convenience of the reader. We may assume that φ is faithful. Since $X(\log X^*X) = (\log XX^*)X$ holds for a matrix X, it follows that

$$S_{\mathrm{BS}}(\omega, \varphi) = -\mathrm{Tr}(D_\varphi \eta(D_\varphi^{-1/2} D_\omega D_\varphi^{-1/2})) ,$$

where $\eta(t)=-t\log t$, $t\geqq 0$. By the operator-concavity of η we get

$$\alpha(\eta(D_\varphi^{-1/2}D_\omega D_\varphi^{-1/2}))\leqq\eta(\alpha(D_\varphi^{-1/2}D_\omega D_\varphi^{-1/2})),$$

where $\alpha(X)=E(D_\varphi)^{-1/2}E(D_\varphi^{1/2}XD_\varphi^{1/2})E(D_\varphi)^{-1/2}$ with the conditional expectation E from $\mathscr{A}$ onto $\mathscr{B}$ with respect to Tr. Hence

$$E(D_\varphi^{1/2}\eta(D_\varphi^{-1/2}D_\omega D_\varphi^{-1/2})D_\varphi^{1/2})\leqq E(D_\varphi)^{1/2}\eta(E(D_\varphi)^{-1/2}E(D_\omega)E(D_\varphi)^{-1/2})E(D_\varphi)^{1/2}.$$

Taking the trace of both sides proves (2).

Now Theorem 2.1 together with Proposition 2.5 has an interesting consequence as follows.

Corollary 2.6. *$S_{\mathrm{BS}}(\omega,\varphi)\geqq S(\omega,\varphi)$ for all states ω and φ of $\mathscr{A}$.*

Proof. Since $S_{\mathrm{BS}}=S$ for commuting states, we have $S_{\mathrm{BS}}(\omega,\varphi)\geqq S_{\mathrm{co}}(\omega,\varphi)$ by (2) of the above proposition. In particular, for every $n\geqq 1$,

$$S_{\mathrm{BS}}(\omega_n,\varphi_n)\geqq S_{\mathrm{co}}(\omega_n,\varphi_n),$$

so that by the above (1),

$$S_{\mathrm{BS}}(\omega,\varphi)\geqq\frac{1}{n}S_{\mathrm{co}}(\omega_n,\varphi_n).$$

Letting $n\to\infty$ we infer the corollary due to Theorem 2.1. □

It is quite remarkable that the corollary is equivalent to the trace inequality

$$\operatorname{Tr}A\log A^{1/2}BA^{1/2}\geqq\operatorname{Tr}A(\log A+\log B) \tag{2.7}$$

for nonnegative matrices A and B. In fact, (2.7) is immediate from the corollary when B is invertible. Take the limit from $B+\varepsilon I$, $\varepsilon>0$, for general nonnegative B. When A and B are positive invertible, one can define $\log AB$ by analytical functional calculus or by power series and get the equality

$$\operatorname{Tr}A\log A^{1/2}BA^{1/2}=\operatorname{Tr}A\log AB$$

because $\operatorname{Tr}A(A^{1/2}BA^{1/2})^n=\operatorname{Tr}A(AB)^n$ for $n\geqq 1$.

3. Proofs of Theorems

In this section we present the proofs of Theorems 2.1–2.3. Let us keep the notations fixed in the previous section. Let Tr denote the canonical trace of $\mathscr{A}$ such that $\operatorname{Tr}(e)=1$ for all one dimensional projections e in $\mathscr{A}$. Let D_φ be the density of φ with respect to Tr, and K be the sum of the sizes of simple summands of $\mathscr{A}$ $\left(\text{i.e. } K=\sum_{l=1}^{L}d_l\right)$. Taking the spectral decomposition of D_φ, one can write $D_\varphi=\sum_{k=1}^{K}\lambda_k e_k$, where e_k are one dimensional projections. Let n be an arbitrary fixed positive integer. For each K-tuple $(n_1,n_2,\ldots,n_K)$ of nonnegative integers with $\sum_{k=1}^{K}n_k=n$, we denote by $I_{n_1,\ldots,n_K}$ the set of all $(i_1,i_2,\ldots,i_n)$ such that $\#\{j:i_j=k\}=n_k$ for $1\leqq k\leqq K$, and define the projection $p_{n_1,\ldots,n_K}$ in $\mathscr{A}_n$ by

$$p_{n_1,\ldots,n_K}=\sum_{(i_1,\ldots,i_n)\in I_{n_1,\ldots,n_K}}e_{i_1}\otimes e_{i_2}\otimes\ldots\otimes e_{i_n}.$$

Then $\sum_{n_1,\ldots,n_K} p_{n_1,\ldots,n_K}=1$ and the density D_{φ_n} of φ_n with respect to $\mathrm{Tr}_n=\bigotimes_1^n \mathrm{Tr}$ is given by

$$D_{\varphi_n}=\bigotimes_1^n D_\varphi=\sum_{n_1,\ldots,n_K}\left(\prod_{k=1}^K \lambda_k^{n_k}\right)p_{n_1,\ldots,n_K}. \tag{3.1}$$

Let E_n denote the conditional expectation from $\mathscr{A}_n$ onto $\bigoplus_{n_1,\ldots,n_K} p_{n_1,\ldots,n_K}\mathscr{A}_n p_{n_1,\ldots,n_K}$ with respect to Tr_n, which is given by

$$E_n(x)=\sum_{n_1,\ldots,n_K} p_{n_1,\ldots,n_K} x p_{n_1,\ldots,n_K}, \qquad x\in\mathscr{A}_n. \tag{3.2}$$

The (von Neumann) entropy $S(\omega)$ of a state ω of $\mathscr{A}_n$ is defined by $S(\omega)=-\mathrm{Tr}_n(D_\omega \log D_\omega)$, where D_ω is the density of ω with respect to Tr_n.

Lemma 3.1. *If ω is a state of $\mathscr{A}_n$ and $\mathscr{B}$ is the commutative subalgebra of $\mathscr{A}_n$ generated by $\{p_{n_1,\ldots,n_K}D_\omega p_{n_1,\ldots,n_K}\}_{n_1,\ldots,n_K}\cup\{p_{n_1,\ldots,n_K}\}_{n_1,\ldots,n_K}$, then*

$$S(\omega,\varphi_n)=S(\omega|\mathscr{B},\varphi_n|\mathscr{B})+S(\omega\circ E_n)-S(\omega).$$

Proof. Let $s(\omega)$ and $s(\varphi_n)$ denote the support projections of ω and φ_n, respectively. Since $D_{\varphi_n}\in\mathscr{B}$ by (3.1), we have $s(\varphi_n)=s(\varphi_n|\mathscr{B})\in\mathscr{B}$. If $s(\omega)\leqq s(\varphi_n)$ is not satisfied, then the desired equality holds because $S(\omega,\varphi_n)=S(\omega|\mathscr{B},\varphi_n|\mathscr{B})=+\infty$. So suppose $s(\omega)\leqq s(\varphi_n)$. In this case we may assume that φ_n is faithful; otherwise consider the restrictions of ω and φ_n to $s(\varphi_n)\mathscr{A}_n s(\varphi_n)$ and $\mathscr{B}s(\varphi_n)$. Since $\mathscr{B}$ is included in the centralizer of φ_n, the conditional expectation $E_\mathscr{B}$ from $\mathscr{A}_n$ onto $\mathscr{B}$ with respect to φ_n exists due to [31]. Hence we get by [14, 22, 23]

$$S(\omega,\varphi_n)=S(\omega|\mathscr{B},\varphi_n|\mathscr{B})+S(\omega,\omega\circ E_\mathscr{B}). \tag{3.3}$$

Since $E_n(D_\omega)\in\mathscr{B}$ and $\mathscr{B}\subset E_n(\mathscr{A}_n)$ by (3.2), we get for every $a\in\mathscr{A}_n$,

$$\begin{aligned}(\omega\circ E_\mathscr{B})(a)&=\mathrm{Tr}_n(D_\omega E_\mathscr{B}(a))=\mathrm{Tr}_n(E_n(D_\omega)E_\mathscr{B}(a))\\&=\varphi_n(D_{\varphi_n}^{-1}E_n(D_\omega)E_\mathscr{B}(a))=\varphi_n(E_\mathscr{B}(D_{\varphi_n}^{-1}E_n(D_\omega)a))\\&=\varphi_n(D_{\varphi_n}^{-1}E_n(D_\omega)a)=\mathrm{Tr}_n(E_n(D_\omega)a),\end{aligned}$$

so that $E_n(D_\omega)$ is the density of $\omega\circ E_\mathscr{B}$ as well as $\omega\circ E_n$ with respect to Tr_n. Therefore $\omega\circ E_\mathscr{B}=\omega\circ E_n$ and

$$S(\omega,\omega\circ E_n)=\mathrm{Tr}_n D_\omega(\log D_\omega-\log E_n(D_\omega))=S(\omega\circ E_n)-S(\omega). \tag{3.4}$$

The desired equality follows from (3.3) and (3.4). □

Lemma 3.2. *For every state ω of $\mathscr{A}_n$,*

$$S(\omega\circ E_n)-S(\omega)\leqq K\log(n+1).$$

Proof. In view of the joint convexity of the relative entropy [2, 16] it suffices by (3.4) to show the case when ω is a pure state. In this case D_ω is a one dimensional projection and $S(\omega)=0$. Since each $p_{n_1,\ldots,n_K}D_\omega p_{n_1,\ldots,n_K}$ is of rank one or zero, it follows that the rank of $E_n(D_\omega)$ is at most

$$\#\left\{(n_1,\ldots,n_K): n_1,\ldots,n_K\geqq 0,\ \sum_{k=1}^K n_k=n\right\}\ (\leqq(n+1)^K).$$

Therefore

$$S(\omega \circ E_n) \leqq \log(n+1)^K = K \log(n+1),$$

as desired. □

Proof of Theorem 2.1. For every n we have

$$S_{\mathrm{co}}(\psi_n, \varphi_n) \leqq S(\psi_n, \varphi_n) \leqq S_{\mathrm{co}}(\psi_n, \varphi_n) + K \log(n+1)$$

by (2.5) and by Lemmas 3.1 and 3.2 applied to ψ_n. This proves (2.2). □

In the sequel of this section we assume that ψ is a completely ergodic state of $\mathscr{A}_\infty$.

Proof of (2.3) of Theorem 2.2. For any $r > -S_{\mathrm{M}}(\psi, \varphi_\infty)$ let us choose $h < S_{\mathrm{M}}(\psi, \varphi_\infty)$ and $\delta > 0$ with $-h + \delta < r$. By Theorem 2.1 there exists a commutative C^*-subalgebra $\mathscr{B}$ of $\mathscr{A}_l$ for some $l \geqq 1$ such that

$$S(\psi_l|\mathscr{B}, \varphi_l|\mathscr{B}) \geqq lh. \tag{3.5}$$

We can consider $\mathscr{B}_k$ as a C^*-subalgebra of $\mathscr{A}_{kl}$, $k \geqq 1$, and $\mathscr{B}_\infty$ as a C^*-subalgebra of $\mathscr{A}_\infty$. Let $\sigma = \gamma^l|\mathscr{B}_\infty$, the right shift on $\mathscr{B}_\infty$. Define $\mu = \varphi_l|\mathscr{B}$, $\nu = \psi|\mathscr{B}_\infty$, $\mu_\infty = \bigotimes_{-\infty}^{\infty} \mu$, $\mu_k = \mu_\infty|\mathscr{B}_k$ and $\nu_k = \nu|\mathscr{B}_k$. These states may be identified with the probability measures on the corresponding underlying spaces. Since ψ is completely ergodic, we can readily see that ν is ergodic for σ. In the following we work in the (commutative) von Neumann algebra $\pi_\nu(\mathscr{B}_\infty)''$ where π_ν is the GNS representation of $\mathscr{B}_\infty$ associated with ν. We denote the normal extensions of ν and σ to $\pi_\nu(\mathscr{B}_\infty)''$ by the same ν and σ.

First suppose $\nu_1 \ll \mu$ (i.e. $s(\nu_1) \leqq s(\mu)$) is not satisfied. Then there is a projection p in $\mathscr{B}$ such that $\mu(p) = 0$ and $\nu_1(p) > 0$. Since the ergodicity of ν implies that $\pi_\nu\left(\bigvee_{i=0}^{k-1} \sigma^i(p)\right) \to 1$ strongly as $k \to \infty$, there exists k_0 such that $\nu\left(\bigvee_{i=0}^{k_0-1} \sigma^i(p)\right) \geqq 1 - \varepsilon$. Set $q = \bigvee_{i=0}^{k_0-1} \sigma^i(p)$ and $n_0 = k_0 l$. Then q is a projection in $\mathscr{A}_{n_0}$ such that $\psi(q) \geqq 1 - \varepsilon$ and

$$\varphi_{n_0}(q) = \mu_{k_0}\left(\bigvee_{i=0}^{k_0-1} \sigma^i(p)\right) \leqq k_0 \mu(p) = 0.$$

Therefore $\beta_\varepsilon(\psi_n, \varphi_n) = -\infty$ for all $n \geqq n_0$, proving (2.3).

Next suppose $\nu_1 \ll \mu$. Then we see that $\nu_k \ll \mu_k$ for every k because $s(\nu_k) \leqq \bigotimes_1^k s(\nu_1) \leqq \bigotimes_1^k s(\mu) = s(\mu_k)$. Let m denote the trace of $\mathscr{B}$ such that $m(e) = 1$ for all atoms e in $\mathscr{B}$ (i.e. m is the counting measure on the underlying space of $\mathscr{B}$). Let us consider the selfadjoint operators H_k, $k \geqq 1$, in $\pi_\nu(\mathscr{B}_\infty)''$ given by

$$\begin{aligned} H_k &= \frac{1}{k} \left\{ \pi_\nu\left(\log \frac{d\nu_k}{dm_k}\right) - \pi_\nu\left(\log \frac{d\mu_k}{dm_k}\right) \right\} \\ &= \frac{1}{k} \pi_\nu\left(\log \frac{d\nu_k}{dm_k}\right) - \frac{1}{k} \sum_{i=0}^{k-1} \sigma^i\left(\pi_\nu\left(\log \frac{d\mu}{dm}\right)\right). \end{aligned} \tag{3.6}$$

By the Shannon-McMillan-Breiman theorem (cf. [3, 21]) the first term of (3.6) converges ν-almost uniformly as $k\to\infty$ to the Kolmogorov-Sinai entropy $h_\nu(\sigma)$ of σ relative to ν. On the other hand, by the mean ergodic theorem the second term of (3.6) converges strongly as $k\to\infty$ to $\nu_1(\log d\mu/dm)$. Thus we see that H_k converges in ν-measure to

$$\begin{aligned} h_0 &= \lim_{k\to\infty} \frac{1}{k} \nu_k \left(\log \frac{d\nu_k}{dm_k} - \log \frac{d\mu_k}{dm_k} \right) \\ &= \lim_{k\to\infty} \frac{1}{k} S(\nu_k, \mu_k) = S_M(\nu, \mu_\infty) . \end{aligned} \tag{3.7}$$

Now for each k let p_k be the projection in $\mathscr{B}_k$ with $p_k \leqq s(\nu_k)$ such that $\pi_\nu(p_k)$ is the spectral projection of H_k corresponding to the interval $(h_0-\delta, h_0+\delta)$. Then there exists k_0 such that $\nu(p_k)\geqq 1-\varepsilon$ for all $k\geqq k_0$. Since $\pi_\nu(p_k)$ is the spectral projection of $\exp(kH_k)$ corresponding to $(e^{k(h_0-\delta)}, e^{k(h_0+\delta)})$, we get

$$\begin{aligned} \pi_\nu(p_k) &\leqq e^{-k(h_0-\delta)}\pi_\nu(p_k)\exp(kH_k) \\ &= e^{-k(h_0-\delta)}\pi_\nu \left(p_k \left(\frac{d\nu_k}{dm_k}\right) \left(\frac{d\mu_k}{dm_k}\right)^{-1} \right), \end{aligned}$$

so that since $p_k \leqq s(\nu_k)$

$$p_k \leqq e^{-k(h_0-\delta)} p_k \left(\frac{d\nu_k}{dm_k}\right) \left(\frac{d\mu_k}{dm_k}\right)^{-1} . \tag{3.8}$$

Therefore

$$\begin{aligned} \mu_k(p_k) &\leqq e^{-k(h_0-\delta)} m_k \left(p_k \frac{d\nu_k}{dm_k} \right) = e^{-k(h_0-\delta)} \nu_k(p_k) \\ &\leqq e^{-k(h_0-\delta)} \leqq e^{-k(lh-\delta)}, \end{aligned} \tag{3.9}$$

because $h_0 \geqq S(\nu_1, \mu) \geqq lh$ by (3.7), (2.1) and (3.5). For each $n \geqq k_0 l$ let $n=kl+j$, where $k\geqq k_0$, $0\leqq j<l$, and put $q_n=p_k$. Then we have $q_n \in \mathscr{A}_n$ and $\psi(q_n)\geqq 1-\varepsilon$, so that

$$\frac{1}{n}\beta_\varepsilon(\psi_n, \varphi_n) \leqq \frac{1}{n}\log\varphi_n(q_n) \leqq \frac{1}{(k+1)l}\log\mu_k(p_k) \leqq -\frac{k}{k+1}h+\delta$$

by (3.9). This proves (2.3) thanks to $-h+\delta<r$. □

Proof of (2.4) of Theorem 2.2. First suppose $S_M(\psi, \varphi_\infty)=0$. Then $\psi=\varphi_\infty$ by (2.1) and hence $\beta_\varepsilon(\psi_n, \varphi_n)\geqq \log(1-\varepsilon)$, $n\geqq 1$, so that (2.4) is immediate. Now suppose $S_M(\psi, \varphi_\infty)>0$. For each n set $\beta_n=\beta_\varepsilon(\psi_n, \varphi_n)$ and choose a projection q_n in $\mathscr{A}_n$ such that $\psi_n(q_n)\geqq 1-\varepsilon$ and $\varphi_n(q_n)<e^{\beta_n+1}$. Letting $0<h<S_M(\psi, \varphi_\infty)$, by (2.3) we have $\beta_n<-nh$ for sufficiently large n. Hence $e^{\beta_n+1}\to 0$ as $n\to\infty$. Define $\mathscr{B}_n=\mathbf{C}q_n+\mathbf{C}(1-q_n)$ which is a two dimensional subalgebra of $\mathscr{A}_n$. Then by monotonicity

$$S(\psi_n, \varphi_n) \geqq S(\psi_n|\mathscr{B}_n, \varphi_n|\mathscr{B}_n) = F(\psi_n(q_n), \varphi_n(q_n)),$$

where

$$F(s,t) = s\log\frac{s}{t} + (1-s)\log\frac{1-s}{1-t}, \qquad 0\leqq s, t\leqq 1 . \tag{3.10}$$

Since for $0<t<s\leqq 1$,

$$\frac{\partial F(s,t)}{\partial t}=\frac{t-s}{t(1-t)}<0,$$

we have for every n large enough

$$\begin{aligned} S(\psi_n,\varphi_n) &\geqq F(\psi_n(q_n), e^{\beta_n+1}) \\ &= -\psi_n(q_n)(\beta_n+1)-(1-\psi_n(q_n))\log(1-e^{\beta_n+1}) \\ &\quad +\psi_n(q_n)\log\psi_n(q_n)+(1-\psi_n(q_n))\log(1-\psi_n(q_n)) \\ &\geqq -(1-\varepsilon)(\beta_n+1)-\log 2. \end{aligned}$$

Therefore

$$(1-\varepsilon)\liminf_{n\to\infty}\frac{1}{n}\beta_n\geqq -S_{\mathrm{M}}(\psi,\varphi_\infty),$$

as desired. □

Proof of Theorem 2.3. It suffices by monotonicity to show that

$$\liminf_{n\to\infty}\frac{1}{n}S_{\mathrm{pr}}(\psi_n,\varphi_n)\geqq S_{\mathrm{M}}(\psi,\varphi_\infty). \tag{3.11}$$

We may suppose $S_{\mathrm{M}}(\psi,\varphi_\infty)>0$. Let $0<\varepsilon<1$ and $0<h<S_{\mathrm{M}}(\psi,\varphi_\infty)$. By (2.3) we have $\beta_\varepsilon(\psi_n,\varphi_n)<-nh$ for sufficiently large n. Hence for each such n we can choose a projection q_n in $\mathscr{A}_n$ such that $\psi_n(q_n)\geqq 1-\varepsilon$ and $\varphi_n(q_n)<e^{-nh}$. Let $\mathscr{B}_n=\mathbf{C}q_n+\mathbf{C}(1-q_n)$ and F be the function in (3.10). Then we have as in the proof of (2.4),

$$\begin{aligned} S_{\mathrm{pr}}(\psi_n,\varphi_n) &\geqq S(\psi_n|\mathscr{B}_n,\varphi_n|\mathscr{B}_n)\geqq F(\psi_n(q_n),\varphi_n(q_n)) \\ &\geqq F(\psi_n(q_n),e^{-nh})\geqq n(1-\varepsilon)h-\log 2 \end{aligned}$$

for every n large enough. Therefore

$$\liminf_{n\to\infty}\frac{1}{n}S_{\mathrm{pr}}(\psi_n,\varphi_n)\geqq(1-\varepsilon)h.$$

We obtain (3.11) letting $\varepsilon\to 0$ and $h\to S_{\mathrm{M}}(\psi,\varphi_\infty)$. □

Indeed it is enough in Theorems 2.2 and 2.3 to assume that ψ is a stationary state of $\mathscr{A}_\infty$ which is ergodic for γ^n for infinitely many n.

4. Extensions

In this section we observe that our theorems in Sect. 2 remain true in AF C^*-algebras or hyperfinite von Neumann algebras. When $\mathscr{A}$ is a general C^*-algebra (always assumed to be unital), given two states ω and φ of $\mathscr{A}$ one can define the relative entropy $S(\omega,\varphi)$ of ω with respect to φ as Uhlmann's relative entropy [32]. But this is also defined through Araki's one [1, 2] for normal states of von Neumann algebras as follows: If π is a representation of $\mathscr{A}$ such that ω and φ have the respective normal extensions $\bar{\omega}$ and $\bar{\varphi}$ to $\pi(\mathscr{A})''$ with $\bar{\omega}\circ\pi=\omega$ and $\bar{\varphi}\circ\pi=\varphi$, then we have $S(\omega,\varphi)=S(\bar{\omega},\bar{\varphi})$ (see [15, 26]).

Let $\mathscr{A}$ be a C^*-algebra with a fixed state φ. As in Sect. 2 we take the C^*-tensor products $\mathscr{A}_\infty = \bigotimes_{-\infty}^{\infty} \mathscr{A}$, $\mathscr{A}_n = \bigotimes_{1}^{n} \mathscr{A}$, $n \geqq 1$, the right shift automorphism of $\mathscr{A}_\infty$ and the product state φ_∞ of $\mathscr{A}_\infty$. Moreover for a stationary state ψ of $\mathscr{A}_\infty$ the mean relative entropy $S_{\mathrm{M}}(\psi, \varphi_\infty)$ is defined as (2.1). To extend the theorems to the case of an AF C^*-algebra, we first give the following lemma.

Lemma 4.1. *Let $\{\mathscr{A}(j)\}$ be an increasing net of C^*-subalgebras of $\mathscr{A}$ such that $\mathscr{A} = \overline{\bigcup_j \mathscr{A}(j)}$. Let ψ be a stationary state of $\mathscr{A}_\infty$. If $\varphi(j) = \varphi|\mathscr{A}(j)$ and $\psi(j) = \psi|\mathscr{A}(j)_\infty$, then $S_{\mathrm{M}}(\psi(j), \varphi(j)_\infty)$ increases to $S_{\mathrm{M}}(\psi, \varphi_\infty)$.*

Proof. For each n, since $\mathscr{A}_n = \overline{\bigcup_j \mathscr{A}(j)_n}$, we have

$$\sup_j S(\psi_n|\mathscr{A}(j)_n, \varphi_n|\mathscr{A}(j)_n) = S(\psi_n, \varphi_n)$$

by the martingale convergence of relative entropy [2, 16] applied under some representation of $\mathscr{A}_n$. Hence

$$\begin{aligned} S_{\mathrm{M}}(\psi, \varphi_\infty) &= \sup_{n \geqq 1} \frac{1}{n} S(\psi_n, \varphi_n) \\ &= \sup_j \sup_{n \geqq 1} \frac{1}{n} S(\psi_n|\mathscr{A}(j)_n, \varphi_n|\mathscr{A}(j)_n) \\ &= \sup_j S_{\mathrm{M}}(\psi(j), \varphi(j)_\infty). \end{aligned}$$

The increasingness is immediate from the monotonicity of relative entropy. □

Now assume that $\mathscr{A}$ is an AF C^*-algebra. Then there is an increasing net $\{\mathscr{A}(j)\}$ of finite dimensional subalgebras of $\mathscr{A}$ such that $\mathscr{A} = \overline{\bigcup_j \mathscr{A}(j)}$. Let ψ be a stationary state of $\mathscr{A}_\infty$ and define $S_{\mathrm{co}}(\psi_n, \varphi_n)$, $n \geqq 1$, as (1.3). For each j and n we have by Lemmas 3.1 and 3.2 applied to $\psi_n|\mathscr{A}(j)_n$,

$$S(\psi_n|\mathscr{A}(j)_n, \varphi_n|\mathscr{A}(j)_n) \leqq S_{\mathrm{co}}(\psi_n, \varphi_n) + K_j \log(n+1),$$

where K_j is the sum of the sizes of simple summands of $\mathscr{A}(j)$. Hence for every j,

$$S_{\mathrm{M}}(\psi(j), \varphi(j)_\infty) \leqq \liminf_{n \to \infty} \frac{1}{n} S_{\mathrm{co}}(\psi_n, \varphi_n),$$

so that Lemma 4.1 shows the equality (2.2). Furthermore by the above argument we can choose, given $h < S_{\mathrm{M}}(\psi, \varphi_\infty)$, a finite dimensional commutative subalgebra $\mathscr{B}$ of $\mathscr{A}_l$ for some $l \geqq 1$ such that (3.5) holds. Hence the proof of (2.3) of Theorem 2.2 works well. Thus we infer that Theorems 2.1–2.3 remain true for every stationary or completely ergodic state ψ of $\mathscr{A}_\infty$ when $\mathscr{A}$ is an AF C^*-algebra.

Our theorems can be formulated in the framework of von Neumann algebras too. Let $\mathscr{M}$ be a von Neumann algebra with a fixed normal state φ. Let $\mathscr{M}_n$ be the n-fold von Neumann tensor product $\overline{\bigotimes}_1^n \mathscr{M}$ for $n \geqq 1$, and $\mathscr{M}_\infty$ the C^*-completion of $\bigcup_{n=1}^{\infty} \left(\overline{\bigotimes}_{-n}^{n} \mathscr{M} \right)$. Then $\{\mathscr{M}_n\}$ is an increasing sequence of von Neumann algebras included in $\mathscr{M}_\infty$. We have the right shift and the product state φ_∞ of $\mathscr{M}_\infty$ as before.

If $\{\mathcal{M}(j)\}$ is an increasing net of von Neumann subalgebras of $\mathcal{M}$ such that $\mathcal{M}=\left(\bigcup_j \mathcal{M}(j)\right)''$ and if ψ is a stationary state of $\mathcal{M}_\infty$ such that $\psi_n=\psi|\mathcal{M}_n$ is normal for every n, then the same conclusion as Lemma 4.1 holds. Now assume that $\mathcal{M}$ is hyperfinite (i.e. approximately finite dimensional). Of course this is the case when $\mathcal{M}=\mathbf{B}(\mathcal{H})$, the algebra of all bounded operators on a Hilbert space $\mathcal{H}$. Then the same results as Theorems 2.1–2.3 hold for every stationary or completely ergodic state ψ of $\mathcal{M}_\infty$ such that ψ_n is normal for every n. In particular we have (2.6) for every normal state ω of $\mathcal{M}$.

Assume again that $\mathcal{A}$ is finite dimensional with a state φ. We finally consider the special setting where ψ is a *tracial* ergodic state of $\mathcal{A}_\infty$. In this case we denote by $\bar{\psi}$ and $\bar{\gamma}$ the respective normal extensions of ψ and γ to $\pi_\psi(\mathcal{A}_\infty)''$. Then $\bar{\psi}$ becomes a faithful normal tracial state of $\pi_\psi(\mathcal{A}_\infty)''$ with $\bar{\psi}\circ\bar{\gamma}=\bar{\psi}$. Let Tr be the canonical trace of $\mathcal{A}$ mentioned at the beginning of Sect. 3 and $\mathcal{Z}(\mathcal{A})$ be the center of $\mathcal{A}$. Although it is a challenging open problem to establish the noncommutative Shannon-McMillan-Breiman theorem for general ergodic states of $\mathcal{A}_\infty$, the special case of the following lemma was given in [20]. In fact, this follows from the classical case for $\psi\Big|\bigotimes_{-\infty}^{\infty}\mathcal{Z}(\mathcal{A})$ because $d\psi_n/d\mathrm{Tr}_n$ belongs to $\bigotimes_1^n\mathcal{Z}(\mathcal{A})$, the center of $\mathcal{A}_n$.

Lemma 4.2. *Let $\mathcal{A}$ and ψ be as above. Then $-\frac{1}{n}\pi_\psi(\log d\psi_n/d\mathrm{Tr}_n)$ converges $\bar{\psi}$-almost uniformly and in $L^1(\bar{\psi})$-norm to a constant h. (h is the dynamical entropy [6] of γ relative to ψ.)*

On the other hand the noncommutative mean ergodic theorem was given in [18]. Based on these convergence results we have the next proposition which is a stronger form of Theorem 2.2 and extends the classical results in [4, 7]. A similar result is given in [19] when φ is the trace Tr and ψ is a product state.

Proposition 4.3. *Let ψ be a tracial ergodic state of $\mathcal{A}_\infty$. Then for every $0<\varepsilon<1$,*

$$\lim_{n\to\infty}\frac{1}{n}\beta_\varepsilon(\psi_n,\varphi_n)=-S_{\mathrm{M}}(\psi,\varphi_\infty).$$

Proof. We shall show that

$$\liminf_{n\to\infty}\frac{1}{n}\beta_\varepsilon(\psi_n,\varphi_n)\geqq -S_{\mathrm{M}}(\psi,\varphi_\infty). \tag{4.1}$$

Our method in the following will allow the proof of (2.3) of Theorem 2.2 to be adapted to give

$$\limsup_{n\to\infty}\frac{1}{n}\beta_\varepsilon(\psi_n,\varphi_n)\leqq -S_{\mathrm{M}}(\psi,\varphi_\infty).$$

When $s(\psi_1)\leqq s(\varphi)$ is not satisfied, we can see arguing as in the proof of (2.3) that $\beta_\varepsilon(\psi_n,\varphi_n)=-\infty$ for every n large enough. This proves the proposition because $S_{\mathrm{M}}(\psi,\varphi_\infty)\geqq S(\psi_1,\varphi)=+\infty$.

Now suppose $s(\psi_1)\leqq s(\varphi)$. Then we have $s(\psi_n)\leqq \bigotimes_1^n s(\psi_1)\leqq \bigotimes_1^n s(\varphi)=s(\varphi_n)$ for every n. Define the selfadjoint operators H_n, $n\geqq 1$, in $\pi_\psi(\mathscr{A}_\infty)''$ by

$$H_n=\frac{1}{n}\left\{\pi_\psi\left(\log\frac{d\psi_n}{d\mathrm{Tr}_n}\right)-\pi_\psi\left(\log\frac{d\varphi_n}{d\mathrm{Tr}_n}\right)\right\}$$
$$=\frac{1}{n}\pi_\psi\left(\log\frac{d\psi_n}{d\mathrm{Tr}_n}\right)-\frac{1}{n}\sum_{i=0}^{n-1}\bar{\gamma}^i\left(\pi_\psi\left(\log\frac{d\varphi}{d\mathrm{Tr}}\right)\right).$$

Lemma 4.2 and the noncommutative mean ergodic theorem imply that H_n converges in $\bar{\psi}$-measure as $n\to\infty$ to

$$h_0=\lim_{n\to\infty}\frac{1}{n}\psi_n\left(\log\frac{d\psi_n}{d\mathrm{Tr}_n}-\log\frac{d\varphi_n}{d\mathrm{Tr}_n}\right)=S_{\mathrm{M}}(\psi,\varphi_\infty).$$

Let $0<\delta<1-\varepsilon$. For each n let p_n be the projection in $\mathscr{A}_n$ with $p_n\leqq s(\psi_n)$ such that $\pi_\psi(p_n)$ is the spectral projection of H_n corresponding to $(h_0-\delta, h_0+\delta)$. Then there exists n_0 such that $\psi(p_n)\geqq 1-\delta$ for all $n\geqq n_0$. Since $d\psi_n/d\mathrm{Tr}_n$, $d\varphi_n/d\mathrm{Tr}_n$ and p_n commute with one another, we get as (3.8),

$$p_n\geqq e^{-n(h_0+\delta)}p_n\left(\frac{d\psi_n}{d\mathrm{Tr}_n}\right)\left(\frac{d\varphi_n}{d\mathrm{Tr}_n}\right)^{-1}. \tag{4.2}$$

For each n choose a projection q_n in $\mathscr{A}_n$ such that $\psi(q_n)\geqq 1-\varepsilon$ and

$$\varphi_n(q_n)<\exp\{\beta_\varepsilon(\psi_n,\varphi_n)+1\}. \tag{4.3}$$

Let $q_n'=q_n\wedge p_n$. Then for $n\geqq n_0$ we get thanks to the traciality of ψ,

$$\psi(1-q_n')\leqq\psi(1-q_n)+\psi(1-p_n)\leqq\varepsilon+\delta. \tag{4.4}$$

Furthermore by (4.2)

$$\varphi_n(q_n)\geqq\varphi_n(q_n')=\mathrm{Tr}_n\left(q_n'\frac{d\varphi_n}{d\mathrm{Tr}_n}\right)$$
$$=\mathrm{Tr}_n\left(q_n'\left(p_n\frac{d\varphi_n}{d\mathrm{Tr}_n}\right)q_n'\right)$$
$$\geqq e^{-n(h_0+\delta)}\mathrm{Tr}_n\left(q_n'\left(p_n\frac{d\psi_n}{d\mathrm{Tr}_n}\right)q_n'\right)$$
$$=e^{-n(h_0+\delta)}\mathrm{Tr}_n\left(q_n'\frac{d\psi_n}{d\mathrm{Tr}_n}\right)=e^{-n(h_0+\delta)}\psi_n(q_n'). \tag{4.5}$$

By (4.3)–(4.5) we have for $n\geqq n_0$

$$\exp\{\beta_\varepsilon(\psi_n,\varphi_n)+1\}\geqq e^{-n(h_0+\delta)}(1-\varepsilon-\delta),$$

so that since $1-\varepsilon-\delta>0$,

$$\frac{1}{n}\{\beta_\varepsilon(\psi_n,\varphi_n)+1\}\geqq -h_0-\delta+\frac{1}{n}\log(1-\varepsilon-\delta).$$

Therefore

$$\liminf_{n\to\infty} \frac{1}{n} \beta_\varepsilon(\psi_n, \varphi_n) \geqq -h_0 - \delta.$$

This implies (4.1). □

Acknowledgements. The first author wishes to thank Professor Hiroshi Nagaoka who suggested to him the asymptotic theory for relative entropy. The second author is grateful to Professors L. Accardi and I. Csiszár for discussions on abelian subalgebras and I-divergency. The authors would also like to thank the referee for many valuable remarks on the first version of this paper.

References

1. Araki, H.: Relative entropy of states of von Neumann algebras. Publ. RIMS, Kyoto Univ. **11**, 809–833 (1976)
2. Araki, H.: Relative entropy for states of von Neumann algebras. II. Publ. RIMS, Kyoto Univ. **13**, 173–192 (1977)
3. Billingsley, P.: Ergodic theory and information. New York: Wiley 1965
4. Blahut, R.E.: Principles and practice of information theory. Reading, MA: Addison-Wesley 1987
5. Belavkin, V.P., Staszewski, P.: C^*-algebraic generalization of relative entropy and entropy. Ann. Inst. H. Poincaré Sect. A **37**, 51–58 (1982)
6. Connes, A., Størmer, E.: Entropy for automorphisms of II_1 von Neumann algebras. Acta Math. **134**, 289–306 (1975)
7. Csiszár, I., Körner, J.: Information theory, coding theorems for discrete memoryless systems. Budapest: Akadémiai Kiadó/Orlando: Academic Press 1981
8. Donald, M.J.: On the relative entropy. Commun. Math. Phys. **105**, 13–34 (1986)
9. Donald, M.J.: Continuity and relative hamiltonians. Commun. Math. Phys. **136**, 625–632 (1991)
10. Doplicher, S., Kastler, D.: Ergodic states in a non-commutative ergodic theory. Commun. Math. Phys. **7**, 1–20 (1968)
11. Ellis, R.S.: Entropy, large deviations, and statistical mechanics. Berlin, Heidelberg, New York: Springer 1985
12. Fujii, J.I., Kamei, E.: Relative operator entropy in noncommutative information theory. Math. Japon. **34**, 341–348 (1989)
13. Helstrom, C.W.: Quantum detection and estimation theory. New York: Academic Press 1976
14. Hiai, F., Ohya, M., Tsukada, M.: Sufficiency, KMS condition and relative entropy in von Neumann algebras. Pacific J. Math. **96**, 99–109 (1981)
15. Hiai, F., Ohya, M., Tsukada, M.: Sufficiency and relative entropy in *-algebras with applications in quantum systems. Pacific J. Math. **107**, 117–140 (1983)
16. Kosaki, H.: Relative entropy of states: a variational expression. J. Operator Theory **16**, 335–348 (1986)
17. Kullback, S., Leibler, R.A.: On information and sufficiency. Ann. Math. Statist. **22**, 79–86 (1951)
18. Kovács, I., Szücs, J.: Ergodic type theorems in von Neumann algebras. Acta Sci. Math. **27**, 233–246 (1966)
19. Ohya, M., Petz, D.: Notes on quantum entropy. Preprint
20. Ohya, M., Tsukada, M., Umegaki, H.: A formulation of noncommutative McMillan theorem. Proc. Jpn. Acad. **63**A, 50–53 (1987)
21. Parry, W.: Entropy and generators in ergodic theory. New York: Benjamin 1969
22. Petz, D.: Properties of quantum entropy. In: Accardi, L., von Waldenfels, W. (eds.) Quantum probability and applications. II. Lect. Notes Math., Vol. 1136, pp. 428–441. Berlin, Heidelberg, New York: Springer 1985
23. Petz, D.: Properties of the relative entropy of states of von Neumann algebras. Acta Math. Hungar. **47**, 65–72 (1986)

24. Petz, D.: Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. Commun. Math. Phys. **105**, 123–131 (1986)
25. Petz, D.: First steps towards a Donsker and Varadhan theory in operator algebras. Lect. Notes Math., Vol. 1442, pp. 311–319. Berlin, Heidelberg, New York: Springer 1990
26. Petz, D.: On certain properties of the relative entropy of states of operator algebras. Math. Z. (to appear)
27. Petz, D.: Characterization of the relative entropy of states of matrix algebras. Preprint
28. Petz, D., Raggio, G.A., Verbeure, A.: Asymptotics of Varadhan-type and Gibbs variational principle. Commun. Math. Phys. **121**, 271–282 (1989)
29. Raggio, G.A., Werner, R.F.: Quantum statistical mechanics of general mean field systems. Helv. Phys. Acta **62**, 980–1003 (1989)
30. Størmer, E.: Large groups of automorphisms of C^*-algebras. Commun. Math. Phys. **5**, 1–22 (1967)
31. Takesaki, M.: Conditional expectations in von Neumann algebras. J. Funct. Anal. **9**, 306–321 (1972)
32. Uhlmann, A.: Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. Commun. Math. Phys. **54**, 21–32 (1977)
33. Umegaki, H.: Conditional expectation in an operator algebra, IV (entropy and information). Kōdai Math. Sem. Rep. **14**, 59–85 (1962)

Communicated by H. Araki