

Commuting squares and a new relative entropy

By Yuki SEO

(Received Aug. 25, 1994)

(Revised May 8, 1995)

1. Introduction.

A. Connes and E. Størmer [3] defined the relative entropy $H(M|N)$ for finite dimensional von Neumann subalgebras M and N of a finite von Neumann algebra L by using the notion of Umegaki relative entropy. When L is commutative, M and N are generated by some finite partitions P and Q , then $H(M|N)$ coincides with the classical conditional entropy $h(P, Q)$ ([2]). Later, M. Pimsner and S. Popa [13] investigated the relation between the relative entropy $H(M|N)$ and the Jones index $[M:N]$, where N is a subfactor of a factor M of type II_1 . Very recently Y. Watatani and J. Wierzbicki [20] computed the relative entropy $H(M|N)$ for two subfactors M and N of a factor of type II_1 without assuming $N \subset M$, which extended the classical formula $h(P, Q) = h(P \vee Q, Q)$ in ergodic theory to the non-commutative case. They showed that the commuting square condition implies $H(M|N) = H(M|M \cap N)$ and the commuting square condition for commutants implies $H(M|N) = H(M \vee N|N)$.

Now, J.I. Fujii and E. Kamei [5] introduced the relative operator entropy $s(a|b)$ for positive operators a, b as a relative version of the Nakamura-Umegaki operator entropy. In the case where a, b are commutative, this relative operator entropy coincides with the Umegaki relative entropy, but in general they do not coincide. On the other hand Belavkin and Staszewski had defined in [1] a relative entropy s_{BS} in C^* -algebra setting. F. Hiai and D. Petz [10] pointed out that $s_{BS}(a, b) \geq -Tr(s(a|b))$ for density matrices a, b where Tr denotes the usual trace matrices. In noncommutative probability theory, F. Hiai investigated the relation between the Umegaki relative entropy and Belavkin and Staszewski relative entropy and showed some remarkable results in [9], [10].

In the previous paper [15, 16], we introduced an entropy $S(M|N)$ of a finite von Neumann algebra M relative to its subalgebra N as a noncommutative version of the Umegaki relative entropy which is not identical with the Connes-Størmer relative entropy $H(M|N)$ and showed a version of the Pimsner-Popa Theorem on the relative entropy and the Jones index for the factors of type II_1 .

In this paper we shall compute the relative entropy $S(M|N)$ without assuming $N \subset M$ and investigate the difference between the relative entropy

$S(M|N)$ and $H(M|N)$.

2. Relative entropy.

Following after [4] the relative operator entropy $s(a|b)$ for positive operators a and b is given by

$$s(a|b) = -\lim_{\varepsilon \rightarrow 0} a^{1/2}(\log a^{1/2}(b+\varepsilon)^{-1}a^{1/2})a^{1/2},$$

if the strong limit exists.

The relative operator entropy for noninvertible positive operators does not always exist, but if there exists $\lambda \geq 0$ such that $b \geq \lambda a$, then $s(a|b)$ exists and

$$(\log \lambda)a \leq s(a|b) \leq -a \log a + (\log \|b\|)a.$$

First we summarize the basic properties concerning the relative operator entropy and use them frequently. If $s(a|b)$ exists, then

- (2-1) monotony: $b \leq c$ imply $s(a|b) \leq s(a|c)$,
- (2-2) transformer inequality: $x*s(a|b)x \leq s(x*ax|x*bx)$ for all x ,
- (2-3) subadditive: $s(a|c) + s(b|d) \leq s(a+b|c+d)$,
- (2-4) upper semicontinuity: $b_n \downarrow b$ implies $s(a|b_n) \downarrow s(a|b)$,

where a, b, c and d are positive operators. Note that in (2-2) $x*s(a|b)x = s(x*ax|x*bx)$ for an invertible operator x . For some general results on the relative operator entropy, see [4], [5].

Let L be a finite von Neumann algebra with a fixed faithful normal normalized trace τ . Let M and N be von Neumann subalgebras of L . We denote by E_M and E_N the unique faithful normal τ -preserving conditional expectations onto M and N . Let $S(L)$ be the set of all finite families (x_1, \dots, x_n) of positive elements of L and satisfying

$$\sum_{i=1}^n x_i = 1.$$

In [15], we defined the relative entropy by

$$S(M|N) = \sup_{S(L)} \sum_i -\tau(s(E_M(x_i)|E_N E_M(x_i))).$$

In the above definition we may replace $S(L)$ by $S(D)$ for any von Neumann algebra D with $M \vee N \subset D \subset L$ without affecting the value $S(M|N)$. Here $M \vee N$ is the von Neumann algebra generated by M and N in L . In fact, since $E_M(x_i) = E_M E_D(x_i)$ for $x_i \in L$, we have

$$\begin{aligned} & \sup_{\langle x_i \rangle \in S(L)} \sum_i -\tau(s(E_M(x_i) | E_N E_M(x_i))) \\ &= \sup_{\langle x_i \rangle \in S(L)} \sum_i -\tau(s(E_M E_D(x_i) | E_N E_M E_D(x_i))) \\ &\leq \sup_{\langle y_i \rangle \in S(D)} \sum_i -\tau(s(E_M(y_i) | E_N E_M(y_i))). \end{aligned}$$

If M is a factor of type II_1 and N is a subfactor of M , then as shown in [16], the relative entropy $S(M|N)$ coincides with the Jones index $[M : N]$ in the sense of that $S(M|N) = \log[M : N]$. In particular we have $H(M|N) \leq S(M|N)$.

First of all, we shall show that this relative entropy extends the classical conditional entropy in commutative probability theory.

Let (X, \mathcal{F}, p) be a probability space and

$$A = \{A_1, \dots, A_m\}, B = \{B_1, \dots, B_n\}, D = \{A_k \cap B_i\}_{k \leq m, i \leq n}$$

are finite partitions of X , then we may consider the following von Neumann algebras:

$$M = L^\infty(X, \sigma(A)), N = L^\infty(X, \sigma(B)), L = L^\infty(X, \sigma(D)).$$

The trace on L corresponds to the expected value of a random variable, $\tau(g) = \int_X g dp$ and E_N, E_M are conditional expectations in the sense of the probability theory. In this case,

$$h(A, B) = \sum_{k=1}^m p(B_k) \sum_{i=1}^n \eta(p(A_i | B_k))$$

is the conditional entropy of the partition A given B in the ergodic theory ([2]).

THEOREM 1. *In the above situation, the relative entropy $S(M|N)$ coincides with the classical conditional entropy $h(A, B)$.*

PROOF. We may suppose that

$$L = \bigoplus_{k,i} L_{k,i}, M = \bigoplus_k M_k, N = \bigoplus_i N_i,$$

where $L_{k,i} \cong C, M_k \cong C, N_i \cong C$ and C is a complex field. We denote by $t_{k,i}$ respectively s_k and u_i the traces of the minimal projections $f^{k,i}$ in L respectively e^k in M and g^i in N . Also we have $\sum_i t_{k,i} = s_k$ and $\sum_k t_{k,i} = u_i$. Note that the conditional expectation $E_M, E_N E_M$ acts as follows:

$$\begin{aligned} E_M(f^{k,i}) &= \frac{t_{k,i}}{s_k} e^k = \frac{t_{k,i}}{s_k} \sum_j f^{k,j} \\ E_N E_M(f^{k,i}) &= \frac{t_{k,i}}{s_k} \sum_j \frac{t_{k,j}}{u_j} g^j. \end{aligned}$$

Let $\{x_l\}$ be a partition of the unity in L and write $x_l = \sum_{k,i} c_{k,i}^l f^{k,i}$ with $c_{k,i}^l \in \mathbf{R}_+$. We get $\sum_l c_{k,i}^l = 1$ by $\sum_l x_l = 1$. By (2-3), we have

$$\begin{aligned} & -\sum_l \tau(s(E_M(x_l) | E_N E_M(x_l))) \\ & \leq -\sum_l \sum_{k,i} \tau(s(E_M(c_{k,i}^l f^{k,i}) | E_N E_M(c_{k,i}^l f^{k,i}))) \\ & = -\sum_l \sum_{k,i} c_{k,i}^l \tau(s(E_M(f^{k,i}) | E_N E_M(f^{k,i}))) \\ & = -\sum_{k,i} \tau(s(E_M(f^{k,i}) | E_N E_M(f^{k,i}))). \end{aligned}$$

In order to prove Theorem 1, it is sufficient to show that the above last term is equal to $h(A, B)$ since $\{f^{k,i}\}$ is a partition of the unity in L .

By the commutativity of L we get that

$$\begin{aligned} & \sum_{k,i} -\tau(s(E_M(f^{k,i}) | E_N E_M(f^{k,i}))) \\ & = -\sum_{k,i} \frac{t_{k,i}}{s_k} \tau\left(s\left(e^k \mid \sum_j \frac{t_{k,j}}{u_j} g^j\right)\right) \\ & = -\sum_{k,i} \frac{t_{k,i}}{s_k} \tau\left(e^k \sum_j \left(\log \frac{t_{k,j}}{u_j}\right) g^j\right) \\ & = -\sum_{k,i} \frac{t_{k,i}}{s_k} \sum_j \left(\log \frac{t_{k,j}}{u_j}\right) t_{k,j} \\ & = -\sum_{k,j} t_{k,j} \log \frac{t_{k,j}}{u_j}. \quad \square \end{aligned}$$

3. Commuting squares and the relative entropy.

Let us denote a diagram $\begin{matrix} M \subset L \\ \cup \quad \cup \\ K \subset N \end{matrix}$ of finite von Neumann algebras with a fixed finite faithful normal trace τ on L by (L, M, N, K) . Then a diagram (L, M, N, K) is a commuting square if $E_M E_N = E_N E_M$ and $K = M \cap N$. ([6])

Also, T. Sano and Y. Watatani [14] introduced a dual notion of co-commuting square in study of angles of two subfactors:

DEFINITION 2. A diagram (L, M, N, K) of finite von Neumann algebras with a fixed finite faithful normal trace τ' on K' is a co-commuting square if their commutants

$$\begin{matrix} M' \subset K' \\ \cup \quad \cup \\ L' \subset N' \end{matrix} \text{ form a commuting square.}$$

Since $L' = M' \cap N'$, it is necessary for $L = M \vee N$ to hold. Throughout the paper we consider only the case of K' being a finite factor.

Though it is true the classical formula $h(A, B) = h(A \vee B, B)$ in ergodic theory, $H(M|N) = H(M \vee N|N)$ is not always true even if M and N are commuting finite dimensional von Neumann subalgebras of L . In [20], Y. Watatani and J. Wierzbicki showed that the commutative case always satisfies the co-commuting square condition by suitable representation and a co-commuting square condition implies $H(M|N) = H(M \vee N|N)$.

The following two theorems are a modification of [20: Theorem 6, 7]. So, our results suggest that the relative entropy $S(M|N)$ may be another generalization of the classical conditional entropy in noncommutative frame. Now we consider the case of commuting squares.

THEOREM 3. *If the diagram (L, M, N, K) is a commuting square of finite von Neumann algebras, then*

$$S(M|N) = S(M|K).$$

PROOF. Since $E_N E_M = E_K$, we have

$$\begin{aligned} S(M|N) &= \sup_{s(L)} \sum_i -\tau(s(E_M(x_i)|E_N E_M(x_i))) \\ &= \sup_{s(L)} \sum_i -\tau(s(E_M(x_i)|E_K E_M(x_i))) \\ &= S(M|K). \end{aligned} \quad \square$$

COROLLARY 4. *If the diagram (L, M, N, K) is a commuting square of factors of type II_1 , then*

$$S(M|N) \geq H(M|N).$$

PROOF. By Theorem 3, [16: Theorem 8] and [13: Corollary 4.1], we have

$$S(M|N) = S(M|K) = \log[M:K] \geq H(M|K) = H(M|N). \quad \square$$

The proof of the next lemma is essentially same as that of [13: Lemma 4.2]. However it is the key result in proving the estimation of $S(M|N)$ from below.

LEMMA 5. *Let the diagram (L, M, N, K) be factors of type II_1 such that $[L:K] < \infty$. If $q \in M$ is a projection such that $E_{K' \cap M}(q) = cf$ for some scalar c and some projection $f \in K' \cap M$, then*

$$S(M|N) \geq -\frac{1}{c} \tau(s(q|E_N(q))).$$

PROOF. Let $x = q - cf$ and $\Omega = \overline{c\omega}^w \{vxv^*; v \in U(K)\}$, where $U(K)$ is group of all unitary operators in K . It follows that Ω is a weakly compact convex

subset of M , therefore there exists a unique $y_o \in \Omega$ such that $\|y_o\|_2 = \inf\{\|y\|_2; y \in \Omega\}$. Hence we have $y_o \in (U(K))' = K'$. Since $E_{K' \cap M}(v_x v_x^*) = 0$ for $v \in U(K)$, we have $E_{K' \cap M}(y) = 0$ for all $y \in \Omega$, in particular $y_o = E_{K' \cap M}(y_o) = 0 \in \Omega$. Therefore for any $\varepsilon > 0$ there exist unitary elements v_1, \dots, v_n in K such that $\|y - f\|_2 < \varepsilon \|f\|_2$ with $y = (1/cn) \sum v_i q v_i^*$. Let $\delta > 0$ and denote by p the spectral projection of y corresponding to $[0, 1 + \delta]$ in the algebra fMf . Put

$$x_i = \frac{1}{(1 + \delta)cn} p \wedge v_i q v_i^*, \quad z_i = v_i q v_i^* - p \wedge v_i q v_i^*.$$

Since $\sum x_i \leq f$, $\tau(z_i) \leq (\varepsilon \delta^{-1})^2 \tau(f)$ and $s(E_M(v_i q v_i^*) | E_N E_M(v_i q v_i^*)) = s(v_i q v_i^* | E_N(v_i q v_i^*)) = v_i s(q | E_N(q)) v_i^*$, we have

$$\begin{aligned} & - \sum \tau(s(E_M(x_i) | E_N E_M(x_i))) \\ \geq & - \sum \frac{1}{(1 + \delta)cn} \tau(s(v_i q v_i^* | E_N(v_i q v_i^*))) - \frac{1}{(1 + \delta)cn} \sum \tau(s(z_i | E_N(z_i))) \\ \geq & - \frac{1}{(1 + \delta)c} \tau(s(q | E_N(q))) - \frac{1}{(1 + \delta)c} (\log \lambda) (\varepsilon \delta^{-1})^2 \tau(f), \end{aligned}$$

where $\lambda = [L : N]^{-1}$. Letting first $\varepsilon \rightarrow 0$ and then $\delta \rightarrow 0$, then we have $S(M | N) \geq -(1/c) \tau(s(q | E_N(q)))$. □

Following after [20], let $K \subset M$ be factors of type II_1 with $[M : K] < \infty$. The inclusion $K \subset M$ is called extremal if it satisfies any of the following conditions:

- (1) $H(M | K) = \log [M : K]$.
- (2) $E_{K' \cap M}(e) = [M : K]^{-1}$, where $e \in P(M)$ satisfies $E_K(e) = [M : K]^{-1}$.
- (3) if τ' is the normalized trace on K' , then $\tau_{|_{K' \cap M}} = \tau'_{|_{K' \cap M}}$.
- (4) the trace preserving conditional expectation coincides with the expectation of minimal index.

THEOREM 6. *If the diagram (L, M, N, K) is a co-commuting square of factors of type II_1 such that $[L : K] < \infty$ and the inclusion $K \subset M$ is extremal, then*

$$S(M | N) = S(L | N) = \log [L : N].$$

PROOF. The proof in the following is a slight modification of that of [20: Theorem 7]. Let Q be the downward basic construction for $K \subset M$ and e its Jones projection. Since $[L : K] < \infty$, we can consider the downward basic construction. Since $E_K(e) = [M : K]^{-1}$ and the inclusion $K \subset M$ is extremal, we have $E_{K' \cap M}(e) = [M : K]^{-1}$. By Lemma 5, it follows that

$$S(M | N) \geq -[M : K] \tau(s(e | E_N(e))).$$

Since the diagram $\begin{matrix} M \subset L \\ \cup & \cup \\ K \subset N \end{matrix}$ is a co-commuting square, by [20: Lemma 5] it follows that $a=[L:N]E_N(e)$ is a projection in N . Therefore e and $E_N(e)$ commute. Then we have

$$\begin{aligned} \tau(s(e|E_N(e))) &= \frac{1}{[L:N]} \tau(s([L:N]e|a)) \\ &= \frac{1}{[L:N]} \tau(-[L:N]e \log [L:N]e + a \log a) \\ &= -\tau(e) \log [L:N]. \end{aligned}$$

Hence we have $S(M|N) \geq [M:K] \tau(e) \log [L:N] = \log [L:N] = S(L|N)$.

As $S(M|N)$ is increasing in M , this gives the theorem. □

COROLLARY 7. *If the diagram (L, M, N, K) is a co-commuting square of factors of type Π_1 such that $[L:K] < \infty$ and the inclusion $K \subset M$ is extremal, then*

$$S(M|N) = H(M|N).$$

PROOF.

$$S(M|N) = S(L|N) = \log [L:N] = H(L|N) = H(M|N). \quad \square$$

COROLLARY 8. *If the diagram (L_1, M_1, N_1, K_1) is a commuting square of factors of type Π_1 and the diagram (L_2, M_2, N_2, K_2) is a co-commuting square of factors of type Π_1 such that $[L_2:K_2] < \infty$ and the inclusion $K_2 \subset M_2$ is extremal, then*

$$S(M_1 \otimes M_2 | N_1 \otimes N_2) \geq H(M_1 \otimes M_2 | N_1 \otimes N_2).$$

PROOF. Let $a_i \in L_{1,+}$ with $\sum_i a_i = 1$ and $b_j \in L_{2,+}$ with $\sum_j b_j = 1$. Then $\sum_{i,j} a_i \otimes b_j = 1 \otimes 1$ and it follows that

$$\begin{aligned} &S(M_1 \otimes M_2 | N_1 \otimes N_2) \\ &\geq -\sum_{i,j} \tau_{L_1 \otimes L_2}(s(E_{M_1 \otimes M_2}(a_i \otimes b_j) | E_{N_1 \otimes N_2} E_{M_1 \otimes M_2}(a_i \otimes b_j))) \\ &= -\sum_{i,j} \tau_{L_1}(s(E_{M_1}(a_i) | E_{N_1} E_{M_1}(a_i))) \tau_{L_2}(E_{M_2}(b_j)) \\ &\quad -\sum_{i,j} \tau_{L_1}(E_{M_1}(a_i)) \tau_{L_2}(s(E_{M_2}(b_j) | E_{N_2} E_{M_2}(b_j))) \\ &= -\sum_i \tau_{L_1}(s(E_{M_1}(a_i) | E_{N_1} E_{M_1}(a_i))) - \sum_j \tau_{L_2}(s(E_{M_2}(b_j) | E_{N_2} E_{M_2}(b_j))). \end{aligned}$$

Hence $S(M_1 \otimes M_2 | N_1 \otimes N_2) \geq S(M_1 | N_1) + S(M_2 | N_2)$.

By Corollary 4, 7 and [20: Example 11] we get

$$\begin{aligned} S(M_1 \otimes M_2 | N_1 \otimes N_2) &\geq S(M_1 | N_1) + S(M_2 | N_2) \\ &\geq H(M_1 | N_1) + H(M_2 | N_2) \\ &= H(M_1 \otimes M_2 | N_1 \otimes N_2). \end{aligned} \quad \square$$

4. Examples.

Here we shall show that there is a slight difference between the relative entropy $S(M|N)$ and $H(M|N)$. By the same example in [20: Example 8], it follows that the formula $S(M|N) = S(M \vee N|N)$ does not hold in general. We note that the equality in Corollary 4 is not always true.

EXAMPLE 9. Let $M = (M_2 \oplus C) \otimes (C \oplus C)$, $L = (M_2 \oplus C) \otimes (M_2 \oplus C)$, $N = (C \oplus C) \otimes (M_2 \oplus C)$ and $K = (C \oplus C) \otimes (C \oplus C)$, where M_2 is the algebra of 2×2 matrix. Since $M \cap N = K$ and $E_M(N) \subset K$, it follows that the diagram (L, M, N, K) is a commuting square. On the other hand, since

$$\begin{aligned} M &\cong (M_2 \oplus C) \oplus (M_2 \oplus C) \\ K &\cong (C \oplus C) \oplus (C \oplus C), \end{aligned}$$

by [15: Example 11] and [13: Theorem 6.2] we have the following:

$$\begin{aligned} H(M|K) &= 2H(M_2 \oplus C | C \oplus C) = \frac{3}{2} \log 3, \\ S(M|K) &= 2S(M_2 \oplus C | C \oplus C) = \log 6. \end{aligned}$$

Therefore we have $S(M|N) = S(M|K) \geq H(M|K) = H(M|N)$.

EXAMPLE 10. For $\lambda < 1/4$, let R_λ be Jones' subfactors of the hyperfinite factor R with $[R : R_\lambda] = \lambda^{-1}$. Then it follows from [13: Corollary 5.3] that $H(R|R_\lambda) = 2\eta t + 2\eta(1-t)$, where $t(1-t) = \lambda$.

Let q be a positive number such that $\lambda^{-1} = 2 + q + q^{-1}$, and define $g = qe_R - (1 - e_R)$. Then by [6: Example 4.2.10] the diagram $(\langle R, e_R \rangle, R, gRg^{-1}, R_\lambda)$ is a commuting square. Therefore we have $H(R|gRg^{-1}) = H(R|R_\lambda) = 2\eta t + 2\eta(1-t)$. Also $S(R|gRg^{-1}) = S(R|R_\lambda) = \log [R : R_\lambda] = \log \lambda^{-1}$. Hence $H(R|gRg^{-1}) \leq S(R|gRg^{-1})$.

References

- [1] V.P. Belavkin and P. Staszewski, C^* -algebraic generalization of relative entropy and entropy, *Ann. Inst. H. Poincaré*, **37** (1982), 51-58.
- [2] P. Billingsly, *Ergodic Theory and Information*, Wiley Sons, 1965.
- [3] A. Connes and E. Størmer, Entropy for automorphisms of II_1 von Neumann algebras, *Acta. Math.*, **134** (1975), 288-306.

- [4] J.I. Fujii, M. Fujii and Y. Seo, An extension of the Kubo-Ando theory : Solidarities, *Math. Japon.*, **35** (1990), 387-396.
- [5] J.I. Fujii and E. Kamei, Relative operator entropy in noncommutative information theory, *Math. Japon.*, **34** (1989), 341-348.
- [6] F. Goodman, P. de la Harpe and V. F. R. Jones, *Coxeter Graphs and Towers of Algebras*, Springer-Verlag, 1989.
- [7] F. Hiai, Minimum index for subfactors and entropy, *J. Operator Theory*, **24** (1990), 301-336.
- [8] F. Hiai, Minimum index for subfactors and entropy, II, *J. Math Soc. Japan*, **43** (1991), 347-379.
- [9] F. Hiai, Relative Hamiltonian and relative entropy, preliminary version.
- [10] F. Hiai and D. Petz, The proper formula for relative entropy and asymptotics in quantum probability, *Comm. Math. Phys.*, **143** (1991), 99-114.
- [11] V. F. R. Jones, Index for subfactors, *Invent. Math.*, **72** (1983), 1-25.
- [12] M. Nakamura and H. Umegaki, A note on the entropy for operator algebras, *Proc. Japan Acad.*, **37** (1961), 149-154.
- [13] M. Pimsner and S. Popa, Entropy and index for subfactors, *Ann. Sci. École Norm. Sup.*, **19** (1986), 57-106.
- [14] T. Sano and Y. Watatani, Angles between two subfactors, *J. Operator Theory*, **32** (1994), 209-241.
- [15] Y. Seo, A new relative entropy for von Neumann algebras, *Math. Japon.*, **36** (1991), 51-57.
- [16] Y. Seo, Relative entropy for finite von Neumann algebras, *J. Operator Theory*, **29** (1993), 373-387.
- [17] M. Takesaki, *Theory of Operator Algebras I*, Springer-Verlag, 1979.
- [18] H. Umegaki, Conditional expectation in an operator algebra, *Kôdai Math. Sem. Rep.*, **14** (1962), 59-85.
- [19] H. Umegaki and M. Ohya, *Quantum Mechanical Entropies* (in Japanese), Kyoritsu Publishing Company, 1984.
- [20] Y. Watatani and J. Wierzbicki, Commuting square and relative entropy for two subfactors, to appear in *J. Funct. Anal.*

Yuki SEO

Tennoji Senior Highschool
attached to Osaka Kyoiku University
Tennoji, Osaka 543
Japan