A

Limits and Linear Algebra

A.1 Limits

In this text, frequently we discuss the asymptotic behaviors in several problems when the number n of prepared systems is sufficiently large. In this situation, we often take the limit $n \to \infty$. In this section, we give a brief summary of the fundamental properties of limits. Given a general sequence $\{a_n\}$, the limit $\lim a_n$ does not necessarily exist. For example, a sequence a_n is a counterexample when a_n diverges to $+\infty$ or $-\infty$. In such a case, it is possible to at least denote these limits as $\lim a_n = +\infty$ or $\lim a_n = -\infty$. However, the sequence a_n has no limit as $n \to infinity$, even allowing possibilities such as $+\infty$ or $-\infty$, when a_n is defined to be 0 when n is even and 1 when it is odd. This is caused by its oscillatory behavior. In this case, we can consider the upper limit $\underline{\lim}\, a_n$ and the lower limit $\overline{\lim}\, a_n$, which are given as $\underline{\lim}\, a_n = 0$ and $\overline{\lim}\, a_n = 1$. More precisely, $\underline{\lim}\, a_n$ and $\overline{\lim}\, a_n$ are defined as follows:

$$\underline{\lim}\, a_n \stackrel{\text{def}}{=} \sup\{a | \forall \epsilon > 0, \exists N, \forall n \geq N, a \leq a_n + \epsilon\},$$
$$\overline{\lim}\, a_n \stackrel{\text{def}}{=} \inf\{a | \forall \epsilon > 0, \exists N, \forall n \geq N, a \geq a_n - \epsilon\}.$$

When $\underline{\lim}\, a_n = \overline{\lim}\, a_n$, the limit $\lim a_n$ exists and is equal to $\underline{\lim}\, a_n = \overline{\lim}\, a_n$. The following three lemmas hold concerning limits.

Lemma A.1 *Let sequences* $\{a_n\}_{n=1}^{\infty}$ *and* $\{b_n\}_{n=1}^{\infty}$ *satisfy*

$$a_n + a_m \leq a_{n+m} + b_{n+m}, \quad \sup_n \frac{a_n}{n} < \infty, \quad \lim_n \frac{b_n}{n} = 0.$$

Then, the limit $\lim \frac{a_n}{n}$ *exists and satisfies*

$$\lim \frac{a_n}{n} = \overline{\lim}\, \frac{a_n}{n} = \sup_n \frac{a_n}{n}. \tag{A.1}$$

If $a_n + a_m \geq a_{n+m} - b_{n+m}$ *and* $\inf_n \frac{a_n}{n} > \infty$, $b_n \to 0$, *then similarly* $\lim \frac{a_n}{n} = \underline{\lim}\, \frac{a_n}{n} = \inf_n \frac{a_n}{n}$, *as shown by considering* $-a_n$.

Proof. Fix the integer m. Then, for any integer n, there uniquely exist integers l_n and r_n such that $0 \leq r_n \leq m-1$ and $n = l_n m + r_n$ for each n. Thus, we have

$$\frac{a_n}{n} = \frac{a_{l_n m+r}}{l_n m + r} \geq \frac{a_{l_n m}}{l_n m + r} + \frac{a_r - b_n}{l_n m + r} \geq \frac{l_n a_m}{l_n m + r} + \frac{a_r - b_n - b_{l_n m}}{l_n m + r}.$$

Since $l_n \to \infty$ as $n \to \infty$, taking the limit $n \to \infty$, we have $\underline{\lim} \frac{a_n}{n} \geq \frac{a_m}{m}$ for arbitrary m. Next, taking the limit $m \to \infty$, we have $\underline{\lim} \frac{a_n}{n} \geq \sup_m \frac{a_m}{m} \geq \overline{\lim}_{m\to\infty} \frac{a_m}{m}$. Since $\overline{\lim} \frac{a_n}{n} \geq \underline{\lim} \frac{a_n}{n}$, we obtain (A.1). ■

Lemma A.2 *Let $\{a_n\}$ and $\{b_n\}$ be two sequences of positive real numbers. Then,*

$$\overline{\lim} \frac{1}{n} \log(a_n + b_n) = \max \left\{ \overline{\lim} \frac{1}{n} \log a_n, \overline{\lim} \frac{1}{n} \log b_n \right\}.$$

Proof. Since $(a_n + b_n) \geq a_n, b_n$ and

$$\overline{\lim} \frac{1}{n} \log(a_n + b_n) \geq \overline{\lim} \frac{1}{n} \log a_n, \overline{\lim} \frac{1}{n} \log b_n,$$

we obtain the $\geq$ part of the proof. Since $2 \max\{a_n, b_n\} \geq (a_n + b_n)$, we have

$$\begin{aligned} \max \left\{ \overline{\lim} \frac{1}{n} \log a_n, \overline{\lim} \frac{1}{n} \log b_n \right\} &= \overline{\lim} \frac{1}{n} \log \max\{a_n, b_n\} \\ = \overline{\lim} \frac{1}{n} \log 2 \max\{a_n, b_n\} &\geq \overline{\lim} \frac{1}{n} \log(a_n + b_n), \end{aligned}$$

which gives the reverse inequality. This completes the proof. ■

Lemma A.3 *Let $\{f_n(x)\}$ be a sequence of functions such that $f_n(x) \leq f_n(y)$ if $x \geq y$, and $f_n(x) \to 0$ if $x > 0$. There exists a sequence $\{\epsilon_n\}$ of positive real numbers converging to zero such that $f_n(x) \to 0$.*

Proof. Let N be a positive integer. Choose positive integers $n(N)$ such that $n(N) < n(N+1)$ and $f_n(\frac{1}{N}) \leq \frac{1}{N}$ for $n \geq n(N)$. We also define $\epsilon_n \stackrel{\text{def}}{=} \frac{1}{N}$ for $n(N) \leq n < n(N+1)$. Then, $\epsilon_n \to 0$. If $n \geq n(N)$, then $f_n(\epsilon_n) \leq \frac{1}{N}$. Therefore, $f_n(\epsilon_n) \to 0$. ■

For any two continuous functions f and g on an open subset $X \subset \mathbb{R}^d$, we define

$$[f, g](a) \stackrel{\text{def}}{=} \min_{x \in V} \{f(x) | g(x) \leq a\}. \tag{A.2}$$

Lemma A.4 *When X is closed and bounded, i.e., compact,*

$$[f, g](a) = \lim_{\epsilon \downarrow 0} [f, g](a + \epsilon). \tag{A.3}$$

Proof. From the definition for $\epsilon > 0$, $[f, g](a) \geq [f, g](a + \epsilon)$. Hence, $[f, g](a) \geq \lim_{\epsilon \downarrow 0} [f, g](a + \epsilon)$. From the compactness, for any $\epsilon_1 > 0$ there

exists $\epsilon_2 > 0$ such that $\|x - x'\| < \epsilon_2 \Rightarrow |f(x) - f(x')| < \epsilon_1$. Further, from the compactness of X we can choose a small number $\epsilon_3 > 0$ such that $\{x|g(x) \leq a + \epsilon_3\} \subset \cup_{x':g(x)\leq a} U_{x',\epsilon_2}$. Hence,

$$\min_{x|g(x)\leq a+\epsilon_3} f(x) \geq \min_{x\in\cup_{x':g(x)\leq a} U_{x',\epsilon_2}} f(x) \geq \min_{x|g(x)\leq a} f(x) - \epsilon_1, \tag{A.4}$$

which implies (A.3). ■

A.2 Singular Value Decomposition and Polar Decomposition

Any $d \times d'$ complex-valued matrix X has the form

$$X = U_1 X' U_2^*$$

with isometric matrices U_1 and U_2 and a diagonal matrix X'. This is called a *singular value decomposition* (the matrix U is an *isometric matrix* if U^*U is the identity matrix; U is a *partially isometric matrix* for the partial space $\mathcal{K}$ if it is a projection onto the partial space $\mathcal{K}$). Choosing a $d \times d'$ partially isometric matrix U in the range $\{X^*Xv|v \in \mathbb{C}^d\}$ of X^*X, we have

$$X = U|X|, \quad |X| \stackrel{\text{def}}{=} \sqrt{X^*X}, \tag{A.5}$$

which is called a *polar decomposition*. If X is Hermitian and is diagonalizable according to $X = \sum_i \lambda_i |u_i\rangle\langle u_i|$, then $|X| = \sum_i |\lambda_i||u_i\rangle\langle u_i|$. Since $X^* = |X|U^*$,

$$XX^* = U|X||X|U^* = UX^*XU^*, \quad \sqrt{XX^*} = U\sqrt{X^*X}U^*, \tag{A.6}$$

$$UX^*U = X. \tag{A.7}$$

Therefore,

$$X = \sqrt{XX^*}U. \tag{A.8}$$

If X is a square matrix (i.e., $d = d'$), then U is unitary. If $d \geq d'$, then U can be chosen as an isometric. If $d \leq d'$, U can be chosen such that U^* is isometric. We now show that these two decompositions exist.

Since X^*X is Hermitian, we may choose a set of mutually orthogonal vectors $u_1, \ldots, u_l$ of norm 1 such that

$$X^*X = \sum_{i=1}^{l} \lambda_i |u_i\rangle\langle u_i|.$$

In the above, we choose $\{\lambda_k\}_{k=1}^l$ such that $\lambda_i \geq \lambda_{i+1} > 0$. Hence, l is not necessarily equal to the dimension of the space because there may exist zero eigenvalues. Defining $v_i \stackrel{\text{def}}{=} \sqrt{\frac{1}{\lambda_i}} X u_i$, we have

$$\langle v_i|v_j\rangle = \sqrt{\frac{1}{\lambda_i}}\sqrt{\frac{1}{\lambda_j}}\langle Xu_i|Xu_j\rangle = \sqrt{\frac{1}{\lambda_i}}\sqrt{\frac{1}{\lambda_j}}\langle u_i|X^*X|u_j\rangle$$
$$= \sqrt{\frac{1}{\lambda_i}}\sqrt{\frac{1}{\lambda_j}}\delta_{i,j}\lambda_j = \delta_{i,j}.$$

Furthermore, from the relation

$$\langle v_i|X|u_j\rangle = \sqrt{\frac{1}{\lambda_i}}\langle Xu_i|X|u_j\rangle = \sqrt{\frac{1}{\lambda_i}}\langle Xu_i|X|u_j\rangle$$
$$= \sqrt{\frac{1}{\lambda_i}}\langle u_i|X^*X|u_j\rangle = \sqrt{\lambda_i}\delta_{i,j}$$

we can show that

$$\sum_i \sqrt{\lambda_i}|v_i\rangle\langle u_i| = \sum_i |v_i\rangle\langle v_i|X\sum_j |u_j\rangle\langle u_j| = X.$$

One may be concerned about the validity of the second equality if X^*X has some eigenvectors u with zero eigenvalue. However, since $\langle u|X^*X|u\rangle = 0$, we have $Xu = 0$. Hence, both sides of the image of vector u coincide with the 0 vector. We define $U_2 \stackrel{\text{def}}{=} (u_i^j)$ and $U_1 \stackrel{\text{def}}{=} (v_i^j)$, which are $l \times d$ and $l \times d'$ isometric matrices, respectively. Let X' be an $l \times l$ diagonal matrix $(\sqrt{\lambda_i}\delta_{i,j})$. This gives us (A.5).

Using the above, we obtain the following lemma.

Lemma A.5 *Let a density matrix ρ be written as*

$$\rho = \sum_{j=1}^{d'} |v_j\rangle\langle v_j|, \tag{A.9}$$

where $\{v_i\}$ is a set of vectors that are not necessarily orthogonal. Let its diagonalization be given by $\rho = \sum_{i=1}^{l} \lambda_i|u_i\rangle\langle u_i|$. Since $\lambda_i > 0$, l is not necessarily equal to the dimension of the space. Then, the vector v_j can be written as $v_j = \sum_{i=1}^{l} w_{j,i}\sqrt{\lambda_i}u_i$ by using an $l \times d'$ isometric matrix $W = (w_{j,i})$ [316].

The set of vectors $\{v_i\}$ satisfying (A.9) is called the *decomposition* of the density matrix ρ.

Proof. Let Y be a $j \times d$ matrix given by (v_i^i). Then,

$$\rho = \sum_{i=1}^{l} \lambda_i|u_i\rangle\langle u_i| = YY^*.$$

Define $w_i \stackrel{\text{def}}{=} \sqrt{\frac{1}{\lambda_i}}Y^*u_i$. Then, $Y^* = \sum_{i=1}^{l} \sqrt{\lambda_i}|w_i\rangle\langle u_i|$. Taking its conjugate, we obtain $Y = \sum_{i=1}^{l} \sqrt{\lambda_i}|u_i\rangle\langle w_i|$. Looking at the jth row, we obtain $|v_j\rangle = \sum_{i=1}^{l}(w_i^j)^*\sqrt{\lambda_i}|u_i\rangle$. Since $\sum_j (w_i^j)^*(w_{i'}^j) = \delta_{i,i'}$, $(w_i^j)^*$ is an isometric matrix. The proof is complete. ■

Next, we consider the case where X is a real $d \times d$ matrix. Since a real symmetric matrix can be diagonalized by an orthogonal matrix, the unitary matrices U_1 and U_2 may be replaced by orthogonal matrices O_1 and O_2. In fact, we may further restrict the orthogonal matrices to orthogonal matrices with determinant 1 (these are called special orthogonal matrices). However, the following problem occurs. Assume that the determinant of O_i $(i = 1, 2)$ is -1. Then, O_i may be redefined by multiplying it by a diagonal matrix with diagonal elements $-1, 1, \ldots, 1$. The redefined matrix is then a special orthogonal matrix, and $O_1^* X O_2$ is diagonal. Choosing O_1 and O_2 in a suitable way, all the diagonal elements of $O_1^* X O_2$ will be positive if $\det X > 0$. On the other hand, if $\det X < 0$, then it is not possible to make all the diagonal elements of $O_1^* X O_2$ positive for special orthogonal matrices O_1, O_2.

Exercises

A.1. Define $J_{i,j} \stackrel{\text{def}}{=} \langle u_i | u_j \rangle$ for a set of linearly independent vectors $u_1, \ldots, u_k$ in $\mathcal{H}$. Show that

$$\sum_{i,j} (J^{-1})^{j,i} |u_i\rangle\langle u_j| = \sum_{i,j} ((J^{-1})^{j,i})^* |u_i\rangle\langle u_j|.$$

Show that this is a projection to the subspace of $\mathcal{H}$ spanned by $u_1, \ldots, u_k$.

A.2. Using relation (A.5), show that

$$AA^* f(AA^*) = Af(A^*A) = A^*. \tag{A.10}$$

A.3 Norms of Matrices

We often focus on the norm between two matrices as a measure of the difference between them. There are two types of norms, the matrix norm and the trace norm. The matrix norm $\|A\|$ of a matrix A is defined as

$$\|A\| \stackrel{\text{def}}{=} \max_{\|x\|=1} \|Ax\|.$$

Since $\|x\| = \max_{\|y\|=1} |\langle y, x\rangle|$, we have $\|A\| = \max_{\|y\|=\|x\|=1} |\langle y, Ax\rangle|$; therefore, $\|A\| = \|A^*\|$. From the definition we have $\|U_1 A U_2\| = \|A\|$ for unitary matrices U_1 and U_2. Defining

$$w(A) \stackrel{\text{def}}{=} \max_{\|x\|=1} |\langle x, Ax\rangle|, \quad \mathrm{spr}(A) \stackrel{\text{def}}{=} \max\{|\lambda| : \lambda \text{ is the eigenvalue of } A\},$$

we obtain

$$\mathrm{spr}(A) \le w(A) \le \|A\|. \tag{A.11}$$

Assume that A is a Hermitian matrix. Then, it may be diagonalized as $A = \sum_{i=1}^{d} \lambda_i |u_i\rangle\langle u_i|$. Thus,

$$|\langle y, Ax\rangle| = \sum_{i=1}^{d} |\lambda_i||\langle y|u_i\rangle\langle u_i|x\rangle| \le \max_i |\lambda_i| \sum_{i=1}^{d} |\langle y|u_i\rangle||\langle u_i|x\rangle|$$
$$\le \max_i |\lambda_i| \sqrt{\sum_{i=1}^{d} |\langle y|u_i\rangle|^2} \sqrt{\sum_{i=1}^{d} |\langle u_i|x\rangle|^2} = \max_i |\lambda_i| = \mathrm{spr}(A).$$

The above inequality implies the equality sign in (A.11). Since $\|A\|^2 = \max_{\|x\|=1} \langle x|A^*A|x\rangle = \mathrm{spr}(A^*A) = (\mathrm{spr}(\sqrt{A^*A}))^2$, then $\|A\| = \|\sqrt{A^*A}\| = \|A^*\| = \|\sqrt{AA^*}\|$.

On the other hand, the trace norm $\|X\|_1$ of a matrix X is defined as

$$\|X\|_1 = \max_{U:\text{unitary}} \mathrm{Tr}\, UX. \tag{A.12}$$

Choosing a unitary matrix U_X such that $X = U_X|X|$ (i.e., a polar decomposition), we obtain [Ex. A.8]

$$\|X\|_1 = \max_{U:\text{unitary}} \mathrm{Tr}\, UX = \mathrm{Tr}\, U_X^* X = \mathrm{Tr}\, |X|. \tag{A.13}$$

Hence, we also have

$$\|X^*\|_1 = \max_{U:\text{unitary}} \mathrm{Tr}\, U^*X^* = \mathrm{Tr}\, U_X X^* = \mathrm{Tr}\, |X^*|.$$

If X is Hermitian, then

$$\|X\|_1 = \max_{T:-I\le T\le I} \mathrm{Tr}\, XT = \mathrm{Tr}\, X(\{X \ge 0\} - \{X < 0\}) = \mathrm{Tr}\, X(I - 2\{X < 0\}). \tag{A.14}$$

Exercises

A.3. Show that the trace norm of a Hermitian matrix $\begin{pmatrix} -a & b \\ b^* & a \end{pmatrix}$ is equal to $2\sqrt{|b|^2 + a^2}$.

A.4. Show that

$$\|X\|_1 \ge \|\,\mathrm{Tr}_B\, X\|_1. \tag{A.15}$$

for a matrix X in $\mathcal{H}_A \otimes \mathcal{H}_B$.

A.5. Let A and B be square matrices of dimension d. Show that the eigenvalues of BA are the same as the eigenvalues of AB including degeneracies if A or B possesses the inverse.

A.6. Show that $\mathrm{spr}(AB) = \mathrm{spr}(BA)$.

A.7. Show that the function $t \mapsto t^{1/2}$ is a matrix monotone function following the steps below.

a Show that $\|A^{1/2}B^{-1/2}\| \leq 1$ when the Hermitian matrices B and A satisfy $B \geq A \geq 0$ and B possesses the inverse.
b Show that $1 \leq \mathrm{spr}(B^{-1/4}A^{1/2}B^{-1/4})$ under the same conditions as **a**.
c Show that

$$B^{1/2} \geq A^{1/2} \tag{A.16}$$

under the same conditions as **a**.
d Show that (A.16) holds even if B does not possess the inverse.

A.8. Prove (A.13) following the steps below.

a Show that $\max_{v:\|v\|=1}\langle v|\,|X|\,|u_i\rangle = \langle u_i|\,|X|\,|u_i\rangle$ for eigenvectors u_i of $|X|$ of length 1.
b Show that $\max_{U:\text{unitary}}\langle u_i|UX|u_i\rangle = \langle u_i|U_X^*X|u_i\rangle = \langle u_i||X||u_i\rangle$.
c Prove (A.13).

A.9. Show that $\|XY\|_1 \leq \|X\|\|Y\|_1$ for two matrices X and Y.

A.10. (*Poincaré inequality*) Let A be a $d \times d$ Hermitian matrix. Let a_i be the eigenvalues of A ordered from largest to smallest. Show that $\min_{x\in\mathcal{K},\|x\|=1}\langle x|A|x\rangle \leq a_k$ for any k-dimensional subspace $\mathcal{K}$.

A.11. Show that $\max_{P:\mathrm{rank}\,P=k}\min_x \dfrac{\langle x|PAP|x\rangle}{\langle x|P|x\rangle} = a_k$ under the same conditions as above.

A.12. Let A and B be Hermitian matrices, and let a_i and b_i be their ordered eigenvalues from largest to smallest. Show that $a_i \geq b_i$ if $A \geq B$.

A.4 Convex Functions and Matrix Convex Functions

Linear functions are often used in linear algebra. On the other hand, functions such as x^2 and $\exp(x)$ do not satisfy the linearity property. If we denote such functions by f, then they instead satisfy

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2), \quad 0 \leq \forall\lambda \leq 1, \forall x_1, x_2 \in \mathbb{R}.$$

A function is called a *convex function* when it satisfies the above inequality. If $-f$ is a convex function, then f is called a *concave function*. In the above, its domain is restricted to real numbers. However, this restriction is not necessary

and may be defined in a more general way. For example, for a vector space, we may define the *convex combination* $\lambda v_1 + (1-\lambda)v_2$ for two vectors v_1 and v_2 with $0 < \lambda < 1$. More generally, a set is called a *convex set* when the convex combination of any two elements is defined. Further, a convex set L is called a *convex cone* if $v \in L$ and $\lambda > 0$ imply $\lambda v \in L$. Therefore, it is possible to define convex and concave functions for functions with a vector space domain and a real number range. Similarly, convex and concave functions may be defined with a convex set domain. Examples of convex sets are the set of probability distributions and the set of density matrices. In particular, an element v of the convex set V is called an *extremal point* if $v_i \in V$ and $v = \lambda v_1 + (1-\lambda)v_2, (0 < \lambda < 1)$ imply $\lambda = 1$ or 0. For example, a pure state is an extremal point in the set of density matrices.

Lemma A.6 *Let f be a convex function on the convex set V. For any element v_0 of V, there exists a linear function g such that*

$$g(v_0) - f(v_0) = \max_{v \in V} g(v) - f(v).$$

When f is differentiable, g coincides with the derivative of f at v_0. Further, for any linear function g and a constant $C_0 \geq 0$, there exists the Lagrange multiplier λ such that

$$\max_{v \in V} f(v) + \lambda g(v) = \max_{v \in V : g(v) \leq C_0} f(v) + \lambda g(v).$$

In this case, λg coincides with the derivative of f at $\operatorname{argmax}_{v \in V : g(v) \leq C_0} f(v)$.

Lemma A.7 *Consider two vector spaces V_1 and V_2 and consider a real-valued function $f(v_1, v_2)$ with the domain $V_1 \times V_2$. If f is convex with respect to v_2 and concave with respect to v_1, then (Chap. VI Prop. 2.3 of [105])*[1]

$$\sup_{v_1 \in S_1} \min_{v_2 \in S_2} f(v_1, v_2) = \min_{v_2 \in S_2} \sup_{v_1 \in S_1} f(v_1, v_2),$$

where S_1 and S_2 are convex subsets of V_1 and V_2.

Next, we focus on the set of probability distributions on $\mathcal{S}(\mathcal{H})$ and denote it by $\mathcal{P}(\mathcal{S}(\mathcal{H}))$. In particular, we consider extremal points of the set $\mathcal{S}(\mathcal{H})$:

$$\mathcal{P}(\rho, \mathcal{S}(\mathcal{H})) \stackrel{\text{def}}{=} \left\{ p \in \mathcal{P}(\mathcal{S}(\mathcal{H})) \middle| \sum_i p_i \rho_i = \rho \right\}.$$

Such extremal points of the above set are characterized as follows.

Lemma A.8 *(Fujiwara and Nagaoka [128]) Let $p \in \mathcal{P}(\rho, \mathcal{S}(\mathcal{H}))$ be an extremal point and $\{\rho_1, \ldots, \rho_k\}$ be the support of p. Then, $\rho_1, \ldots, \rho_k$ are linearly independent. Hence, the number of supports of p is less than* $\dim \mathcal{T}(\mathcal{H}) = (\dim \mathcal{H})^2$.

[1] This relation holds even if V_1 is infinite dimensional, as long as S_2 is a closed and bounded set.

Note that we obtain the same result when we replace $\mathcal{P}(\rho, \mathcal{S}(\mathcal{H}))$ by $\mathcal{P}(\mathcal{S}(\mathcal{H}))$.
Proof. Assume that $\rho_1, \ldots, \rho_k$ are linearly dependent. That is, we choose real numbers $\lambda_1, \ldots, \lambda_k$ such that $\sum_{i=1}^k \lambda_i \rho_i = 0$ and $\sum_i \lambda_i = 0$. Define two distributions q^+ and q^- with the same support by

$$q_i^{\pm} \stackrel{\text{def}}{=} p_i \pm \epsilon \lambda_i. \tag{A.17}$$

Then, we have $p = \frac{1}{2}q^+ + \frac{1}{2}q^-$ and $q^+ \neq q^-$. It is a contradiction. ■

Indeed, applying this lemma to ρ_{mix}, we can see that any extremal POVM has at most $(\dim \mathcal{H})^2$ elements. Further, we focus on the cost functions $f_1, \ldots, f_l$ on $\mathcal{S}(\mathcal{H})$ and treat the following sets:

$$\mathcal{P}_{=(\leq)c}(\rho, f, \mathcal{S}(\mathcal{H})) \stackrel{\text{def}}{=} \left\{ p \in \mathcal{P}(\rho, \mathcal{S}(\mathcal{H})) \left| \sum_i p_i f_j(\rho_i) = (\leq) c \forall j = 1, \ldots, l \right. \right\}$$

$$\mathcal{P}_{=(\leq)c}(f, \mathcal{S}(\mathcal{H})) \stackrel{\text{def}}{=} \left\{ p \in \mathcal{P}(\mathcal{S}(\mathcal{H})) \left| \sum_i p_i f_j(\rho_i) = (\leq) c \forall j = 1, \ldots, l \right. \right\}.$$

Lemma A.9 *(Fujiwara and Nagaoka [128]) Let p be an extremal point of one of the above sets. Then, the number of supports of p is less than $(l + 1)(\dim \mathcal{H})^2$.*

The concept of "convex function" can be extended to functions of matrices. If a function f with the range $[0, \infty]$ satisfies

$$\lambda f(A) + (1 - \lambda) f(B) \geq f(\lambda A + (1 - \lambda) B),$$

for arbitrary Hermitian matrices A, B with eigenvalues in $[0, \infty]$, it is called a *matrix convex function*. See Sect. 1.5 for the definition of $f(A)$. Also, the function f is called a *matrix concave function* when the function $-f$ is a matrix convex function. The following equivalences are known [49]:

① $f(t)$ is matrix monotone.
② $t/f(t)$ is matrix monotone.
③ $f(t)$ is matrix concave.

Furthermore, it is known that if the function f satisfies one of the above conditions, $1/f(t)$ is matrix convex [49]. Hence, since the functions $t^s, -t^{-s} (s \in [0, 1])$, and $\log t$ are matrix monotone, the functions t^s $(s \in [-1, 0] \cup [1, 2])$, $-t^s$ $(s \in [0, 1])$, $-\log t$, and $t \log t$ are matrix convex functions.

Theorem A.1 *The following conditions are equivalent for a function f [49].*

① *$f(t)$ is matrix convex.*
② *When a matrix C satisfies $C^*C = I$, any Hermitian matrix A with eigenvalues in $[0, \infty]$ satisfies $f(C^*AC) \leq C^* f(A) C$.*

③ *When matrices* $C_1, \ldots, C_k$ *satisfy* $\sum_i C_i^* C_i = I$*, any Hermitian matrices* $A_1, \ldots, A_k$ *with eigenvalues in* $[0, \infty]$ *satisfy* $f(\sum_i C_i^* A_i C_i) \le \sum_i C_i^* f(A_i) C_i$.

Now, we prove important inequalities.

Proof of (6.11), (5.40), and (5.41). In what follows, we focus on the linear space $\mathcal{M}_A$ of matrices on $\mathcal{H}_A$ and the linear space $\mathcal{M}_A$ of matrices on $\mathcal{H}_A$. For a given density ρ on $\mathcal{H}_A$, we also define the map L_ρ and R_ρ as a linear map on the matrix space $\mathcal{M}_A$:

$$L_\rho(A) \stackrel{\text{def}}{=} \rho A, \quad R_\rho(A) \stackrel{\text{def}}{=} A\rho.$$

The map L_ρ is positive Hermitian under the inner product $\langle Y, X \rangle = \operatorname{Tr} Y^* X$ because

$$\begin{aligned} \langle Y, L_\rho X \rangle_{\rho,r}^{(e)} &= \operatorname{Tr} Y^* \rho \rho X = \operatorname{Tr} (\rho Y)^* \rho X = \langle L_\rho Y, X \rangle_{\rho,r}^{(e)}, \\ \langle X, L_\rho X \rangle &= \operatorname{Tr} X^* \rho^2 X \ge 0. \end{aligned}$$

For another state σ, the map R_σ is positive Hermitian under the inner product $\langle Y, X \rangle_{\rho,r}^{(e)}$. It is checked as follows:

$$\begin{aligned} \langle Y, R_\sigma X \rangle_{\rho,r}^{(e)} &= \operatorname{Tr} Y^* \rho X \sigma = \operatorname{Tr} (Y\sigma)^* \rho X = \langle R_\sigma Y, X \rangle_{\rho,r}^{(e)}, \\ \langle X, R_\sigma X \rangle_{\rho,r}^{(e)} &= \operatorname{Tr} X^* \rho X \sigma \ge 0. \end{aligned}$$

Since L_ρ and R_σ are commutative, $L_\rho^{-1} R_\sigma$ is also positive Hermitian. Now, for any state $\rho^{A,B}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$, and the TP-CP map $\kappa = \operatorname{Tr}_B$, we focus on the map $\kappa_{\rho^{A,B},r}$ from $\mathcal{M}_{A,B}$ to $\mathcal{H}_A$ (this map is defined in Sect. 6.1). The dual map $\kappa_{\rho^{A,B},r}^*$ is written as $\kappa_{\rho^{A,B},r}^*(Y) = Y \otimes I$ because

$$\begin{aligned} &\langle \kappa_{\rho^{A,B},r}^*(Y), X \rangle_{\rho^{A,B},r}^{(e)} = \langle Y, \kappa_{\rho^{A,B},r}(X) \rangle_{\operatorname{Tr}_B \rho^{A,B},r}^{(e)} \\ =& \operatorname{Tr}_A Y^* (\operatorname{Tr}_B \rho^{A,B}) \kappa_{\rho^{A,B},r}(X) = \operatorname{Tr}_A Y^* \operatorname{Tr}_B (\rho^{A,B} X) \\ =& \operatorname{Tr}_{A,B} Y^* \otimes I_B \rho^{A,B} X = \langle Y \otimes I_B, X \rangle_{\rho^{A,B},r}^{(e)}. \end{aligned}$$

Hence, $\kappa_{\rho^{A,B},r} \kappa_{\rho^{A,B},r}^*(Y) = Y$. Now, let f be a matrix convex function. Applying Condition ② in Theorem A.1, we obtain

$$\begin{aligned} &\operatorname{Tr} X^* f(\kappa_{\rho^{A,B},r} L_{\rho^{A,B}}^{-1} R_{\sigma^{A,B}} \kappa_{\rho^{A,B},r}^*)(\operatorname{Tr}_B \rho^{A,B}) X \\ &\quad = \langle X, f(\kappa_{\rho^{A,B},r} L_{\rho^{A,B}}^{-1} R_{\sigma^{A,B}} \kappa_{\rho^{A,B},r}^*) X \rangle_{\operatorname{Tr}_B \rho^{A,B},r}^{(e)} \\ &\quad \le \langle X, \kappa_{\rho^{A,B},r} f(L_{\rho^{A,B}}^{-1} R_{\sigma^{A,B}}) \kappa_{\rho^{A,B},r}^* X \rangle_{\operatorname{Tr}_B \rho^{A,B},r}^{(e)} \\ &\quad = \operatorname{Tr} (X \otimes I)^* f(L_{\rho^{A,B}}^{-1} R_{\sigma^{A,B}}) \rho^{A,B} (X \otimes I). \end{aligned}$$

Thus, substituting $\rho^{A,B}$ $(-x^\lambda)$ into $\sigma^{A,B}$ $(f(x))$, we have

$$\begin{aligned} &- \operatorname{Tr}_A X^* (\operatorname{Tr}_B \rho^{A,B})^{1-\lambda} X (\operatorname{Tr}_B \rho^{A,B})^{\lambda} \\ \le &- \operatorname{Tr}_{A,B} (X \otimes I_B)^* (\rho^{A,B})^{1-\lambda} (X \otimes I_B) (\rho^{A,B})^{\lambda} \end{aligned}$$

because $f(L^{-1}_{\rho^{A,B}} R_{\sigma^{A,B}}) X = (\rho^{A,B})^{-\lambda} X (\sigma^{A,B})^{\lambda}$. Hence, we obtain (6.11). Further, substituting X $(-x^{\lambda})$ into I $(f(x))$, we obtain

$$- \operatorname{Tr}_A (\operatorname{Tr}_B \rho^{A,B})^{1-\lambda} (\operatorname{Tr}_B \sigma^{A,B})^{\lambda} \le - \operatorname{Tr}_{A,B} (\rho^{A,B})^{1-\lambda} (\sigma^{A,B})^{\lambda}$$

for $\le \lambda \le 1$, which implies (5.40) in the partial trace case. From the Stinespring representation we obtain (5.40) in the general case. Also, substituting X (x^{λ}) into I $(f(x))$, we obtain

$$\operatorname{Tr}_A (\operatorname{Tr}_B \rho^{A,B})^{1-\lambda} (\operatorname{Tr}_B \sigma^{A,B})^{\lambda} \le \operatorname{Tr}_{A,B} (\rho^{A,B})^{1-\lambda} (\sigma^{A,B})^{\lambda}, \quad -1 \le \lambda \le 0,$$

which implies (5.41) in the partial trace case. Therefore, we obtain (5.41) in the general case. ■

Exercises

A.13. Show the concavity of the von Neumann entropy (5.49) using the matrix convexity of $x \log x$.

A.14. Show the joint convexity of (5.31) using the matrix convexity of $-\log x$ and $x \log x$.

A.5 Proof and Construction of Stinespring and Choi–Kraus Representations

In this section, we will prove Theorem 5.1 and construct the Stinespring and Choi–Kraus representations. First, let us consider the following theorem for completely positive maps, without the trace-preserving condition.

Theorem A.2 *Given a linear map κ from the set of Hermitian matrices on the d-dimensional system $\mathcal{H}_A$ to that on the d'-dimensional system $\mathcal{H}_B$, the following conditions are equivalent.*

① *κ is a completely positive map.*
② *κ^* is a completely positive map.*
③ *κ is a $\min\{d, d'\}$-positive map.*
④ *The matrix $K(\kappa)$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ is positive semidefinite.*
⑤ *(Stinespring representation) There exist a Hilbert space $\mathcal{H}_C$ with the same dimension as $\mathcal{H}_B$, a pure state $\rho_0 \in \mathcal{S}(\mathcal{H}_B \otimes \mathcal{H}_C)$, and a matrix W in $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ such that $\kappa(X) = \operatorname{Tr}_{A,C} W(X \otimes \rho_0) W^*$.*
⑥ *(Choi–Kraus representation) There exist dd' linear maps $F_1, \dots, F_{dd'}$ from $\mathcal{H}_A$ to $\mathcal{H}_B$ such that $\kappa(X) = \sum_i F_i X F_i^*$.*

We also define Conditions ⑤′ and ⑥′ by deforming Conditions ⑤ and ⑥ as:

⑤′ *There exist Hilbert spaces $\mathcal{H}_C$ and $\mathcal{H}'_C$, a positive semidefinite state $\rho'_0 \in \mathcal{S}(\mathcal{H}_C)$, and a linear map W from $\mathcal{H}_A\mathcal{H}_C$ to $\mathcal{H}_A\mathcal{H}'_C$ such that $\kappa(X) = \mathrm{Tr}_{C'} W(X \otimes \rho'_0)W^*$.*

⑥′ *There exist linear maps $F_1, \ldots, F_k$ from $\mathcal{H}_A$ to $\mathcal{H}_B$ such that $\kappa(X) = \sum_i F_i X F_i^*$.*

Proof. We now show that ②⇔ ①⇒③⇒④⇒ ⑤⇒⑥⇒⑥′⇒ ① and ⑤⇒⑤′⇒①. Since ①⇒③, ⑤⇒⑤′, and ⑥⇒⑥′ by inspection, we prove the remaining relations.

We first prove ①⇔②. The n-positivity of κ is equivalent to

$$\mathrm{Tr}\, \kappa \otimes \iota_n(X) \geq 0 \tag{A.18}$$

for any positive semidefinite Hermitian matrix X on $\mathcal{H}_A \otimes \mathbb{C}^n$, which is equivalent to

$$\mathrm{Tr}\, \kappa \otimes \iota_n(X)Y \geq 0$$

for any positive semidefinite Hermitian matrix X (Y) on $\mathcal{H}_A \otimes \mathbb{C}^n$ ($\mathcal{H}_B \otimes \mathbb{C}^n$). Since $(\kappa \otimes \iota_n)^* = \kappa^* \otimes \iota_n$, we have $\mathrm{Tr}\, \kappa \otimes \iota_n(X)Y = \mathrm{Tr}\, X\kappa^* \otimes \iota_n(Y)$. Therefore, the n-positivity of κ is equivalent to the n-positivity of κ^*. Hence, the complete positivity of κ is equivalent to the complete positivity of κ^*.

Next, we derive ③⇒④. We show ④ for $d' \leq d$. Since κ is a d-positive map, $\kappa \otimes \iota_B$ is a positive map (ι_B is the identity map in $\mathcal{T}(\mathcal{H}_B)$). Let X be a positive semidefinite Hermitian matrix on $\mathcal{H}_A \otimes \mathcal{H}_B$. Assume that $X = \sum_{i,k,j,l,} x^{(i,k),(j,l)} |e_i^A \otimes e_k^B\rangle\langle e_j^A \otimes e_l^B|$. Since $(\kappa \otimes \iota_B)(X) \geq 0$, we have

$$\begin{aligned} 0 &\leq \langle I_B|(\kappa \otimes \iota_B)(X)|I_B\rangle \\ &= \sum_{i,j,k,l} x^{(i,k),(j,l)} \langle I_B| \left(\kappa(|e_i^A\rangle\langle e_j^A|) \otimes |e_k^B\rangle\langle e_l^B|\right) |I_B\rangle \\ &= \sum_{i,j,k,l} x^{(i,k),(j,l)} \langle e_k^B|\kappa(|e_i^A\rangle\langle e_j^A|)|e_l^B\rangle \\ &= \sum_{i,j,k,l} x^{(i,k),(j,l)} K(\kappa)^{(j,l),(i,k)} = \mathrm{Tr}\, XK(\kappa). \end{aligned} \tag{A.19}$$

Therefore, $K(\kappa) \geq 0$, and we obtain ④. In the above, we denote the vector $\sum_{k=1}^{d'} e_k^B \otimes e_k^B$ in the space $\mathcal{H}_B \otimes \mathcal{H}_B$ by I_B. In the derivation of (A.19), we used the fact that

$$\langle I_B| \left(|e_k^B\rangle\langle e_l^B| \otimes |e_s^B\rangle\langle e_t^B|\right) |I_B\rangle = \langle I_B| \left(|e_k^B \otimes e_s^B\rangle\langle e_l^B \otimes e_t^B|\right) |I_B\rangle = \delta_{k,s}\delta_{l,t}.$$

Using a discussion in the proof of ①⇔②, we can show that Condition ③ implies that κ^* is a d-positive map if $d \leq d'$. From this fact we can derive

Condition ④ for κ^*. This gives us $K(\kappa^*) \geq 0$, which is equivalent to $K(\kappa) \geq 0$. Thus, we obtain ④ for κ.

We now derive ④⇒⑤. Since $K(\kappa) \geq 0$, $\sqrt{K(\kappa)}$ exists. In what follows, we consider a space $\mathcal{H}_C$ with a basis $e_1^C, \ldots, e_{d'}^C$. Note that the space $\mathcal{H}_C$ is isometric to the space $\mathcal{H}_B$. Defining $U_{C,B} \stackrel{\text{def}}{=} \sum_{k=1}^{d'} e_k^C \otimes e_k^B$, we have

$$\begin{aligned}&\operatorname{Tr} |e_{i'}^A\rangle\langle e_{j'}^A| \otimes (|U_{C,B}\rangle\langle U_{C,B}|) \, |e_j^A \otimes e^C{}_k \otimes e_s^B\rangle\langle e_i^A \otimes e^C{}_l \otimes e_t^B| \\ =&\delta_{j',j}\delta_{i',i}\delta_{l,t}\delta_{k,s},\end{aligned} \tag{A.20}$$

where the order of the tensor product is $\mathcal{H}_A \otimes \mathcal{H}_C \otimes \mathcal{H}_B$. Although $K(\kappa)$ is originally a Hermitian matrix on $\mathcal{H}_A \otimes \mathcal{H}_B$, we often regard it as a Hermitian matrix on $\mathcal{H}_A \otimes \mathcal{H}_C$ because $\mathcal{H}_C$ is isometric to $\mathcal{H}_B$. Using (A.20), we have

$$\begin{aligned}\operatorname{Tr} \kappa(X)Y &= \operatorname{Tr}(X \otimes Y)K(\kappa) = \operatorname{Tr} \left(X \otimes |U_{C,B}\rangle\langle U_{C,B}|\right) K(\kappa) \otimes Y \\ &= \operatorname{Tr} \left(X \otimes |U_{C,B}\rangle\langle U_{C,B}|\right) \left(\sqrt{K(\kappa)} \otimes I_B\right) (I_{A,C} \otimes Y) \left(\sqrt{K(\kappa)} \otimes I_B\right) \\ &= \operatorname{Tr}_B \operatorname{Tr}_{A,C} \left(\left(\sqrt{K(\kappa)} \otimes I_B\right) (X \otimes |U_{C,B}\rangle\langle U_{C,B}|) \left(\sqrt{K(\kappa)} \otimes I_B\right)\right) Y\end{aligned}$$

for $\forall X \in \mathcal{T}(\mathcal{H}_A), \forall Y \in \mathcal{T}(\mathcal{H}_B)$. Therefore, we can show that

$$\kappa(X) = \operatorname{Tr}_{A,C} \left[\left(\sqrt{d'K(\kappa)} \otimes I_B\right) \left(X \otimes \frac{|U_{C,B}\rangle\langle U_{C,B}|}{d'}\right) \left(\sqrt{d'K(\kappa)} \otimes I_B\right)\right]. \tag{A.21}$$

Letting $\rho_0 = \frac{|U_{C,B}\rangle\langle U_{C,B}|}{d'}$ and $W = \sqrt{d'K(\kappa)} \otimes I_B$, we obtain ⑤.

Next, we show that ⑤⇒⑥. Let ρ_0 be $|x\rangle\langle x|$, P be a projection from $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ to $\mathcal{H}_A \otimes |x\rangle$, and $P_{i,k}$ be a projection from $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ to $\mathcal{H}_B \otimes |e_i^A \otimes e^C{}_k\rangle$. Using formula (1.26) of the partial trace, we have

$$\begin{aligned}\kappa(X) &= \operatorname{Tr}_{A,C} W(X \otimes \rho_0)W^* = \sum_{i=1}^{d} \sum_{k=1}^{d'} P_{i,k} W P X P W^* P_{i,k} \\ &= \sum_{i=1}^{d} \sum_{k=1}^{d'} (P_{i,k} W P) X (P_{i,k} W P)^*.\end{aligned}$$

We thus obtain ⑥.

Finally, we show ⑤′⇒①. From Condition ⑤′ any positive semidefinite Hermitian matrix X on $\mathcal{H}_A \otimes \mathbb{C}^n$ satisfies

$$\kappa \otimes \iota_n(X) = \operatorname{Tr}_{\mathcal{H}_A \otimes \mathcal{H}_C} (W \otimes I_n)(X \otimes \rho_0')(W^* \otimes I_n) \geq 0,$$

where I_n is an identity matrix in $\mathbb{C}^n$. From condition (A.18), therefore, κ is an n-positive map for arbitrary n. It follows that κ is a completely positive map from which we obtain ①.

Concerning a proof of ⑥′ ⇒ ①, we have

$$\kappa \otimes \iota_n(X) = \sum_i (F_i \otimes I_n) X (F_i^* \otimes I_n) \geq 0,$$

for a semipositive definite Hermitian matrix X on $\mathcal{H}_A \otimes \mathbb{C}^n$. Thus, we obtain ①. ■

Next, we prove Theorem 5.1. Thanks to Theorem A.2, it is sufficient to show the equivalence of Conditions ① to ⑥′ in Theorem 5.1 when κ is a completely positive map. Indeed, ① ⇒ ③, ⑤ ⇒ ⑤′, and ⑥ ⇒ ⑥′ by inspection. Concerning ⑤′ ⇒ ① and ⑥′ ⇒ ①, it is sufficient to show the trace-preserving property because of Theorem A.2. Therefore, we only show ③ ⇒ ④ ⇒ ⑤ ⇒ ⑥ as follows.

We first show ③ ⇒ ④. From definition (1.23) of the partial trace we obtain

$$\mathrm{Tr}_A \rho = \mathrm{Tr}_B \kappa(\rho) = \mathrm{Tr}_{A,B}(\rho \otimes I_B) K(\kappa) = \mathrm{Tr}_A \rho \, (\mathrm{Tr}_B K(\kappa))$$

for arbitrary $\rho \in \mathcal{S}(\mathcal{H}_A)$. Hence, $\mathrm{Tr}_B K(\kappa) = I_A$, and thus we obtain ④.

Next, we show ④ ⇒ ⑤. Employing the notation used in the proof of ④ ⇒ ⑤ in Theorem A.2, we let P be the projection from $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ to $\mathcal{H}_A \otimes |U_{C,B}\rangle$. Since any $\rho \in \mathcal{S}(\mathcal{H}_A)$ satisfies

$$\begin{aligned} \mathrm{Tr}\, \rho &= \mathrm{Tr}_B \, \mathrm{Tr}_{A,C} \left(\left(\sqrt{d' K(\kappa)} \otimes I_B \right) P \rho P \left(\sqrt{d' K(\kappa)} \otimes I_B \right) \right) \\ &= \mathrm{Tr}_{A,C,B} \left(\left(\sqrt{d' K(\kappa)} \otimes I_B \right) P \rho P \left(\sqrt{d' K(\kappa)} \otimes I_B \right) \right), \end{aligned}$$

we obtain

$$\mathrm{Tr}_{A,C,B} \left(\left(\sqrt{d' K(\kappa)} \otimes I_B \right) P \right)^* \left(\sqrt{d' K(\kappa)} \otimes I_B \right) P = P.$$

Let $\mathcal{H}_R$ be the range of $\left(\sqrt{d' K(\kappa)} \otimes I_B \right) P$ for $\mathcal{H}_A \otimes |U_{C,B}\rangle$. Then, the dimension of $\mathcal{H}_R$ is equal to that of $\mathcal{H}_A$. $\left(\sqrt{d' K(\kappa)} \otimes I_B \right) P$ can be regarded as a map from $\mathcal{H}_A \otimes |U_{C,B}\rangle$ to $\mathcal{H}_R$.

Let ${\mathcal{H}_R}^\perp$ be the orthogonal complementary space of $\mathcal{H}_R$ in $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, and ${\mathcal{H}_A}^\perp$ be the orthogonal complementary space of $\mathcal{H}_A \otimes |U_{C,B}\rangle$. Since the dimension of ${\mathcal{H}_R}^\perp$ is equal to that of ${\mathcal{H}_A}^\perp$, there exists a unitary (i.e., metric-preserving) linear mapping U' from ${\mathcal{H}_R}^\perp$ to ${\mathcal{H}_A}^\perp$. Then, $U_\kappa \stackrel{\mathrm{def}}{=} \left(\sqrt{d' K(\kappa)} \otimes I_B \right) P \oplus U'$ is a unitary linear map from $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C = (\mathcal{H}_A \otimes |U_{C,B}\rangle) \oplus {\mathcal{H}_A}^\perp$ to $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C = \mathcal{H}_R \oplus {\mathcal{H}_R}^\perp$. Therefore, from (A.21) we have $\kappa(\rho) = \mathrm{Tr}_{A,C} U_\kappa \rho \otimes \frac{|U_{C,B}\rangle\langle U_{C,B}|}{d'} U_\kappa$, which gives Condition ⑤.

Next, we show ⑤ ⇒ ⑥ by employing the notation used in the proof of ⑤ ⇒ ⑥ in Theorem A.2. Since

$$\operatorname{Tr}\rho = \operatorname{Tr}\kappa(\rho) = \operatorname{Tr}_B \operatorname{Tr}_{A,C} U_\kappa(\rho\otimes\rho_0)U_\kappa^* = \sum_{i=1}^{d}\sum_{k=1}^{d'} \operatorname{Tr}_B P_{i,k}WP\rho PW^*P_{i,k}$$

$$= \sum_{i=1}^{d}\sum_{k=1}^{d'} \operatorname{Tr}_B (P_{i,k}WP)\rho(P_{i,k}WP)^* = \operatorname{Tr}_A \sum_{i=1}^{d}\sum_{k=1}^{d'} (P_{i,k}WP)^*(P_{i,k}WP)\rho,$$

we obtain $\sum_{i=1}^{d}\sum_{k=1}^{d'}(P_{i,k}WP)^*(P_{i,k}WP) = I_A$. Therefore, we obtain ⑥. Further, from the proof ⑤ ⇒ ⑥, we obtain (5.1).

Finally, we directly construct Stinespring representation ⑤′ from Choi–Kraus representation ⑥′. Define the map W from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathbb{C}^k$ as

$$W(x) \stackrel{\text{def}}{=} \sum_{i=1}^{k} F_i(x)\otimes e_i.$$

Then, W satisfies

$$\operatorname{Tr}_{\mathbb{C}^k} W\rho W^* = \sum_{i=1}^{k} F_i\rho F_i^*.$$

We obtain Condition ⑤′ from ⑥′ in Theorem A.2. In Theorem 5.1, we have to check the unitarity. From the condition $\sum_{i=1}^{k} F_i^*F_i = I$, we obtain $W^*W = I$, i.e., W is an isometry map. Hence, it is possible to deform map W to a unitary map by extending the input space. In this case, the state in the environment $\kappa^E(\rho)$ equals $\operatorname{Tr}_B W\rho W^* = (\operatorname{Tr} F_j^*F_i\rho)_{i,j}$. Thus, we obtain Lemma 5.1.

B

Proofs of Theorems and Lemmas

B.1 Proof of Theorem 3.1

In this proof, we only consider the case in which there exists an element $x \in L$ such that $A(x) = b$.[1] Otherwise, since both sides are equal to $-\infty$, the theorem holds.

When $x \in L$ satisfies $A(x) = b$ and y satisfies $A^T(y) - c \in L^T$, we have $0 \leq \langle A^T(y) - c, x \rangle = \langle y, A(y) \rangle - \langle c, x \rangle = \langle y, b \rangle - \langle c, x \rangle$. Hence, we can check that

$$\max_{x \in V_1} \{\langle c, x \rangle | x \in L, A(x) = b\} \leq \min_{y \in V_2} \{\langle y, b \rangle | A^T(y) - c \in L^T\}. \tag{B.1}$$

Furthermore,

$$\begin{aligned} &\min_{y \in V_2} \{\langle y, b \rangle | A^T(y) - c \in L^T\} \\ = &\min_{(\mu, y) \in \mathbb{R} \times V_2} \{\mu | \exists y \in V_2, \forall x \in L, \langle y, b \rangle - \langle A^T(y) - c, x \rangle \leq \mu\}. \end{aligned}$$

This equation can be checked as follows. When $y \in V_2$ satisfies $A^T(y) - c \in L^T$, the real number $\mu = \langle y, b \rangle$ satisfies the condition on the right-hand side (RHS). Hence, we obtain the $\geq$ part. Next, we consider a pair (μ, y) satisfying the condition on the RHS. Then, we can show that $\langle A^T(y) - c, x \rangle$ is greater than zero for all $x \in L$, by reduction to absurdity. Assume that there exists an element $x \in L$ such that $\langle A^T(y) - c, x \rangle$ is negative. By choosing a sufficiently large number $t > 0$, $tx \in L$, but $\langle y, b \rangle - \langle A^T(y) - c, tx \rangle \leq \mu$ does not hold. It is a contradiction. This proves the $\leq$ part.

Let $\eta_0 \stackrel{\text{def}}{=} \max_{x \in V_1} \{\langle c, x \rangle | x \in L, A(x) = b\}$. Then $(\eta_0, 0)$ is a point that lies on the boundary of the convex set $\{(\langle c, x \rangle, A(x) - b)\}_{x \in L} \subset \mathbb{R} \times V_2$. Choosing an appropriate $y_0 \in V_2$ and noting that $(1, -y_0) \in \mathbb{R} \times V_2$, we have

[1] Our proof follows [397].

$$\eta_0 = \eta_0 - \langle y, 0\rangle \geq \langle c, x\rangle - \langle y_0, A(x) - b\rangle, \quad \forall x \in L.$$

From this fact we have

$$\eta_0 \geq \min_{(\mu,y)\in\mathbb{R}\times V_2} \{\mu | \exists y \in V_2, \forall x \in L, \langle y, b\rangle - \langle A^T(y) - c, x\rangle \leq \mu\}.$$

This proves the reverse inequality of (B.1) and completes the proof.

B.2 Proof of Theorem 8.2

We prove this theorem in the following steps: ①⇒② ⇒③ ⇒①, ②⇒④ ⇒①. The proof given here follows from Bhatia [49].

We first show ①⇒② for dimension d by induction. Let $t \stackrel{\text{def}}{=} (y_1 - x_1)/(y_1 - y_2) = (x_2 - y_2)/(y_1 - y_2)$ for $d = 2$. Since $x \preceq y$, we have $0 \leq t \leq 1$. Further, the relation

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1-t & t \\ t & 1-t \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{B.2}$$

proves the case for $d = 2$. In the following proof, assuming that the result holds for $d \leq n-1$, we prove the case for $d = n$. Any permutation is expressed by by a product of T transforms. Hence, it is sufficient to show ② when $x_1 \geq x_2 \geq \ldots \geq x_n$ and $y_1 \geq y_2 \geq \ldots \geq y_n$. Since $x \preceq y$, we have $y_n \leq x_1 \leq y_1$. Choosing an appropriate k, we have $y_k \leq x_1 \leq y_{k-1}$. When t satisfies $x_1 = ty_1 + (1-t)y_k$, the relation $0 \leq t \leq 1$ holds. Let T_1 be the T transform among the first and kth elements defined by t. Define

$$x' \stackrel{\text{def}}{=} (x_2, \ldots, x_n)^T, \tag{B.3}$$

$$y' \stackrel{\text{def}}{=} (y_2, \ldots, y_{k-1}, (1-t)y_1 + ty_k, y_{k+1}, \ldots, y_n)^T. \tag{B.4}$$

Then, $T_1 y = (x_1, y')$. Since $x' \preceq y'$ (as shown below), from the assumptions of the induction there exist T transforms $T_f, \ldots, T_2$ such that $T_f \cdots T_2 y' = x'$. Therefore, $T_f \cdots T_2 T_1 y = T_f \cdots T_2(x_1, y') = (x_1, x') = x$, which completes the proof for this part. We now show that $x' \preceq y'$. For an integer m satisfying $2 \leq m \leq k-1$, we have

$$\sum_{j=2}^{m} x_j \leq \sum_{j=2}^{m} y_j. \tag{B.5}$$

If $k \leq m \leq n$, then

$$\begin{aligned}\sum_{j=2}^{m} {y'}_j &= \left(\sum_{j=2}^{k-1} y_j\right) + (1-t)y_1 + ty_k + \left(\sum_{j=k+1}^{m} y_j\right) \\ &= \left(\sum_{j=1}^{m} y_j\right) - ty_1 + (t-1)y_k = \sum_{j=1}^{m} y_j - x_1 \geq \sum_{j=1}^{m} x_j - x_1 = \sum_{j=2}^{m} x_j,\end{aligned}$$

which shows that $x' \preceq y'$.

Next, we show ②⇒③. The product of two double stochastic transition matrices A_1 and A_2 is also a double stochastic transition matrix A_1A_2. Since a T transform is a double stochastic transition matrix, we obtain ③.

For proving ③⇒①, it is sufficient to show that

$$\sum_{t=1}^{k}\sum_{j=1}^{d} x^{i_t,j}a_j \le \sum_{j=1}^{k} a_j^{\downarrow}$$

for an arbitrary integer k and a set of k arbitrary integers $i_1,\dots,i_k$ from 1 to d. This can be shown from the fact that $\sum_{j=1}^{d}\sum_{t=1}^{k} x^{i_t,j} = k$ and $\sum_{t=1}^{k} x^{i_t,j} \le 1$ for each j.

We now show ②⇒④. For simplicity, we consider $d=2$ and let

$$B = \begin{pmatrix} \frac{(y_1/y_2)^2-(x_2/x_1)(y_1/y_2)}{(y_1/y_2)^2-1} & \frac{(y_1/y_2)(x_1/x_2)-1}{(y_1/y_2)^2-1} \\ \frac{(x_2/x_1)(y_1/y_2)-1}{(y_1/y_2)^2-1} & \frac{(y_1/y_2)^2-(y_1/y_2)(x_1/x_2)}{(y_1/y_2)^2-1} \end{pmatrix}. \tag{B.6}$$

It can be verified that this is a stochastic transition matrix. Since

$$\begin{pmatrix} \frac{(y_1/y_2)^2-(x_2/x_1)(y_1/y_2)}{(y_1/y_2)^2-1}x_1 \\ \frac{(y_1/y_2)(x_1/x_2)-1}{(y_1/y_2)^2-1}x_2 \end{pmatrix} = \frac{(y_1/y_2)x_1 - x_2}{(y_1/y_2)^2-1}\begin{pmatrix} (y_1/y_2) \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} \frac{(x_2/x_1)(y_1/y_2)-1}{(y_1/y_2)^2-1}x_1 \\ \frac{(y_1/y_2)^2-(y_1/y_2)(x_1/x_2)}{(y_1/y_2)^2-1}x_2 \end{pmatrix} = \frac{(y_1/y_2)x_2 - x_1}{(y_1/y_2)^2-1}\begin{pmatrix} 1 \\ (y_1/y_2) \end{pmatrix},$$

we observe that $B^1 \circ x \approx B^2 \circ x \approx y$.

Let T_0 be a T transform defined with respect to t between kth and lth elements ($k<l$), and define B^1 and B^2 as

$$\begin{pmatrix} b^{1,k} & b^{1,l} \\ b^{2,k} & b^{2,l} \end{pmatrix} = \begin{pmatrix} \frac{(y_k/y_l)^2-(x_l/x_k)(y_k/y_l)}{(y_k/y_l)^2-1} & \frac{(y_k/y_l)(x_k/x_l)-1}{(y_k/y_l)^2-1} \\ \frac{(x_l/x_k)(y_k/y_l)-1}{(y_k/y_l)^2-1} & \frac{(y_k/y_l)^2-(y_k/y_l)(x_k/x_l)}{(y_k/y_l)^2-1} \end{pmatrix}$$

$$b^{1,i} = \frac{(y_k/y_l)x_k - x_l}{(y_k/y_l)-1}, \quad b^{2,i} = \frac{(y_k/y_l)x_l - x_k}{(y_k/y_l)-1} \quad \text{if } i \ne k,l.$$

Then, $B^1 \circ x \approx B^2 \circ x \approx y$, if $x = T_0 y$.

Further, if two stochastic transition matrices B, C satisfy $y \approx (B^j)^* \circ x$ and $z \approx (C^i)^* \circ y$ for arbitrary integers i and j, then there exists an appropriate substitution $s(j)$ such that

$$y \propto s(j)((B^j)^* \circ x),$$

where we identify the permutation $s(j)$ and the matrix that represents it. Since $s(j)^* = (s(j))^{-1}$,

$$\begin{aligned} z &\approx (C^i)^* \circ y = s(j)\left(((s(j))^{-1}(C^i)^*) \circ (s(j))^{-1}y\right) \\ &\propto s(j)\left(((s(j))^{-1}(C^i)^*) \circ (B^j)^* \circ x\right) = s(j)\left((C^i s(j))^* \circ (B^j)^* \circ x\right) \\ &= s(j)\left((C^i s(j))^* \circ (B^j)^* \circ x\right) \approx (C^i s(j))^* \circ (B^j)^* \circ x. \end{aligned}$$

Therefore,

$$\sum_{i,j}(C^i s(j))^* \circ (B^j)^* = \sum_j (\sum_i C^i s(j))^* \circ (B^j)^*$$
$$= \sum_j (e^* s(j))^* \circ (B^j)^* = \sum_j e \circ (B^j)^* = \sum_j (B^j)^* = e.$$

When we define the matrix D by $(D^{i,j})^* \stackrel{\text{def}}{=} (C^i s(j))^* \circ (B^j)^*$ (note that the pair i, j refers to one column), this matrix is a stochastic transition matrix and satisfies

$$(D^{i,j})^* \circ x = z. \tag{B.7}$$

Using this and the previous facts, we obtain ②⇒④.

Finally, we show ④⇒①. It is sufficient to show the existence of a d-dimensional vector $c = (c_i)$ with positive real elements such that

$$c_i \le 1, \quad \sum_{i=1}^d c_i = k, \quad \sum_{j=1}^k y_{i_j} \ge \sum_{j=1}^k x_j^{\downarrow} \tag{B.8}$$

for arbitrary k. For this purpose, we choose k different integers $i_1, \ldots, i_k$ such that

$$\sum_{j=1}^k x_j^{\downarrow} = \sum_{j=1}^k x_{i_j}.$$

For each j, we choose the permutation $s(j)$ and the positive real number d_j such that $(B^j)^* \circ x = d_j s(j) y$. Note that $\sum_{j=1}^d d_j = 1$. Since

$$\sum_{j=1}^d b^{j,i} x_i = x_i,$$

we have

$$\sum_{j=1}^k x_{i_j} = \sum_{t=1}^d \sum_{j=1}^k b^{t,i_j} x_{i_j} = \sum_{t=1}^d \sum_{j=1}^k d_t (s(t)y)_{i_j} = \sum_{t=1}^d \sum_{j=1}^k \sum_{l=1}^d d_t s(t)_{i_j,l} y_l.$$

Since

$$\sum_{j=1}^k s(t)_{i_j,l} \le 1, \quad \sum_{l=1}^d \sum_{j=1}^k s(t)_{i_j,l} = k,$$

we obtain

$$\sum_t \sum_{j=1}^k d_t s(t)_{i_j,l} \le 1, \quad \sum_{l=1}^d \sum_t \sum_{j=1}^k s(t)_{i_j,l} = k$$

where we used $\sum_t d_t = 1$. This shows the existence of a vector $c = (c_i)$ satisfying (B.8).

B.3 Proof of Theorem 8.3

Let ρ be a separable state on $\mathcal{H}_A \otimes \mathcal{H}_B$. We can choose an appropriate set of vectors $\{u_i\}_i$ in $\mathcal{H}_A$ and $\{v_i\}_i$ in $\mathcal{H}_B$ such that $\rho = \sum_i |u_i \otimes v_i\rangle\langle u_i \otimes v_i| = \sum_j \lambda_j |e_j\rangle\langle e_j|$, where the RHS is the diagonalized form of ρ. From Lemma A.5 we can take an isometric matrix $W = (w_{i,j})$ such that $u_i \otimes v_i = \sum_j w_{i,j}\sqrt{\lambda_j} e_j$. Since $W^*W = I$, we have

$$\sum_i w_{i,j}^* u_i \otimes v_i = \sqrt{\lambda_j} e_j. \tag{B.9}$$

Similarly, we diagonalize $\mathrm{Tr}_B \rho$ such that $\mathrm{Tr}_B \rho = \sum_k \lambda'_k |f_k\rangle\langle f_k|$. Then, we can take an isometric matrix $W' = (w'_{i,k})$ such that $u_i = \sum_k w'_{i,k}\sqrt{\lambda'_k} f_k$.

Substituting this into (B.9), we obtain

$$\sqrt{\lambda_j} e_j = \sum_i \sum_k w'_{i,k} w_{i,j}^* \sqrt{\lambda'_k} f_k \otimes v_i.$$

Taking the norm on both sides, we have

$$\lambda_j = \sum_k D_{j,k} \lambda'_k, \quad D_{j,k} \stackrel{\mathrm{def}}{=} \left(\sum_{i,i'} w'_{i,k} w_{i,j}^* (w'_{i',k})^* w_{i',j} \langle v_{i'} | v_i \rangle \right).$$

If we can show that $D_{j,k}$ is a double stochastic transition matrix, Condition ③ in Theorem 8.2 implies $(\lambda'_k) \preceq (\lambda_j)$. Since

$$\left(\sum_{i,i'} w'_{i,k} w_{i,j}^* (w'_{i',k})^* w_{i',j} \langle v_{i'} | v_i \rangle \right) = \left\langle \sum_{i'} w'_{i',k} w_{i',j}^* v_{i'} \middle| \sum_i w'_{i,k} w_{i,j}^* v_i \right\rangle \geq 0$$

and $W'^* W' = I, W^* W = I$, we obtain

$$\begin{aligned} \sum_k \left(\sum_{i,i'} w'_{i,k} w_{i,j}^* (w'_{i',k})^* w_{i',j} \langle v_{i'} | v_i \rangle \right) &= \sum_{i,i'} \delta_{i,i'} w_{i,j}^* w_{i',j} \langle v_{i'} | v_i \rangle \\ &= \sum_i w_{i,j}^* w_{i,j} = 1. \end{aligned}$$

We may similarly show that $\sum_j D_{j,k} = 1$. Hence, $D_{j,k}$ is a double stochastic transition matrix.

B.4 Proof of Theorem 8.8 for Mixed States

We show the $\leq$ part of (8.87) for a general state ρ. Let $\{E_{A,i} \otimes E_{B,i}\}_i$ be the Choi–Kraus representation of an S-TP-CP map κ. Then,

$$\kappa(|\Phi_L\rangle\langle\Phi_L|) = \sum_i (E_{A,i} \otimes E_{B,i}) |\Phi_L\rangle\langle\Phi_L| (E_{A,i} \otimes E_{B,i})^* .$$

Now, choose y_i such that

$$\begin{aligned} p'_i &\stackrel{\text{def}}{=} \mathrm{Tr}\, (E_{A,i} \otimes E_{B,i}) |\Phi_L\rangle\langle\Phi_L| (E_{A,i} \otimes E_{B,i})^* \\ p'_i |y_i\rangle\langle y_i| &= (E_{A,i} \otimes E_{B,i}) |\Phi_L\rangle\langle\Phi_L| (E_{A,i} \otimes E_{B,i})^* . \end{aligned}$$

From Corollary 8.1 there exists a probabilistic decomposition $\{(p_i, x_i)\}$ of ρ such that

$$F(\kappa(|\Phi_L\rangle\langle\Phi_L|), \rho) = \sum_i \sqrt{p_i p'_i} |\langle x_i | y_i \rangle|.$$

Since the Schmidt rank of y_i is at most L,

$$|\langle x_i | y_i \rangle| \leq \sqrt{P(x_i, L)}. \tag{B.10}$$

From the Schwarz inequality,

$$\begin{aligned} F(\kappa(|\Phi_L\rangle\langle\Phi_L|), \rho) &\leq \sum_i \sqrt{p_i p'_i} \sqrt{P(x_i, L)} \\ &\leq \sqrt{\sum_i p'_i} \sqrt{\sum_i p_i P(x_i, L)} = \sqrt{\sum_i p_i P(x_i, L)}. \end{aligned} \tag{B.11}$$

Thus, we obtain the $\leq$ part of (8.87).

Conversely, if there exists a vector y_i with a Schmidt rank of at most L that satisfies the equality in (B.10) and

$$p'_i = \frac{p_i P(x_i, L)}{\sum_j p_j P(x_j, L)},$$

the equality in (B.11) holds. Therefore, according to Theorem 8.4, there exists a one-way LOCC satisfying the RHS of (8.85).

B.5 Proof of Theorem 8.9 for Mixed States

B.5.1 Proof of Direct Part

The second equality in (8.94) holds according to (8.84) and Lemma A.1. We therefore show the $\leq$ part of the first equality. Let us first show that

$$\min_{(p_i, x_i)} \left\{ \sum_i p_i (1 - P(x_i, [e^{nR}])) \middle| \sum_i p_i |x_i\rangle\langle x_i| = \rho^{\otimes n} \right\}$$

converges to zero exponentially for $R > E_f(\rho)$. The convergence of this expression to zero is equivalent to that of the value inside the $\sqrt{\ }$ on the RHS of (8.87) to one. Hence, we consider the latter quantity, i.e., the value inside the $\sqrt{\ }$. Choose a decomposition $\{(p_i, x_i)\}$ such that $R > \sum_i p_i E(|x_i\rangle\langle x_i|)$. Let $\rho_i \stackrel{\text{def}}{=} \mathrm{Tr}_B |x_i\rangle\langle x_i|$. From (8.88),

$$\sum_i p_i P(x_i, [e^R]) \le \sum_i p_i e^{\frac{\log \mathrm{Tr}\, \rho_i^{1-s} - sR}{1-s}}.$$

In particular, since $\frac{\log \mathrm{Tr}(\rho_i \otimes \rho_j)^{1-s} - 2sR}{1-s} = \frac{\log \mathrm{Tr}\, \rho_i^{1-s} - sR}{1-s} + \frac{\log \mathrm{Tr}\, \rho_j^{1-s} - sR}{1-s}$, we obtain

$$\begin{aligned}\sum_{i^n} p_{i^n}^n P(x_{i^n}^n, [e^{nR}]) \le & \sum_{i^n} p_{i^n}^n e^{\frac{\log \mathrm{Tr}(\rho_{i^n}^n)^{1-s} - snR}{1-s}} \\ = & \left(\sum_i p_i e^{\frac{\log \mathrm{Tr}\, \rho_i^{1-s} - sR}{1-s}} \right)^n, \end{aligned} \tag{B.12}$$

where we define $x_{i^n}^n \stackrel{\text{def}}{=} x_{i_1} \otimes \cdots \otimes x_{i_n}$, $\rho_{i^n}^n \stackrel{\text{def}}{=} \rho_{i_1} \otimes \cdots \otimes \rho_{i_n}$ with respect to $i^n \stackrel{\text{def}}{=} (i_1, \ldots, i_n)$, and p^n is the independent and identical distribution of p. Further, we obtain

$$\begin{aligned}\lim_{s \to 0} \frac{1}{s} \log \left(\sum_i p_i e^{\frac{\log \mathrm{Tr}\, \rho_i^{1-s} - sR}{1-s}} \right) &= \frac{d}{ds} \log \left(\sum_i p_i e^{\frac{\log \mathrm{Tr}\, \rho_i^{1-s} - sR}{1-s}} \right) \Bigg|_{s=0} \\ &= \sum_i p_i (H(\rho_i) - R) < 0. \end{aligned}$$

Note that the inside of the logarithmic on the left-hand side (LHS) of the above equation is equal to 1 when $s = 0$. Taking an appropriate $1 > s_0 > 0$, we have $\log \left(\sum_i p_i e^{\frac{\log \mathrm{Tr}\, \rho_i^{1-s_0} - s_0 R}{1-s_0}} \right) < 0$. Thus, the RHS of (B.12) exponentially converges to zero. Therefore, we obtain $E_c^{\rightarrow}(\rho) \le E_f(\rho)$. Similarly, $E_c^{\rightarrow}(\rho^{\otimes k}) \le E_f(\rho^{\otimes k})$.

Next, we choose a sequence $\{m_n\}$ such that $(m_n - 1)k \le n \le m_n k$ with respect to n. Denote the partial trace of $(\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes m_n k - n}$ by C_n. Then,

$$F(\rho^{\otimes m_n k}, \kappa_{m_n}(|\Phi_{L_{m_n}}\rangle\langle\Phi_{L_{m_n}}|)) \ge F(\rho^{\otimes n}, C_n \circ \kappa_{m_n}(|\Phi_{L_{m_n}}\rangle\langle\Phi_{L_{m_n}}|)) \tag{B.13}$$

for κ_m, L_m. Therefore, if the LHS of (B.13) converges to zero, then the RHS also converges to zero. Since

$$\overline{\lim} \frac{1}{n} \log L_{m_n} = \frac{1}{k} \overline{\lim}_{m \to \infty} \frac{1}{m} \log L_m,$$

C_n is a local quantum operation, and $E_c^{\rightarrow}(\rho^{\otimes k}) \le E_f(\rho^{\otimes k})$, and we have

$$E_c^{\rightarrow}(\rho) \le \frac{E_c^{\rightarrow}(\rho^{\otimes k})}{k} \le \frac{E_f(\rho^{\otimes k})}{k}.$$

Considering $\inf_k$, we obtain the $\le$ part of (8.94).

B.5.2 Proof of Converse Part

Let us first consider the following lemma as a preparation.

Lemma B.1 *Let p be a probability distribution $p = \{p_i\}_{i=1}^d$. Then,*

$$\sum_{i=L+1}^{d} p_i^{\downarrow} \ge \frac{H(p) - \log L - \log 2}{\log(d-L) - \log L}. \tag{B.14}$$

Proof. By defining the double stochastic transition matrix $A = (a_{i,j})$

$$a_{i,j} \stackrel{\text{def}}{=} \begin{cases} \frac{1}{L} & \text{if } i,j \le L \\ \frac{1}{d-L} & \text{if } i,j > L \\ 0 & \text{otherwise,} \end{cases}$$

the image Ap satisfies $(Ap)_i = \begin{cases} \frac{P(p,L)}{L} & \text{if } i \le L \\ \frac{1-P(p,L)}{d-L} & \text{if } i > L \end{cases}$. From Condition ③ in Theorem 8.2 we have $Ap \preceq p$. Therefore,

$$H(p) \ge H(Ap) = -P(p,L)\log\frac{P(p,L)}{L} - (1-P(p,L))\log\frac{1-P(p,L)}{d-L}.$$

Since the binary entropy $h(x)$ is less than $\log 2$, we have

$$\begin{aligned} &P^c(p,L)(\log(d-L) - \log L) + \log 2 \\ \ge &P^c(p,L)(\log(d-L) - \log L) + h(P(p,L)) \ge H(p) - \log L. \end{aligned}$$

We thus obtain (B.14). ■

We now show the $\ge$ part of Theorem 8.9 by using Lemma B.1. Consider the sequence of S-TP-CP maps $\{\kappa_n\}$ and the sequence of maximally entangled states $\{|\Phi_{L_n}\rangle\}$ satisfying

$$F(\kappa_n(|\Phi_{L_n}\rangle\langle\Phi_{L_n}|), \rho^{\otimes n}) \to 1. \tag{B.15}$$

Combining (8.86) and (8.87) in Theorem 8.8 and Lemma B.1, we have

$$\begin{aligned} &1 - F^2(\kappa_n(|\Phi_{L_n}\rangle\langle\Phi_{L_n}\|), \rho^{\otimes n}) \\ &\quad \ge \min_{(p_i,x_i)} \left\{ \sum_i p_i P^c(x_i, L_n) \middle| \sum_i p_i |x_i\rangle\langle x_i| = \rho^{\otimes n} \right\} \\ &\quad \ge \min_{(p_i,x_i)} \left\{ \sum_i p_i \frac{E(|x_i\rangle\langle x_i|) - \log L_n - \log 2}{\log(d^n - L_n) - \log L_n} \middle| \sum_i p_i |x_i\rangle\langle x_i| = \rho^{\otimes n} \right\} \\ &\quad = \frac{E_f(\rho^{\otimes n}) - \log L_n - \log 2}{\log(d^n - L_n) - \log L_n} = \frac{\frac{E_f(\rho^{\otimes n})}{n} - \frac{\log L_n}{n} - \frac{\log 2}{n}}{\frac{\log(d^n - L_n)}{n} - \frac{\log L_n}{n}}. \end{aligned}$$

Using (B.15) and Lemma A.1, we obtain

$$\begin{aligned}0 &= \lim \left(\frac{\log(d^n - L_n)}{n} - \frac{\log L_n}{n} \right) \left(1 - \left(F(\kappa_n(|\Phi_{L_n}\rangle\langle\Phi_{L_n}|), \rho^{\otimes n})\right)^2\right) \\ &\geq \overline{\lim} \left(\frac{E_f(\rho^{\otimes n})}{n} - \frac{\log L_n}{n} - \frac{\log 2}{n} \right) = \lim \frac{E_f(\rho^{\otimes n})}{n} - \underline{\lim} \frac{\log L_n}{n}.\end{aligned}$$

Thus, we obtain

$$\lim \frac{E_f(\rho^{\otimes n})}{n} \leq \underline{\lim} \frac{\log L_n}{n} \leq \overline{\lim} \frac{\log L_n}{n} \leq E_c(\rho),$$

which completes the proof of Theorem 8.9.

B.6 Proof of Theorem 9.3

First we prove Lemma 9.1.

Proof of Lemma 9.1.
S $\Rightarrow$ **A:** From (9.26), **S** implies (9.30), *i.e.*, **A**.
S $\Rightarrow$ **L:** From (9.27),

$$\begin{aligned}&\min_{\rho^{1,2}:\mathrm{Tr}\,\rho^{1,2}(X^1+X^2)\leq K} f^{1,2}(\rho^{1,2}) \\ &\quad\geq \min_{\rho^{1,2}:\mathrm{Tr}\,\rho^{1,2}(X^1+X^2)\leq K} f^1(\rho^1) + f^2(\rho^2) \\ &\quad= \min_{\rho^1,\rho^2:\mathrm{Tr}\,\rho^1 X^1+\mathrm{Tr}\,\rho^2 X^2\leq K} f^1(\rho^1) + f^2(\rho^2) \\ &\quad= \min_{0\leq\lambda\leq 1} \min_{\rho^1:\mathrm{Tr}\,\rho^1 X^1\leq\lambda K} f^1(\rho^1) + \min_{\rho^2:\mathrm{Tr}\,\rho^2 X^2\leq(1-\lambda)K} f^2(\rho^2).\end{aligned}$$

On the other hand, since $f^1(\rho^1) + f^2(\rho^2) \geq f^{1,2}(\rho^1 \otimes \rho^2)$, we have

$$\begin{aligned}&\min_{\rho^1,\rho^2:\mathrm{Tr}\,\rho^1 X^1+\mathrm{Tr}\,\rho^2 X^2\leq K} f^1(\rho^1) + f^2(\rho^2) \\ &\quad\geq \min_{\rho^1,\rho^2:\mathrm{Tr}\,\rho^1 X^1+\mathrm{Tr}\,\rho^2 X^2\leq K} f^{1,2}(\rho^1 \otimes \rho^2) \\ &\quad\geq \min_{\rho^{1,2}:\mathrm{Tr}\,\rho^{1,2}(X^1+X^2)\leq K} f^{1,2}(\rho^{1,2}).\end{aligned}$$

Hence, we obtain (9.29).
L $\Rightarrow$ **C:** Choose $\rho_0^{1,2}$ such that $\mathrm{Tr}\,\rho_0^{1,2}(X^1 + X^2) - f^{1,2}(\rho_0^{1,2}) = \max_{\rho^{1,2}} \mathrm{Tr}\,\rho^{1,2}(X^1+X^2) - f^{1,2}(\rho^{1,2})$. Then, the real number $K \stackrel{\mathrm{def}}{=} \mathrm{Tr}\,\rho_0^{1,2}(X^1+X^2)$ satisfies

$$\begin{aligned}
&\max_{\rho^{1,2}} \operatorname{Tr} \rho^{1,2}(X^1 + X^2) - f^{1,2}(\rho^{1,2}) \\
&\quad = \max_{\rho^{1,2}:\operatorname{Tr} \rho^{1,2}(X^1+X^2)\ge K} \operatorname{Tr} \rho^{1,2}(X^1 + X^2) - f^{1,2}(\rho^{1,2}) \\
&\quad = K + \max_{\rho^{1,2}:\operatorname{Tr} \rho^{1,2}(X^1+X^2)\ge K} -f^{1,2}(\rho^{1,2}) \\
&\quad = K + \max_{\rho^{1},\rho^{2}:\operatorname{Tr} \rho^{1}X^1+\operatorname{Tr} \rho^{2}X^2\ge K} -f^{1}(\rho^{1}) - f^{2}(\rho^{2}) \\
&\quad = \max_{\rho^{1},\rho^{2}:\operatorname{Tr} \rho^{1}X^1+\operatorname{Tr} \rho^{2}X^2\ge K} \operatorname{Tr} \rho^{1}X^1 - f^{1}(\rho^{1}) + \operatorname{Tr} \rho^{2}X^2 - f^{2}(\rho^{2}) \\
&\quad \le \max_{\rho^{1},\rho^{2}} \operatorname{Tr} \rho^{1}X^1 - f^{1}(\rho^{1}) + \operatorname{Tr} \rho^{2}X^2 - f^{2}(\rho^{2}).
\end{aligned}$$

Conversely, from (9.26),

$$\begin{aligned}
&\max_{\rho^{1,2}} \operatorname{Tr} \rho^{1,2}(X^1 + X^2) - f^{1,2}(\rho^{1,2}) \\
&\quad \ge \max_{\rho^{1},\rho^{2}} \operatorname{Tr} \rho^{1,2}(X^1 + X^2) - f^{1,2}(\rho^{1} \otimes \rho^{2}) \\
&\quad \ge \max_{\rho^{1},\rho^{2}} \operatorname{Tr} \rho^{1}X^1 - f^{1}(\rho^{1}) + \operatorname{Tr} \rho^{2}X^2 - f^{2}(\rho^{2}).
\end{aligned}$$

Hence, we obtain (9.28).

C $\Rightarrow$ **S:** For any $\rho_0^{1,2}$, from Lemma A.6, we choose Hermitian matrices X^1 and X^2 such that $\operatorname{Tr} \rho_0^i X^i - f^i(\rho_0^i) = \max_{\rho^i} \operatorname{Tr} \rho^i X^i - f^i(\rho^i)$. Hence,

$$\begin{aligned}
&\sum_{i=1}^2 \operatorname{Tr} \rho_0^i X^i - f^i(\rho_0^i) = \sum_{i=1}^2 \max_{\rho^i} \operatorname{Tr} \rho^i X^i - f^i(\rho^i) \\
&\quad = \max_{\rho^{1,2}} \operatorname{Tr} \rho^{1,2}(X^1 + X^2) - f^{1,2}(\rho^{1,2}) \ge \operatorname{Tr} \rho_0^{1,2}(X^1 + X^2) - f^{1,2}(\rho_0^{1,2}).
\end{aligned}$$

Since $\operatorname{Tr} \rho_0^{1,2}(X^1 + X^2) = \operatorname{Tr} \rho_0^1 X^1 + \operatorname{Tr} \rho_0^2 X^2$, we have (9.27). ■

Proof of HM $\Rightarrow$ HC. First, we assume that there exists a channel $\kappa_{X,p}$ for any channel κ, any positive semi-definite Hermitian matrix X on the input system $\mathcal{H}_A$, and any probability p, such that

$$C_c(\kappa_{X,p}) = \max_{\rho} (1-p)(\chi_\kappa(\rho)) + p \operatorname{Tr} H\rho, \tag{B.16}$$

and

$$\begin{aligned}
&C_c(\kappa^1_{X^1,p} \otimes \kappa^2_{X^2,p}) \\
&\quad = \max_{\rho} \Big((1-p)^2 \chi_{\kappa^1 \otimes \kappa^2}(\rho) + (1-p)p(\chi_{\kappa^1}(\rho^1) + \operatorname{Tr} X^2\rho^2) \\
&\qquad + (1-p)p(\chi_{\kappa^2}(\rho^2) + \operatorname{Tr} X^1\rho^1) + p^2(\operatorname{Tr} X^1\rho^1 + \operatorname{Tr} X^2\rho^2) \Big).
\end{aligned} \tag{B.17}$$

The channel $\kappa_{X,p}$ is called *Shor Extension* [374] of κ. Apply Condition **HM** to the channel $\kappa^1_{\frac{1}{p}X^1,p} \otimes \kappa^2_{\frac{1}{p}X^2,p}$, then we have

$$\begin{aligned}
&\max_{\rho^{1,2}}\Big((1-p)^2\chi_{\kappa^1\otimes\kappa^2}(\rho^{1,2})+(1-p)p(\chi_{\kappa^1}(\rho^1)+\operatorname{Tr}\frac{1}{p}X^2\rho^2)\\
&\quad+(1-p)p(\chi_{\kappa^2}(\rho^2)+\operatorname{Tr}\frac{1}{p}X^1\rho^1)+p^2(\operatorname{Tr}\frac{1}{p}X^1\rho^1+\operatorname{Tr}\frac{1}{p}X^2\rho^2)\Big)\\
&\le \max_{\rho^1}(1-p)(\chi_{\kappa^1}(\rho^1)+\operatorname{Tr}\frac{1}{p}X^1\rho^1)+\max_{\rho^2}(1-p)(\chi_{\kappa^2}(\rho^2)+\operatorname{Tr}\frac{1}{p}X^2\rho^2).
\end{aligned}$$

Taking the limit $p\to 0$, we obtain

$$\begin{aligned}
&\max_{\rho^{1,2}}\chi_{\kappa^1\otimes\kappa^2}(\rho^{1,2})+\operatorname{Tr}(X^1+X^2)\rho^{1,2}\\
&\quad\le \max_{\rho^1}(\chi_{\kappa^1}(\rho^1)+\operatorname{Tr}X^1\rho^1)+\max_{\rho^2}(\chi_{\kappa^2}(\rho^2)+\operatorname{Tr}X^2\rho^2),
\end{aligned}$$

which implies Condition **HC**.

Next, we define the channel $\kappa_{X,p}$ with the input system $\mathcal{H}_A\otimes\mathbb{C}^k$, where $k\ge\|X\|$, and check (B.17). First, we generate one-bit random number X with probability $P_0=1-p$ and $P_1=p$. When $X=0$, the output state is $\kappa(\operatorname{Tr}_{\mathbb{C}^k}\rho)$ for the input state ρ. Otherwise, we perform the measurement of the spectral decomposition of X, and send the receiver the state $\tilde{\kappa}^y(\operatorname{Tr}_A\rho)$ depending on its measured data y, which is eigenvalue of X. Here, we defined the channel $\tilde{\kappa}^y$ and the stochastic transition matrix Q_l^j such that

$$\tilde{\kappa}^y(\sigma)=\sum_l\sum_j Q_l^j|u_l\rangle\langle u_l|\langle u_j|\sigma|u_j\rangle,\quad y=C_c(Q)=I(p_{\mathrm{mix}},Q).$$

In this case, we assume that the receiver received the information X and y. Then, the relation (B.16) holds. From these discussions, we can check the equation (B.17). ■

Proof of EM $\Rightarrow$ EC. We define the channel $\tilde{\kappa}_{H,p}$ with the input system $\mathcal{H}_A$ as follows First, we generate one-bit random number X with probabilities $P_0=1-p$ and $P_1=p$. When $X=0$, the output state is $\kappa(\operatorname{Tr}_{\mathbb{C}^k}\rho)$ for the input state ρ. When $X=1$, we perform the measurement of the spectral decomposition of H, and obtain the eigenvalue y of H. Then, the output state is ρ_y, where ρ_y satisfies $H(\rho_y)=y$. In this case, the receiver is assumed to receive the information X and y. Then, the output entropy of the channel $\tilde{\kappa}_{H,p}$ can be calculated as

$$H(\tilde{\kappa}_{X,p}(\rho))=(1-p)H(\kappa(\rho))+p\operatorname{Tr}X\rho+h(p)-pH(\mathrm{P}_\rho^{E_X}).$$

Further,

$$\begin{aligned}
&H(\tilde{\kappa}^1_{X^1,p}\otimes\tilde{\kappa}^2_{X^2,p}(\rho))\\
&\quad=(1-p)^2(H(\kappa^1\otimes\kappa^2(\rho)))+p(1-p)(\operatorname{Tr}X^1\rho^1+H(\kappa^2(\rho^2)))\\
&\qquad+p(1-p)(\operatorname{Tr}X^2\rho^2+H(\kappa^1(\rho^1)))+p^2(\operatorname{Tr}X^1\rho^1+\operatorname{Tr}X^2\rho^2)+2h(p)\\
&\qquad-pH(\mathrm{P}_{\rho^1}^{E_{X^1}})-pH(\mathrm{P}_{\rho^2}^{E_{X^2}}).
\end{aligned}$$

Condition **EM** implies

$$\min_{\rho^{1,2}} H_{\tilde{\kappa}^1_{\frac{1}{p}X^1,p}\otimes\tilde{\kappa}^2_{\frac{1}{p}X^2,p}}(\rho^{1,2}) = \min_{\rho^1} H_{\tilde{\kappa}^1_{\frac{1}{p}X^1,p}}(\rho^1) + \min_{\rho^2} H_{\tilde{\kappa}^2_{\frac{1}{p}X^2,p}}(\rho^2).$$

Since $H(\mathrm{P}^{E_{X^2}}_{\rho^2}) \le \log d_A d_B$, taking the limit $p \to 0$, we have

$$H(\tilde{\kappa}_{\frac{1}{p}X,p}, \rho) \to H_\kappa(\rho) + \operatorname{Tr} X\rho$$
$$H_{\tilde{\kappa}^1_{\frac{1}{p}X^1,p}\otimes\tilde{\kappa}^2_{\frac{1}{p}X^2,p}}(\rho) \to H_{\kappa^1\otimes\kappa^2}(\rho) + (\operatorname{Tr} X^1\rho^1 + \operatorname{Tr} X^2\rho^2).$$

Since the set of density matrices is compact, we obtain

$$\begin{aligned}&\min_{\rho^{1,2}} H_{\kappa^1\otimes\kappa^2}(\rho^{1,2}) + \operatorname{Tr}(X^1+X^2)\rho^{1,2} \\ &\quad = \min_{\rho^1,\rho^2}(H_{\kappa^1}(\rho^1) + H_{\kappa^2}(\rho^2) + \operatorname{Tr} X^1\rho^1 + \operatorname{Tr} X^2\rho^2,\end{aligned}$$

which implies **EC**. ■

Proof of FA ⇒ FS. See Pomeransky [350].

B.7 Proof of Lemma 9.4

In this section, we prove Lemma 9.4 from Lemma B.2 given below. For this proof, we use arguments similar to that in the proof of Theorem 9.7.

Following these arguments, we can choose ML elements $\{x_{m,l}\}_{(m,l)\in\{1,\dots,M\}\times\{1,\dots,L\}}$ of $\mathcal{X}$ and a POVM $\{Y_{m,l}\}$ (the elements $x_{m,l}$ are not necessarily different from each other) such that

$$\begin{aligned}&\frac{8}{ML}\sum_{m,l}\epsilon_{B,m,l} + \frac{1}{M}\sum_{m=1}^{M}\epsilon_{E,m} \le (\text{RHS of (9.83)}) \\ &\epsilon_{B,m,l} \stackrel{\text{def}}{=} 1 - \operatorname{Tr} W_{B,x_{m,j}} Y_{m,l}, \quad \epsilon_{E,m} \stackrel{\text{def}}{=} \left\| \frac{1}{L}\sum_{l=1}^{L} W_{E,x_{m,l}} - \sum_x p_x W_{E,x} \right\|.\end{aligned} \tag{B.18}$$

In what follows, we prove Lemma 9.4 using the above argument and Lemma B.2. The chosen element $x_{m,l}$ may in general contain some degeneracies. However, to apply Lemma B.2, no degeneracies should be present. In order to resolve this problem, we define new ML elements $x'_{m,l}$ satisfying the following conditions: (1) the elements $x'_{m,l}$ have no degeneracies. (2) $\{x_{m,l}\} \subset \{x'_{m,l}\}$. (3) For any element $x_{m,l}$, there exits an index (m_1, l_1) such that $x_{m_1,l_1} = x'_{m_1,l_1}$ among indices (m_1, l_1) satisfying $x_{m_1,l_1} = x_{m,l}$. The POVM $\{Y_{m,l}\}$ is also rewritten such that

$$Y'_{m,l} \stackrel{\text{def}}{=} \begin{cases} \sum_{x_{m',l'}=x_{m,l}} Y_{m',l'} & x_{m,l} = x'_{m,l} \\ 0 & x_{m,l} \neq x'_{m,l}. \end{cases}$$

Denoting the error probability by $\epsilon'_{B,m,l}$ in a manner similar to (B.18), we have

$$\sum_{m,l} \epsilon_{B,m,l} = \sum_{m,l} \epsilon'_{B,m,l}.$$

The number of elements in $\{x_{m,l} \neq x'_{m,l}\}$ is less than $\sum_{m,l} \epsilon'_{B,m,l}$. Therefore,

$$\frac{1}{M} \sum_{m=1}^{M} \left\| \frac{1}{L} \sum_{l=1}^{L} W^E_{x_{m,l}} - \frac{1}{L} \sum_{l=1}^{L} W^E_{x'_{m,l}} \right\| \leq \frac{2}{ML} \sum_{m,l} \epsilon_{B,m,l}.$$

Define $\epsilon'_{E,m}$ with respect to the pair $\{x'_{m,l}\}$ in a manner similar to (B.18). Then, we have

$$\begin{aligned} &\frac{6}{ML} \sum_{m,l} \epsilon'_{B,m,l} + \frac{1}{M} \sum_{m=1}^{M} \epsilon'_{E,m} \leq \frac{8}{ML} \sum_{m,l} \epsilon_{B,m,l} + \frac{1}{M} \sum_{m=1}^{M} \epsilon_{E,m} \\ &\quad \leq (\text{RHS of (9.83)}). \end{aligned}$$

Hence, Lemma 9.4 can be obtained by applying Lemma B.2 to $u_{x'_{m,l}}$.

Lemma B.2 *Let $\{u^A_{m,l}\}_{(m,l)\in\{1,\dots,M\}\times\{1,\dots,L\}}$ be ML mutually orthogonal normalized vectors in $\mathcal{H}_A$, $U_A^{B,E}$ be an isometric map from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_E$, $Y = \{Y_{m,l}\}_{(m,l)\in\{1,\dots,M\}\times\{1,\dots,L\}}$ be a POVM in $\mathcal{H}_B$, and κ be a TP-CP map from $\mathcal{H}_A$ to $\mathcal{H}_B$ defined according to $U_A^{B,E}$. We define*

$$\begin{aligned} \epsilon^B_{m,l} &\stackrel{\text{def}}{=} 1 - \operatorname{Tr} U_A^{B,E} |u^A_{m,l}\rangle\langle u^A_{m,l}| (U_A^{B,E})^* (Y_{m,l} \otimes I_E) \\ \epsilon^E_m &\stackrel{\text{def}}{=} \left\| \left(\frac{1}{L} \sum_{l=1}^{L} W^E_{m,l} \right) - W^E \right\|_1, \ W^E_{m,l} \stackrel{\text{def}}{=} \operatorname{Tr}_B U_A^{B,E} |u^A_{m,l}\rangle\langle u^A_{m,l}| (U_A^{B,E})^* \end{aligned}$$

with respect to a state W^E in $\mathcal{H}_E$.

Now let $\mathcal{H}_C \stackrel{\text{def}}{=} \langle u^C_1, \dots, u^C_M \rangle$ and $\mathcal{H}_D \stackrel{\text{def}}{=} \langle u^D_1, \dots, u^D_M \rangle$ be spaces with a dimension of M. When we choose an encoding τ_α from $\mathcal{H}_C$ to $\mathcal{H}_A$ and a decoding ν_α from $\mathcal{H}_B$ to $\mathcal{H}_D$, depending on the random variable $\alpha = (\alpha_1, \dots, \alpha_M)$ by following the method below, we have

$$1 - \mathrm{E}_\alpha F_e(\rho_{\mathrm{mix},C}, \nu_\alpha \circ \kappa \circ \tau_\alpha) \leq \frac{6}{ML} \sum_{m,l} \epsilon^B_{m,l} + \frac{1}{M} \sum_{M} \epsilon^E_m, \tag{B.19}$$

where α_m $(m = 1, \dots, M)$ are independent random variables subject to the uniform distribution on the integers $0, \dots, L-1$, and $\mathcal{H}_C$ and $\mathcal{H}_D$ are identified due to the natural correspondence $u^C_m \mapsto u^D_m$.

The encoder τ_α and the decoder ν_α are constructed as follows. Define the isometric map $U^A_{C,\alpha}$ from $\mathcal{H}_C$ to $\mathcal{H}_A$ by

$$U^A_{C,\alpha} u^C_m \stackrel{\rm def}{=} \left(u^A_{m,\alpha} \stackrel{\rm def}{=} \sum_{l=1}^L e^{(2l\alpha_m \pi i)/L} u^A_{m,l} \right).$$

We also define τ_α according to $\tau_\alpha(\rho) \stackrel{\rm def}{=} U^A_{C,\alpha} \rho (U^A_{C,\alpha})^*$. Next, we choose the space $\mathcal{H}_{B'}$, containing $\mathcal{H}_B$, to be sufficiently large such that the purification $\hat{u}$ of W^E can be taken as an element of $\mathcal{H}_E \otimes \mathcal{H}_{B'}$. We also choose the isometric map $U^{B,D}_B$ from $\mathcal{H}_B$ to $\mathcal{H}_B \otimes \mathcal{H}_D$, the unitary matrix $U^{B'}_{m,\alpha}$ in $\mathcal{H}_{B'}$, and the unitary matrix $U^{B',D}_\alpha$ in $\mathcal{H}_{B'} \otimes \mathcal{H}_D$ such that

$$\begin{aligned}
&{\rm Tr}\, \rho Y_m = \langle u^D_m | U^{B,D}_B \rho (U^{B,D}_B)^* | u^D_m \rangle, \quad Y_m \stackrel{\rm def}{=} \sum_{l=1}^L Y_{m,l}, \qquad \text{(B.20)} \\
&U^{B'}_{m,\alpha} \stackrel{\rm def}{=} \mathop{\rm argmax}_U |\langle u^D_m \otimes (U^* \otimes I_E) \hat{u} | U^{B,D}_B U^{B,E}_A u^A_{m,\alpha} \rangle|, \\
&U^{B',D}_\alpha \stackrel{\rm def}{=} \sum_{m=1}^M |u^D_m\rangle\langle u^D_m| \otimes U^{B'}_{m,\alpha}.
\end{aligned}$$

The first condition (B.20) is the condition that $U^{B,D}_B$ gives a Naĭmark extension of $\{Y_m\}_{m=1}^M$. We choose $U^{B'}_{m,\alpha}$ such that $\langle u^D_m \otimes ((U^{B'}_{m,\alpha})^* \otimes I_E) \hat{u} | U^{B,D}_B U^{B,E}_A u^A_{m,\alpha} \rangle$ is a nonnegative real number. Let ν_α be given by

$$\nu_\alpha(\rho) \stackrel{\rm def}{=} {\rm Tr}_{B'}\, U^{B',D}_\alpha U^{B,D}_B \rho (U^{B,D}_B)^* (U^{B',D}_\alpha)^*. \qquad \text{(B.21)}$$

This discussion gives the construction of the encoder τ_α and the decoder ν_α.

Proof. Let $\mathcal{H}_R \stackrel{\rm def}{=} \langle u^R_1, \ldots, u^R_M \rangle$ be the environment of $\mathcal{H}_C$. Then,

$$\begin{aligned}
&{\rm E}_\alpha F_e(\rho_{{\rm mix},C}, \nu_\alpha \circ \kappa \circ \tau_\alpha) \\
&= {\rm E}_\alpha F \left(\frac{1}{\sqrt{M}} \sum_{m=1}^M u^R_m \otimes u^C_m, \nu_\alpha \circ \kappa \circ \tau_\alpha \left(\frac{1}{\sqrt{M}} \sum_{m=1}^M u^R_m \otimes u^C_m \right) \right) \\
&\le F \left(\left(\frac{1}{\sqrt{M}} \sum_{m=1}^M u^R_m \otimes u^C_m \right) \otimes \hat{u}, \frac{1}{\sqrt{M}} \sum_{m=1}^M u^R_m \otimes U^{B',D}_\alpha U^{B,D}_B U^{B,E}_A u^A_{m,\alpha} \right) \\
&= \left| \left\langle \frac{1}{\sqrt{M}} \sum_{m'=1}^M u^R_{m'} \otimes u^C_{m'} \otimes \hat{u} \right| \frac{1}{\sqrt{M}} \sum_{m=1}^M u^R_m \otimes U^{B',D}_\alpha U^{B,D}_B U^{B,E}_A u^A_{m,\alpha} \right\rangle \Bigg| \\
&= \left| \frac{1}{M} \sum_{m=1}^M \langle u^C_m \otimes \hat{u} \left| U^{B',D}_\alpha U^{B,D}_B U^{B,E}_A u^A_{m,\alpha} \right\rangle \right|
\end{aligned}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \left\langle u_m^C \otimes \hat{u} \left| U_\alpha^{B',D} U_B^{B,D} U_A^{B,E} u_{m,\alpha}^A \right\rangle \right.$$
$$= \frac{1}{M} \sum_{m=1}^{M} F(U_B^{B,D} U_A^{B,E} u_{m,\alpha}^A, u_m^C \otimes (U_{m,\alpha}^{B'})^* \hat{u}). \tag{B.22}$$

Next, we evaluate $1 - F(U_B^{B,D} U_A^{B,E} u_{m,\alpha}^A, u_m^C \otimes (U_{m,\alpha}^{B'})^* \hat{u})$. Note that the relation $b^2(\rho, \sigma) = 1 - F(\rho, \sigma)$ will be used frequently in what follows. For this evaluation, we choose a sufficiently large space $\mathcal{H}_{B''}$ containing $\mathcal{H}_{B'}$ and a PVM $E^m = \{E_l^m\}_{l=1}^L$ on $\mathcal{H}_{B''}$ such that

$$\begin{aligned} &\mathrm{Tr}\, \rho Y_{m,l} (= \mathrm{Tr}\, \sqrt{Y_m} \rho \sqrt{Y_m} (\sqrt{Y_m})^{-1} \rho Y_{m,l} (\sqrt{Y_m})^{-1}) \\ &\quad = \mathrm{Tr}(|u_m^D\rangle\langle u_m^D| \otimes E_l^m) U_B^{B,D} \rho (U_B^{B,D})^* \\ &\qquad \left(= \mathrm{Tr}(I_D \otimes E_l^m)(|u_m^D\rangle\langle u_m^D| \otimes I_B) U_B^{B,D} \rho (U_B^{B,D})^* (|u_m^D\rangle\langle u_m^D| \otimes I_B) \right). \end{aligned}$$

Define a normalized vector $u_{m,l}^{B'',E}$ in $\mathcal{H}_{B'',E}$ by

$$u_m^D \otimes u_{m,l}^{B'',E} \stackrel{\mathrm{def}}{=} \frac{1}{\sqrt{1 - \epsilon_{m,l}^E}} (|u_m^D\rangle\langle u_m^D| \otimes E_l^m) U_B^{B,D} U_A^{B,E} u_{m,l}^A$$

and a unitary $U_{m,\alpha}^{B''}$ by

$$U_{m,\alpha}^{B''} \stackrel{\mathrm{def}}{=} \operatorname*{argmax}_U F \left(\frac{1}{\sqrt{L}} \sum_{l=1}^{L} e^{(2l\alpha_m \pi i)/L} u_m^D \otimes u_{m,l}^{B'',E}, u_m^C \otimes (U)^* \hat{u} \right).$$

Then,

$$\begin{aligned} &b^2(U_B^{B,D} U_A^{B,E} u_{m,\alpha}^A, u_m^C \otimes (U_{m,\alpha}^{B'})^* \hat{u}) \\ &\quad \le b^2(U_B^{B,D} U_A^{B,E} u_{m,\alpha}^A, u_m^C \otimes (U_{m,\alpha}^{B''})^* \hat{u}) \\ &\quad \le 2b^2 \left(\frac{1}{\sqrt{L}} \sum_{l=1}^{L} e^{(2l\alpha_m \pi i)/L} u_m^D \otimes u_{m,l}^{B'',E}, u_m^C \otimes (U_{m,\alpha}^{B''})^* \hat{u} \right) \\ &\qquad + 2b^2 \left(U_B^{B,D} U_A^{B,E} u_{m,\alpha}^A, \frac{1}{\sqrt{L}} \sum_{l=1}^{L} e^{(2l\alpha_m \pi i)/L} u_m^D \otimes u_{m,l}^{B'',E} \right). \end{aligned} \tag{B.23}$$

Now we evaluate the first term in (B.23). Since

$$\begin{aligned}
&\mathrm{Tr}_{B''}\left|\frac{1}{\sqrt{L}}\sum_{l=1}^{L} e^{(2l\alpha_m\pi i)/L}u_{m,l}^{B'',E}\right\rangle\left\langle\frac{1}{\sqrt{L}}\sum_{l'=1}^{L} e^{(2l'\alpha_m\pi i)/L}u_{m,l'}^{B'',E}\right| \\
&= \mathrm{Tr}_{B''}\sum_{l''=1}^{L}(E_{l''}^{m}\otimes I_E) \\
&\quad\left(\frac{1}{L}\sum_{l=1}^{L}\sum_{l'=1}^{L} e^{(2(l-l')\alpha_m\pi i)/L}\left|u_{m,l}^{B'',E}a\right\rangle\left\langle u_{m,l'}^{B'',E}\right|\right)(E_{l''}^{m}\otimes I_E) \\
&= \mathrm{Tr}_{B''}\sum_{l=1}^{L}\frac{1}{L}\left|u_{m,l}^{B'',E}\right\rangle\left\langle u_{m,l}^{B'',E}\right|,
\end{aligned}$$

from the definition of $U_{m,\alpha}^{B''}$, we obtain

$$\begin{aligned}
&b^2\left(\frac{1}{\sqrt{L}}\sum_{l=1}^{L} e^{(2l\alpha_m\pi i)/L}u_m^D\otimes u_{m,l}^{B'',E}, u_m^C\otimes(U_{m,\alpha}^{B''})^*\hat{u}\right) \\
&= b^2\left(\frac{1}{\sqrt{L}}\sum_{l=1}^{L} e^{(2l\alpha_m\pi i)/L}u_{m,l}^{B'',E}, (U_{m,\alpha}^{B''})^*\hat{u}\right) \\
&= b^2\left(\mathrm{Tr}_{B''}\sum_{l=1}^{L}\frac{1}{L}\left|u_{m,l}^{B'',E}\right\rangle\left\langle u_{m,l}^{B'',E}\right|, W_E\right) \\
&\le 2b^2\Big(\mathrm{Tr}_{B''}\sum_{l=1}^{L}\frac{1}{L}\left|u_{m,l}^{B'',E}\right\rangle\left\langle u_{m,l}^{B'',E}\right|, \\
&\qquad\sum_{l=1}^{L}\frac{1}{L}\mathrm{Tr}_{B,D}\left|U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right\rangle\left\langle U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right|\Big) \\
&\quad + 2b^2\left(\sum_{l=1}^{L}\frac{1}{L}W_{m,l}^{E}, W_E\right),
\end{aligned}\tag{B.24}$$

where we use Lemma 8.2 and the fact that $\mathrm{Tr}_{B,D}\left|U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right\rangle\left\langle U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right| = W_{m,l}^{E}$. For the second term in (B.24), inequality (5.42) yields

$$b^2\left(\sum_{l=1}^{L}\frac{1}{L}W_{m,l}^{E}, W_E\right) \le \frac{1}{2}\left\|\sum_{l=1}^{L}\frac{1}{L}W_{m,l}^{E} - W_E\right\|_1 \le \frac{1}{2}\epsilon_m^E. \tag{B.25}$$

The first term in (B.24) is evaluated as

$$
\begin{aligned}
& b^2\Big(\mathrm{Tr}_{B''}\sum_{l=1}^{L}\frac{1}{L}\left|u_{m,l}^{B'',E}\right\rangle\left\langle u_{m,l}^{B'',E}\right|, \\
& \qquad \sum_{l=1}^{L}\frac{1}{L}\,\mathrm{Tr}_{B,D}\left|U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right\rangle\left\langle U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right|\Big) \\
& \le \sum_{l=1}^{L}\frac{1}{L}b^2\Big(\mathrm{Tr}_{B''}\left|u_{m,l}^{B'',E}\right\rangle\left\langle u_{m,l}^{B'',E}\right|, \\
& \qquad \mathrm{Tr}_{B'',D}\left|U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right\rangle\left\langle U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right|\Big) \\
& \le \sum_{l=1}^{L}\frac{1}{L}b^2\left(u_m^D\otimes u_{m,l}^{B'',E}, U_B^{B,D}U_A^{B,E}u_{m,l}^{A}\right) \\
& = \sum_{l=1}^{L}\frac{1}{L}\left(1-\sqrt{1-\epsilon_{m,l}^{B}}\right) \le \sum_{l=1}^{L}\frac{1}{L}\epsilon_{m,l}^{B},
\end{aligned}
\tag{B.26}
$$

where we use the fact that $\mathrm{Tr}_{B''}\left|u_{m,l}^{B'',E}\right\rangle\left\langle u_{m,l}^{B'',E}\right| = \mathrm{Tr}_{B'',D}\left|u_m^D\otimes u_{m,l}^{B'',E}\right\rangle\left\langle u_m^D\otimes u_{m,l}^{B'',E}\right|$. The second term of (B.23) can be evaluated as

$$
\begin{aligned}
& \mathrm{E}_\alpha F\left(U_B^{B,D}U_A^{B,E}u_{m,\alpha}^{A}, \sum_{l=1}^{L}\frac{e^{(2l\alpha_m\pi i)/L}}{\sqrt{L}}u_m^D\otimes u_{m,l}^{B'',E}\right) \\
& = \mathrm{E}_\alpha\left\langle\sum_{l=1}^{L}\frac{e^{(2l\alpha_m\pi i)/L}}{\sqrt{L}}U_B^{B,D}U_A^{B,E}u_{m,l}^{A}, \sum_{l=1}^{L}\frac{e^{(2l\alpha_m\pi i)/L}}{\sqrt{L}}u_m^D\otimes u_{m,l}^{B'',E}\right\rangle \\
& = \frac{1}{L}\sum_{l=1}^{L}\sum_{l'=1}^{L}(\mathrm{E}_\alpha e^{(2(l-l')\alpha_m\pi i)/L})\left\langle U_B^{B,D}U_A^{B,E}u_{m,l'}^{A}, u_m^D\otimes u_{m,l}^{B'',E}\right\rangle \\
& = \frac{1}{L}\sum_{l=1}^{L}\langle U_B^{B,D}U_A^{B,E}u_{m,l}^{A}, u_m^D\otimes u_{m,l}^{B'',E}\rangle \\
& = \frac{1}{L}\sum_{l=1}^{L}\left\langle U_B^{B,D}U_A^{B,E}u_{m,l}^{A}, \frac{1}{\sqrt{1-\epsilon_{m,l}^{E}}}(|u_m^D\rangle\langle u_m^D|\otimes E_l^m)U_B^{B,D}U_A^{B,E}u_{m,\alpha}^{A}\right\rangle \\
& = \frac{1}{L}\sum_{l=1}^{L}\sqrt{1-\epsilon_{m,l}^{E}} \ge \frac{1}{L}\sum_{l=1}^{L}(1-\epsilon_{m,l}^{E}).
\end{aligned}
\tag{B.27}
$$

Combining (B.23)–(B.27) with (B.22), we obtain (B.19). ■

B.8 Proof of Lemma 10.3

We first prove the following lemma.

Lemma B.3 *A visible encoder may be represented by a map from $\mathcal{X}$ to $\mathcal{S}(\mathcal{K})$. Consider the convex combination of codes T and T':*

$$(\lambda T + (1-\lambda)T')(x) \stackrel{\rm def}{=} \lambda T(x) + (1-\lambda)T'(x), \quad 0 < \forall \lambda < 1.$$

Then, the set of visible encoders is a convex set, and the set of extremal points (see Sect. A.4 for the definition of an extremal point) is equal to

$$\{T \,|T(x) \text{ is a pure state } \forall x \in \mathcal{X}\}. \tag{B.28}$$

Proof. When $T(x)$ is a pure state for every input x, T is therefore an extremal point because it is impossible to represent the encoder T as a convex combination of other encoders. Hence, to complete the proof, it is sufficient to show that an arbitrary visible encoder $T(x) = \sum_{j_x} s_{j_x} |\phi_{j_x}\rangle\langle\phi_{j_x}|$ can be represented as a convex combination of encoders satisfying the condition in (B.28). Define a visible encoder $T(j_1, j_2, \dots, j_n)$ by

$$T(j_1, j_2, \dots, j_n|i) = |\phi_{j_x}\rangle\langle\phi_{j_x}|.$$

Then, this encoder belongs to the set (B.28). Since $T = \sum_{j_1,j_2,\dots,j_n} s_{j_1} s_{j_2} \cdots s_{j_n}$ $T(j_1, j_2, \cdots, j_n)$, the proof is completed. ■

We also require the following lemma for the proof of Lemma 10.3. This lemma is equivalent to Theorem 8.3, which was shown from the viewpoint of entanglement in Sect. 8.4.

Lemma B.4 *Let $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be separable. Then,*

$$\begin{aligned} &\max\{\mathrm{Tr}\, P\rho_A | P : \text{Projection in } \mathcal{H}_A \text{ with rank } k\} \\ &\quad \geq \max\{\mathrm{Tr}\, P\rho | P : \text{Projection in } \mathcal{H}_A \otimes \mathcal{H}_B \text{ with rank } k\} \end{aligned}$$

holds for any integer k.

Proof of Lemma 10.3. According to Lemma B.3, it is sufficient to show (10.22) for a visible encoder T in (B.28). From Condition ⑥ in Theorem 5.1, there exist a space $\mathcal{H}'$ with the same dimension of $\mathcal{H}$, a pure state ρ_0 in $\mathcal{H}' \otimes \mathcal{H}$, and a unitary matrix U in $\mathcal{K} \otimes \mathcal{H}' \otimes \mathcal{H}$ such that $\nu(\rho) = \mathrm{Tr}_{\mathcal{K},\mathcal{H}'} U(\rho \otimes \rho_0)U^*$, and the state

$$\rho_x \stackrel{\rm def}{=} \frac{(W_x \otimes I)\, U\, (T(x) \otimes \rho_0)\, U^*\, (W_x \otimes I)}{\mathrm{Tr}\, U\, (T(x) \otimes \rho_0)\, U^*\, (W_x \otimes I)} \in \mathcal{S}(\mathcal{K} \otimes \mathcal{H} \otimes \mathcal{H}')$$

is a pure state. Since $UT(x) \otimes \rho_0 U^*$ is a pure state and $(W_x \otimes I)$ is a projection, we have

$$\operatorname{Tr}\nu(T(x))W_x = \operatorname{Tr}UT(x)\otimes\rho_0 U^*\,(W_x\otimes I) = \operatorname{Tr}U(T(x)\otimes\rho_0)\,U^*\rho_x. \quad \text{(B.29)}$$

Since $\operatorname{Tr}_{\mathcal{K},\mathcal{H}'}\rho_x = W_x$, we may write $\rho_x = W_x\otimes\sigma_x$ by choosing an appropriate pure state $\sigma_x\in\mathcal{S}(\mathcal{K}\otimes\mathcal{H}')$. Hence, the state $\rho_p \stackrel{\text{def}}{=} \sum_{i\in\mathcal{X}} p(x)\rho_x = \sum_{x\in\mathcal{X}} p(x)W_x\otimes\sigma_x$ is separable and satisfies $W_p = \operatorname{Tr}_{\mathcal{H},\mathcal{K}'}\rho_p$. Since $I_{\mathcal{K}}\geq T(x)$, we have $U\,(I_{\mathcal{K}}\otimes\rho_0)\,U^*\geq U\,(T(x)\otimes\rho_0)\,U^*$. Thus, from (B.29) we have

$$\begin{aligned}&\sum_{x\in\mathcal{X}} p(x)\operatorname{Tr}\nu(T(x))W_x = \sum_{x\in\mathcal{X}} p(x)\operatorname{Tr}_{\mathcal{H}}\operatorname{Tr}_{\mathcal{K}\otimes\mathcal{H}'} U\,(T(x)\otimes\rho_0)\,U^*\rho_x\\ &\leq \sum_{x\in\mathcal{X}} p(x)\operatorname{Tr}U\,(I_{\mathcal{K}}\otimes\rho_0)\,U^*\rho_x = \operatorname{Tr}U\,(I_{\mathcal{K}}\otimes\rho_0)\,U^*\rho_p. \quad \text{(B.30)}\end{aligned}$$

According to $I\geq U\,(I_{\mathcal{K}}\otimes\rho_0)\,U^*\geq 0$ and $\operatorname{Tr}U\,(I_{\mathcal{K}}\otimes\rho_0)\,U^* = \operatorname{Tr}I_{\mathcal{K}} = \dim\mathcal{K}$, we obtain

$$\begin{aligned}&\operatorname{Tr}U\,(I_{\mathcal{K}}\otimes\rho_0)\,U^*\rho_p \leq \max\left\{\operatorname{Tr}P\rho_p \,\middle|\, \begin{array}{l}P:\text{Projection in }\mathcal{K}\otimes\mathcal{H}\otimes\mathcal{H}',\\ \operatorname{rank}P=\dim\mathcal{K}\end{array}\right\}\\ &\leq \max\{\operatorname{Tr}PW_p | P:\text{Projection in }\mathcal{H}, \operatorname{rank}P=\dim\mathcal{K}\}. \quad \text{(B.31)}\end{aligned}$$

(B.31) may be obtained from Lemma B.4 and the separability of ρ_p. The projection P on $\mathcal{H}$ satisfies

$$\operatorname{Tr}(W_p - a)P \leq \operatorname{Tr}(W_p - a)\{W_p - a\geq 0\}.$$

If the rank of P is $\dim\mathcal{K}$ (i.e., if $\operatorname{Tr}P = \dim\mathcal{K}$), then

$$\operatorname{Tr}W_p P \leq a\dim\mathcal{K} + \operatorname{Tr}W_p\{W_p - a\geq 0\}. \quad \text{(B.32)}$$

From (B.30)–(B.32),

$$\begin{aligned}1-\varepsilon(\Psi) &= \sum_{x\in\mathcal{X}} p(x)\operatorname{Tr}\nu(T(x))W_x\\ &\leq \max\{\operatorname{Tr}PW_p | P:\ \text{Projection in }\mathcal{H},\ \operatorname{rank}P=\dim\mathcal{K}\}\\ &\leq a\dim\mathcal{K} + \operatorname{Tr}W_p\{W_p - a\geq 0\}.\end{aligned}$$

We therefore obtain (10.22). ■

C

Hints and Brief Solutions to Exercises

Exercise 1.1. Use the fact that the discriminant of $\langle u + rcv | u + rcv \rangle$ concerning r is negative.

Exercise 1.3. Note that $X^T = \sum_{i,j} \overline{x^{i,j}} |u_i\rangle\langle u_j|$. Show that $\overline{\langle x|Xx\rangle} = \langle \overline{x}|X^T\overline{x}\rangle \geq 0$, where $x = \sum_i x^i u_i$ and $\overline{x} = \sum_i \overline{x^i} u_i$.

Exercise 1.4. Consider the derivative of the product of the matrix elements.

Exercise 1.13. **d**: Take the partial trace on A or B.

Exercise 1.19. A matrix cannot be semipositive definite if its determinant is negative.

Exercise 1.23. Show $\mathrm{Tr}\{Y \geq 0\}X\{Y \geq 0\} \geq \mathrm{Tr}\{Y \geq 0\}Y\{Y \geq 0\}$ and $\mathrm{Tr}\{Y < 0\}X\{Y < 0\} \geq -\mathrm{Tr}\{Y < 0\}Y\{Y < 0\}$.

Exercise 2.2. Consider the case $\mathrm{P}_Y(1) = \lambda, \mathrm{P}_Y(0) = 1 - \lambda, \mathrm{P}_{X|Y=1} = p, \mathrm{P}_{X|Y=0} = p'$.

Exercise 2.6. Apply a stochastic transition matrix of rank 1 to Theorem 2.1.

Exercise 2.8. Use the fact that $\sum_j \sum_i Q_j^i |p_i - q_i| \geq \sum_j |\sum_i Q_j^i (p_i - q_i)|$.

Exercise 2.9. Consider the $x \geq y$ and $x < y$ cases separately.

Exercise 2.10. **a**: Use $|p_i - q_i| = |\sqrt{p_i} - \sqrt{q_i}||\sqrt{p_i} + \sqrt{q_i}|$. **b**: Use $p_i + q_i \geq 2\sqrt{p_i}\sqrt{q_i}$.

Exercise 2.11. Assume that the datum i generates with the probability distribution p_i. Apply Jensen's inequality to the random variable $\sqrt{q_i/p_i}$ and the convex function $-\log x$.

Exercise 2.13. Check that $\phi'(s|p\|q) = \dfrac{\sum_i p_i^{1-s} q_i^s (\log q_i - \log p_i)}{\sum_i p_i^{1-s} q_i^s}$.

Exercise 2.14. Check that $\phi''(s|p\|q) =$

$$\frac{(\sum_i p_i^{1-s} q_i^s)(\sum_i p_i^{1-s} q_i^s (\log q_i - \log p_i)^2) - (\sum_i p_i^{1-s} q_i^s (\log q_i - \log p_i))^2}{(\sum_i p_i^{1-s} q_i^s)^2}.$$

Next, use Schwarz's inequality between two vectors 1 and $(-\log p_i + \log q_i)$.

Exercise 2.18. Consider the spectral decomposition $\boldsymbol{M}$ of X and apply Jensen's inequality to $\mathrm{P}_\rho^{\boldsymbol{M}}$.

Exercise 2.21. Note that $2(x^2+y^2) \geq (x+y)^2$.

Exercise 2.22. **b**: Take the average by integrating between $[0, 2\pi]$ for each θ_i. Note that $\langle u|v\rangle$ is continuous for each θ_i.

Exercise 2.24. **d**: Use Schwarz's inequality with respect to the inner product $\operatorname{Tr} XY^*$ with two vectors $\rho^{(1-s)/2}(\log\sigma - \log\rho)\sigma^{s/2}$ and $\rho^{(1-s)/2}\sigma^{s/2}$.

Exercise 2.25. **a**: Use the Schwarz inequality. **c**: Choose U such that $|\rho^{1/2}\sigma^{1/2}| = U\rho^{1/2}\sigma^{1/2}$. Note that $|\operatorname{Tr} U\rho^{1/2}M_i\sigma^{1/2}| \geq \operatorname{Tr} U\rho^{1/2}M_i\sigma^{1/2}$.

Exercise 2.26. Note that the spectral decomposition $\sum_i \lambda M_i$ of $\rho^{1/2}U^*\sigma^{-1/2}$ satisfies $M_i^{1/2}\sigma^{1/2} = \lambda_i M_i^{1/2}\rho^{1/2}U^*$.

Exercise 2.27. **b**: See the hint for Exercise 2.26.

Exercise 2.31. Use the approximation $\sqrt{p_{\theta+\epsilon}(\omega)} \cong \sqrt{p_\theta(\omega)}\sqrt{1 + l_\theta(\omega)\epsilon + \frac{1}{2}\frac{d^2 p_\theta(\omega)}{d\theta^2}\epsilon^2}$.

Exercise 2.34. **c**: Use $\frac{d\theta}{d\eta} = J_\eta$. **g**: Use the uniqueness of the solution of the differential equation.

Exercise 2.35.
$$\begin{aligned}
&D(q_{\eta_j=X_j(\omega)}\|p_{\theta'}) - D(q_{\eta_j=X_j(\omega)}\|p_\theta) \\
&= \textstyle\sum_{\omega'} q_{\eta_j=X_j(\omega)}(\omega')\left(\left(\log q_{\eta_j=X_j(\omega)}(\omega') - \log p_{\theta'}(\omega')\right)\right. \\
&\quad \left. - \left(\log q_{\eta_j=X_j(\omega)}(\omega') - \log p_\theta(\omega')\right)\right) \\
&= \textstyle\sum_{\omega'} q_{\eta_j=X_j(\omega)}(\omega')\left(\log p_\theta(\omega') - \log p_{\theta'}(\omega')\right) \\
&= \textstyle\sum_{\omega'} q_{\eta_j=X_j(\omega)}(\omega')\left(\sum_i(\theta^i - \theta'^i)X_i(\omega') + \mu(\theta') - \mu(\theta')\right) \\
&= \left(\textstyle\sum_i(\theta^i - \theta'^i)X_i(\omega) + \mu(\theta') - \mu(\theta')\right) = \log p_\theta(\omega) - \log p_{\theta'}(\omega).
\end{aligned}$$

Exercise 2.36. Show that $\frac{p_\theta(\omega)}{d\theta} = 0$ if and only if $\eta(\theta) = X(\omega)$.

Exercise 2.37. Combine (2.9) and (2.77).

Exercise 2.38. The case of $n \geq m$ can be obtained from $n, n-1, \ldots, m+1 \geq m$. The $n < m$ case may be obtained from $\frac{1}{m}, \frac{1}{m-1}, \ldots, \frac{1}{n+1} \leq \frac{1}{n}$.

Exercise 2.41. **o**: Replace $H(p)$ and $\psi(s)$ with $-D(q\|p')$ and $\phi(s|p\|p')$, respectively.

Exercise 3.1. Note that $\operatorname{Tr}|X|$ is equal to the sum of the absolute values of the eigenvalues of X.

Exercise 3.3. $\|\rho_{\mathrm{mix}} - \rho\|_1 = 2\operatorname{Tr}(\rho_{\mathrm{mix}} - \rho)\{\rho_{\mathrm{mix}} - \rho \geq 0\}$. Hence, $2 - \|\rho_{\mathrm{mix}} - \rho\|_1 = 2 - 2\operatorname{Tr}(\rho_{\mathrm{mix}} - \rho)\{\rho_{\mathrm{mix}} - \rho \geq 0\} \leq 2 - 2\operatorname{Tr}\rho_{\mathrm{mix}}\{\rho_{\mathrm{mix}} - \rho \geq 0\} = 2\operatorname{Tr}\rho_{\mathrm{mix}}\{\rho_{\mathrm{mix}} - \rho < 0\} \leq 2\operatorname{Tr}\rho_{\mathrm{mix}}\{0 < \rho\} = 2\frac{\operatorname{rank}\rho}{d}$.

Exercise 3.4. $\sum_{i=1}^k \frac{1}{k}\operatorname{Tr}\rho_i M_i \leq \sum_{i=1}^k \frac{1}{k}\operatorname{Tr} M_i = \frac{1}{k}\operatorname{Tr}\sum_{i=1}^k M_i = \frac{1}{k}\operatorname{Tr} I = \frac{d}{k}$. Further, $\sum_{i=1}^k \frac{1}{k}\operatorname{Tr}\rho_i M_i \leq \sum_{i=1}^k \frac{1}{k}\|M_i\|_1\|\rho_i\| = \frac{1}{k}\max_{i'}\|\rho_{i'}\|\sum_{i=1}^k\|M_i\|_1 = \frac{1}{k}\max_{i'}\|\rho_{i'}\|\sum_{i=1}^k\operatorname{Tr} M_i = \frac{d}{k}\max_i\|\rho_i\|$.

Exercise 3.5. Use Cramér's theorem with $X = -\log \frac{p(\omega)}{\bar{p}(\omega)}$, $\theta = s$, $x = 0$, and show that $\lim \frac{-1}{n} \log p^n \left\{ \frac{-1}{n} \log \left(\frac{p^n(x^n)}{\bar{p}^n(x^n)} \right) \geq 0 \right\} = \max_{s \geq 0} -\phi(s)$ and $\lim \frac{-1}{n} \log \bar{p}^n \left\{ \frac{-1}{n} \log \left(\frac{p^n(x^n)}{\bar{p}^n(x^n)} \right) \leq R \right\} = \max_{s \leq 1} -\phi(s)$.

Exercise 3.9. Use the convexity of $\phi(s)$ and the symmetry $\phi(s) = \phi(1-s)$.

Exercise 3.11. **a**: For the derivations of (3.20) and (3.21), substitute $\theta = s$ and $X = -\log \left(\frac{p(x)}{\bar{p}(x)} \right)$ in (2.117) and (2.119), respectively. For the derivation of (3.22), substitute $\theta = s - 1$ $X = -\log \left(\frac{p(x)}{\bar{p}(x)} \right)$ in (2.119). **e**: Lemma 3.1 guarantees that the optimal test is always a likelihood test. When $R \leq \phi'(s_r)$, (3.22) $\geq r$, and when $R = \phi'(s_r)$, (3.20) $= \sup_{0 \leq s \leq 1} \frac{-sr - \phi(s|p\|\bar{p})}{1-s}$.

Exercise 3.12.

$$
\begin{aligned}
& D\left(\mathrm{P}_{\rho^{\otimes n}}^{\boldsymbol{M}^n} \middle\| \mathrm{P}_{\sigma^{\otimes n}}^{\boldsymbol{M}^n} \right) \\
&= \sum_{\omega_n} \left(\prod_{k=1}^n \operatorname{Tr} M_{k,\omega_n}^n \rho \right) \log \left(\prod_{k=1}^n \operatorname{Tr} M_{k,\omega_n}^n \rho \right) - \log \left(\prod_{k=1}^n \operatorname{Tr} M_{k,\omega_n}^n \sigma \right) \\
&= \sum_{\omega_n} \left(\prod_{k=1}^n \operatorname{Tr} M_{k,\omega_n}^n \rho \right) \sum_{k=1}^n \left(\log \operatorname{Tr} M_{k,\omega_n}^n \rho - \log \operatorname{Tr} M_{k,\omega_n}^n \sigma \right) \\
&= \sum_{\omega_n} \sum_{k=1}^n \operatorname{Tr} a_{k,\omega_n} M_{k,\omega_n}^n \rho \left(\log \operatorname{Tr} a_{k,\omega_n} M_{k,\omega_n}^n \rho - \log \operatorname{Tr} a_{k,\omega_n} M_{k,\omega_n}^n \sigma \right) \\
&= \sum_{k=1}^n \sum_{\omega_n} \operatorname{Tr} a_{k,\omega_n} M_{k,\omega_n}^n \rho \left(\log \operatorname{Tr} a_{k,\omega_n} M_{k,\omega_n}^n \rho - \log \operatorname{Tr} a_{k,\omega_n} M_{k,\omega_n}^n \sigma \right) \\
&= \sum_{k=1}^n D\left(\mathrm{P}_{\rho}^{\boldsymbol{M}^{n,k}} \middle\| \mathrm{P}_{\sigma}^{\boldsymbol{M}^{n,k}} \right).
\end{aligned}
$$

Exercise 3.13. **a**: Use the fact that $\{\sigma^{\otimes n} \leq e^{-na}\} \sigma^{\otimes n} \{\sigma^{\otimes n} \leq e^{-na}\} \leq \{\sigma^{\otimes n} \leq e^{-na}\} (\sigma^{\otimes n})^s e^{-n(1-s)a} \{\sigma^{\otimes n} \leq e^{-na}\}$ for (3.36). Apply Lemma 3.6 for (3.35). **b**: Use the fact that $\frac{-sr-\phi(s)}{1-s} = -(\phi(s) + sa)$ if $r = -(\phi(s) - (1-s)a)$. **d**: Show that $(\kappa_{\sigma^{\otimes n}}(\rho^{\otimes n}))^{-s} \leq (n+1)^d (\rho^{\otimes n})^{-s}$. **e**: See the hint for **b**.

Exercise 3.15. Lemma 3.5 guarantees that $\lim \operatorname{Tr} \tau^{\otimes n}(I - T_n) = 1$ because $D(\tau\|\sigma) \leq \underline{\lim} \frac{-1}{n} \log \operatorname{Tr} \sigma^{\otimes n}(I - T_n)$. Hence, applying Lemma 3.5 again, we have $\underline{\lim} \frac{-1}{n} \log \operatorname{Tr} \rho^{\otimes n} T_n \leq D(\tau\|\rho)$.

Exercise 3.16. **a**: Join equation (3.17) and inequality (3.40). **c**: Consider the case $r = r_s$ in **a**.

Exercise 3.17. **a**: Use (3.17) with $\kappa_{\sigma^{\otimes n}}$ and $\sigma^{\otimes n}$. **b**: Show that the function $s \mapsto \frac{-sr-\phi(s|p\|\bar{p})}{1-s}$ is monotone increasing in $(-\infty, s_r)$ and is monotone decreasing in (s_r, ∞). **c**: Use (3.43).

Exercise 4.1. Since $H(p|u\rangle\langle u|+(1-p)|v\rangle\langle v|) = H((1-p)|u\rangle\langle u|+p|v\rangle\langle v|)$, the concavity implies $H(1/2|u\rangle\langle u| + 1/2|v\rangle\langle v|) \geq H((1-p)|u\rangle\langle u| + p|v\rangle\langle v|)$. Show that the larger eigenvalue of $1/2|u\rangle\langle u| + 1/2|v\rangle\langle v|$ is $\frac{1+|\langle v|u\rangle|}{2}$.

Exercise 4.2.

$$
\begin{aligned}
&I(p_A, W^A) + I(p_B, W^B) - I(p, W^A \otimes W^B) \\
&\quad = \sum_{x_A,x_B} p_A(x_A)p_B(x_B) D(W^A_{x_A} \otimes W^B_{x_B} \| W^A_{p_A} \otimes W^B_{p_B}) \\
&\qquad - \sum_{x_A,x_B} p(x_A,x_B) D(W^A_{x_A} \otimes W^B_{x_B} \| W^A_{p_A} \otimes W^B_{p_B}) \\
&\qquad + \sum_{x_A,x_B} p(x_A,x_B) D(W^A_{x_A} \otimes W^B_{x_B} \| W^A_{p_A} \otimes W^B_{p_B}) \\
&\qquad - \sum_{x_A,x_B} p(x_A,x_B) D(W^A_{x_A} \otimes W^B_{x_B} \| \left(W^A \otimes W^B\right)_p) \\
&\quad = \sum_{x_A,x_B} (p_A(x_A)p_B(x_B) - p(x_A,x_B)) \left(D(W^A_{x_A} \| W^A_{p_A}) + D(W^B_{x_B} \| W^B_{p_B})\right) \\
&\qquad + \sum_{x_A,x_B} p(x_A,x_B) \Bigg(- \operatorname{Tr}\left(W^A_{x_A} \otimes W^B_{x_B}\right) \log \left(W^A_{p_A} \otimes W^B_{p_B}\right) \\
&\qquad + \operatorname{Tr}\left(W^A_{x_A} \otimes W^B_{x_B}\right) \log \left(W^A \otimes W^B\right)_p \Bigg) \\
&\quad = D\left(W^A_{p_A} \otimes W^B_{p_B} \,\middle\|\, \left(W^A \otimes W^B\right)_p\right) \geq 0.
\end{aligned}
$$

Exercise 4.3. The $\leq$ part in (4.19) follows from $I(\boldsymbol{M}, p, W) \leq I(p, W)$. Use the Fano inequality noting the definition of $C(W)$ for the proof of the $\geq$ part.

Exercise 4.4. Use the method of Lagrange multipliers.

Exercise 4.5. From $(A - cB)^*(A - cB) \geq 0$ we have $A^*B + B^*A \leq c^{-1}\, A^*A + c\, B^*B$.

Exercise 4.6. Consider the case of $c = \sqrt{\beta/(\alpha+\beta)}$.

Exercise 4.7. Apply Lemma 3.6 to the first term on the RHS of (4.35).

Exercise 4.9. Let $\nu_i = W_{\varphi(i)}$ in (4.39).

Exercise 4.10. Apply lemma 3.1 to the RHS of (4.35).

Exercise 4.11. **a**: Define $A = -nR$. **b**: Define $T_n = \{\kappa_{S^{\otimes n}}(R^{\otimes n}) - N_n S^{\otimes n} \geq 0\}$ in (4.30) and use the arguments in **c,d** of Exercise 3.13.

Exercise 4.12. Order the N_n signals from smallest to largest, and note that the error probability of the first $N_n/2$ signals is less than twice the average error probability.

Exercise 4.13. First, show that

$$\mathrm{P}_X\Big\{\varepsilon[\Phi_X] > M \sum_{x\in\mathcal{X}} p(x)\big(2\,\mathrm{Tr}\, W_x^i \{W_x^i - 2NW_p^i \le 0\} + 4N\,\mathrm{Tr}\, W_p^i \{W_x^i - 2NW_p^i > 0\}\big)\Big\} < \frac{1}{M}.$$

Use this inequality.

Exercise 4.15. **a**: Apply the Markov inequality to the uniform distribution on the message set $\{1, \ldots, N_n\}$.

Exercise 5.1. Let $\mathcal{H}_D$ be $\mathcal{H}_C \otimes \mathcal{H}_B$, and consider the unitary matrix corresponding to the replacement $W : u \otimes v \mapsto v \otimes u$ in $\mathcal{H}_A \otimes \mathcal{H}_B$, and define $V \stackrel{\mathrm{def}}{=} (W \otimes I_C)U$.

Exercise 5.3. Use Exercise 5.2.

Exercise 5.5. Consider the Hilbert space $\mathcal{H}_B$ produced by $|\omega\rangle$, and apply Condition ⑤ of Theorem 5.1 to the entanglement-breaking channel $W_\omega = |\omega\rangle\langle\omega|$. Finally, consider the measurement $\{|\omega\rangle\langle\omega| \otimes I_C\}$.

Exercise 5.17. When we choose the coordinate $u_1 = u_1^A \otimes u_1^B, u_2 = u_1^A \otimes u_2^B, u_3 = u_2^A \otimes u_1^B, u_4 = u_2^A \otimes u_2^B$, we have $\mathrm{Inv}_\lambda \otimes \iota_{\mathbb{C}^2}(|\Phi_2\rangle\langle\Phi_2|) = \begin{pmatrix} 1-\lambda & 0 & 0 & 1-2\lambda \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 1-2\lambda & 0 & 0 & 1-\lambda \end{pmatrix}$. This matrix is positive if and only if $(1-\lambda)^2 - (1-2\lambda)^2 \ge 0$, i.e., $\frac{2}{3} \ge \lambda \ge 0$.

Exercise 5.19. See the proof of $d_1(\rho,\sigma) \ge b^2(\rho,\sigma)$.

Exercise 5.20. Let $\boldsymbol{M}$ be a POVM that satisfies the equality in (2.60). Applying (2.18) to $\mathrm{P}_\rho^{\boldsymbol{M}}$ and $\mathrm{P}_\sigma^{\boldsymbol{M}}$, we obtain $d_1(\mathrm{P}_\rho^{\boldsymbol{M}}, \mathrm{P}_\sigma^{\boldsymbol{M}}) \ge b^2(\mathrm{P}_\rho^{\boldsymbol{M}}, \mathrm{P}_\sigma^{\boldsymbol{M}})$. Finally, adding (2.61), we obtain $d_1(\rho,\sigma) \ge b^2(\rho,\sigma)$.

Exercise 5.23. **c**: Apply the information-processing inequality of the quantum relative entropy to the two-valued measurement $\{P, I-P\}$.

Exercise 5.24. For any POVM $\boldsymbol{M} = \{M_i\}$ on $\mathcal{H}^{\otimes n}$, $\{(\kappa^{\otimes n})^*(M_i)\}$ is also a POVM. Using this fact, show that $B^*(r|\rho\|\sigma) \le B^*(r|\kappa(\rho)\|\kappa(\sigma))$. Next, choose r such that $B^*(r|\kappa(\rho)\|\kappa(\sigma)) = \frac{-sr-\bar{\phi}(s|\kappa(\rho)\|\kappa(\sigma))}{1-s}$ for any $s \le 0$.

Exercise 5.27. Let ρ_{mix} be a completely mixed state in $\mathcal{H}_B$. Consider the relative entropy $D(\rho_{A,B,C}\|\rho_{\mathrm{mix}} \otimes \rho_{A,C})$ and the partial trace of $\mathcal{H}_C$.

Exercise 5.30. Consider the state $\begin{pmatrix} p_1\rho_1 & & 0 \\ & \ddots & \\ 0 & & p_k\rho_k \end{pmatrix}$ in $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$.

Exercise 5.31. Let the purification of $\rho_{A,B}$ be $\rho_{A,B,C}$ for a reference state $\mathcal{H}_R$. From the subadditivity (5.58),

$$H(\rho_{A,B}) - H(\rho_A) + H(\rho_B) = H(\rho_C) - H(\rho_{B,C}) + H(\rho_B) \geq 0.$$

Exercise 5.29. **a**: Use (5.52).

Exercise 5.35. **a**: Differentiate both sides with respect to ϵ.

Exercise 5.36. **a**: Use **b** of Exercise 5.35. **b**: Note that $|\eta(a_1) - \eta(b_1)| \leq \eta(1/e) = 1/e$. **c**: Use $\frac{1}{2}\log 2 = 0.34658 < 1/e = 0.36792$.

Exercise 5.37. Note that $I(p, W) = \sum_x p(x)(H(W_p) - H(W_x))$ and η_0 is concave.

Exercise 5.43. Show this inequality when κ_A and κ_B are partial traces using (5.75). Next, show it in the general case.

Exercise 5.44. Use (8.47) and Lemma 3.7.

Exercise 6.4. Assume Condition ①. Since **d** of Exercise 6.1 implies $[\rho, A] = 0$, we have $A = \rho X' = X'\rho$ for a suitable X'. We can choose $A = \rho \circ X$ for a Hermitian matrix X that commutes with each M_i. Since $\rho \circ (X - X') = 0$, $\rho(X - X') = (X - X')\rho = 0$. Therefore, $A = \rho X = X\rho$.

Exercise 6.5. Note that there exists a unitary matrix U_m such that $\|X\rho Y\|_1 = \operatorname{Tr} X\rho Y U_m$.

Exercise 6.7. Letting $\sigma = \operatorname{Tr}_{\mathcal{H}'} |y\rangle\langle y|$, we have $(\|X \otimes I_{\mathcal{H}'}\|^{(e)}_{|y\rangle\langle y|,b})^2 = |\operatorname{Tr} X\sigma|^2 = |\langle I, X\rangle^{(e)}_{\sigma,b}|^2 \leq (\|X\|^{(e)}_{\sigma,b})^2$. The equality holds only when $X = I$.

Exercise 6.10. Apply Theorem 6.2 to the entanglement-breaking channel $\rho \mapsto \sum_i (\operatorname{Tr} M_i \rho)|u_i\rangle\langle u_i|$ with the CONS $\{u_i\}$.

Exercise 6.12. **a**: Combine **a** and **d** of Exercise 6.2. **b**: Combine Exercise 6.11 and **a**. **c**: Use the fact that $\kappa_{\boldsymbol{M},\rho_\theta,s}$ is a projection, as shown in Exercise 6.2.

Exercise 6.13. **b**: Use $\exp(X(\theta)) = \sum_{n=0}^{\infty} \frac{X(\theta)^n}{n!}$ and **a**.

Exercise 6.14. First, show that $\frac{d\rho_\theta^{\otimes n}}{d\theta} = \sqrt{n}\left(\frac{d\rho_\theta}{d\theta}\right)^{(n)}$. Use this and (6.9).

Exercise 6.16. Use the fact that $\boldsymbol{J}_{\theta,s;i,j} = \operatorname{Tr} L^*_{\theta,i,s} \frac{\partial \rho_\theta}{\partial \theta^j}$.

Exercise 6.18. Use $L_{\theta,b} = \frac{d \log \rho_\theta}{d\theta}$.

Exercise 6.19. Compare the e and m representations of the derivative of the quantum state family $\{\rho_\theta = e^{-i\theta Y}\rho e^{i\theta Y}\}$ in Exercise 6.18.

Exercise 6.20. **b**: Use (6.14) and (5.23).

Exercise 6.21. Consider the TP-CP map $\lambda\rho_\theta^1 \oplus (1-\lambda)\rho_\theta^2 \to \lambda\rho_\theta^1 + (1-\lambda)\rho_\theta^2$.

Exercise 6.23. First, for a given SLD geodesic $\mathit{\Pi}^\theta_{L,s}\sigma$, choose a unitary matrix U_1 such that ULU^* is equal to the constant times of S_1. Then, the SLD geodesic $U\mathit{\Pi}^\theta_{L,s}\sigma U^*$ has the form given in Exercise 6.21. Next, choose another unitary matrix U_2 such that

$$U_2 S_1 U_2^* = S_1,\ U_2(x_2 S_2 + x_3 S_3)U_2^* = \sqrt{x_2^2 + x_3^2} S_3. \tag{C.1}$$

Exercise 6.24. **b**: $\int_0^1 \mathrm{Tr}(\sigma-\rho)^2(\rho + t(\sigma-\rho))^{-1}dt$
$= \int_0^1 \mathrm{Tr}\,\rho(\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1} - I)(I + t(\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1} - I))^{-1}(\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1} - I)dt$
$= \mathrm{Tr}\,\rho\left((\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1} - I) - \log(I + (\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1} - I))\right)$
$= -\mathrm{Tr}\,\rho\log(\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1}) = \mathrm{Tr}\,\rho\log(\sqrt{\rho}\sigma^{-1}\sqrt{\rho})$.

Exercise 6.26. **a**: Use the partial integration formula twice. **c**: Use (6.22). **d**: Use **a**,**b**,**c** and the fact that $\frac{d^2\rho_\theta}{d\theta^2} = 0$.

Exercise 6.27. Show the equivalence of ② and ③ based on Exercise 6.4.

Exercise 6.28. Show that $\lim \frac{1}{n} D_x^{(m)}((\rho_1\otimes\rho_2)^{\otimes n}\|(\sigma_1\otimes\sigma_2)^{\otimes n}) =$ $\lim \frac{1}{n} D_x^{(m)}(\rho_1^{\otimes n}\|\sigma_1^{\otimes n}) + \lim \frac{1}{n} D_x^{(m)}(\rho_2^{\otimes n}\|\sigma_2^{\otimes n})$. Use (6.50) and (5.36).

Exercise 6.30. Consider a state family where $A = \frac{d\rho_\theta}{d\theta}$ and $\rho_{\theta_0} = \rho$, and let κ be given by a POVM M. From property (6.17) and $\|\rho^{-1}\|_{\rho,x}^{(m)} = 1$, we have $J_{\theta_0}^M = \|\kappa(A)\|_{\kappa(\rho),x}^{(m)} \le \|A\|_{\rho,x}^{(m)}$. Then, use Exercise 6.29.

Exercise 6.31. See the hint for Exercise 2.34.

Exercise 6.32. **a**: Let K be the difference K between $O(\boldsymbol{M},\hat\theta) - \theta$ and $\frac{1}{J_{\theta,s}} L_{\theta,s}$. Then, $K\rho + \rho K = 0$ when Condition ① holds. **b**: Use Condition (6.68).

Exercise 6.35. **b**: Define $B_0 \stackrel{\mathrm{def}}{=} \left\{\hat\theta \le \theta\right\}$, $B_i \stackrel{\mathrm{def}}{=} \left\{\theta + \frac{\delta(i-1)}{m} \le \hat\theta \le \theta + \frac{\delta i}{m}\right\}$, $(i = 1,\dots,m)$, and $B_{m+1} \stackrel{\mathrm{def}}{=} \left\{\theta + \delta \le \hat\theta\right\}$, and consider a POVM $M_i \stackrel{\mathrm{def}}{=} MB_i$ comprising $m+1$ measurements.

Exercise 6.38. **a**: Show that $\frac{\langle M(\omega), M(\omega)\rangle_{\rho_\theta,s}^{(e)}}{\langle M(\omega), I\rangle_{\theta,s}^{(e)}} = \mathrm{Tr}\,M(\omega)$. **c**: Use Exercise A.1. **d**: Use Exercise 2.37.

Exercise 6.39. If $(\boldsymbol{M},\hat\theta)$ is a locally unbiased estimator, then show that $\boldsymbol{V}_\theta(\boldsymbol{M},\hat\theta) \ge (\boldsymbol{J}_\theta^{\boldsymbol{M}})^{-1}$. Then, show that for each POVM $\boldsymbol{M}$ there exists a function $\hat\theta$ such that $(\boldsymbol{M},\hat\theta)$ is a locally unbiased estimator and $\boldsymbol{V}_\theta(\boldsymbol{M},\hat\theta) = (\boldsymbol{J}_\theta^{\boldsymbol{M}})^{-1}$.

Exercise 6.40. Use the method of Lagrange multipliers.

Exercise 6.41. Show that $\left(\|L(u)\|_{\rho_\theta,s}^{(e)}\right)^2 = \langle u|\boldsymbol{J}_{\theta,s}|u\rangle$. Then, show that $\langle x|\boldsymbol{J}_\theta^{\boldsymbol{M}^u}|x\rangle = \langle L(x)|\kappa_{\boldsymbol{M},\rho_\theta,s}|L(x)\rangle_{\rho_\theta,s}^{(e)} \ge \langle L(x)|\frac{|L(u)\rangle_{\rho_\theta,s}^{(e)}\langle L(u)|}{\langle u|\boldsymbol{J}_{\theta,s}|u\rangle}|L(x)\rangle_{\rho_\theta,s}^{(e)}$ for $x\in\mathbb{R}^d$.

Exercise 6.44. Let $u_1,\dots,u_d$ be the eigenvectors of $\boldsymbol{J}_{\theta,s}$, and let p_i be the eigenvalues of $\frac{1}{\mathrm{tr}\,\boldsymbol{J}_{\theta,s}^{-\frac12}}\boldsymbol{J}_{\theta,s}^{-\frac12}$. Then, the RHS of (6.94) is equal to $\frac{1}{\mathrm{tr}\,\boldsymbol{J}_{\theta,s}^{-\frac12}}\boldsymbol{J}_{\theta,s}^{\frac12}$.

Exercise 6.45. Choose the new coordinate such that the SLD Fisher information matrix is $\sqrt{G}^{-1}\boldsymbol{J}_{\theta,s}\sqrt{G}^{-1}$.

Exercise 6.52. **c**: See **b** of Exercise 6.50. **e**: Consider the estimator $(\boldsymbol{M}, \hat{\theta})$ in the extended system given in **b** for the set of vectors (y^i). Let P be the projection to the original space. Consider the POVM $\{PM_kP\}$ for $\boldsymbol{M} = \{M_k\}$. **g**: Note that the set of vectors (u^i) satisfying $\langle u^i|x_j\rangle = \delta^i_j$ is restricted to $u^i = ((\mathbf{Re}\, J)^{-1})^{i,j}x_j$.

Exercise 7.1. Consider the unitary matrix $\left(\sum_\omega |u_\omega\rangle\langle u_\omega| \otimes U_{\kappa'_\omega}\right)(I_{B,C} \otimes U)$ and $\rho_1 \otimes \rho_0$.

Exercise 7.2. Let $\mathcal{H}_D$ be $\mathcal{H}_C \otimes \mathcal{H}_B$, consider the unitary matrix corresponding to the replacement $W : u \otimes v \mapsto v \otimes u$ in $\mathcal{H}_A \otimes \mathcal{H}_B$, and define $V \stackrel{\rm def}{=} (W \otimes I_C)U$.

Exercise 7.3. Equation (5.4) yields that $\mathrm{Tr}\, \rho M'_{\omega,\omega} = \mathrm{Tr}\, \rho \kappa^*_\omega(M_{\omega'}) = \mathrm{Tr}\, \kappa^*_\omega(\rho) M_{\omega'}$.

Exercise 7.5. **b:** Let $\sum_i p_i \rho_i^A \otimes \rho_i^B = (\kappa \otimes \iota_{A'})(|x_M\rangle\langle x_M|)$. Then, $K(\kappa) = d\sum_i p_i \rho_i^A \otimes \rho_i^B$ from (5.3). From the definition of $K(\kappa)$, we have $\mathrm{Tr}\, \kappa(\rho)\sigma = \mathrm{Tr}\, K(\kappa)\rho \otimes \sigma = d\sum_i p_i (\mathrm{Tr}\, \rho\rho_i^A)(\mathrm{Tr}\, \sigma \rho_i^B)$. Thus, we obtain $\kappa(\rho) = d\sum_i p_i (\mathrm{Tr}\, \rho\rho_i^A)\rho_i^B$.

Exercise 7.7. Expand the RHS of (7.17) and rearrange.

Exercise 7.8. Similarly, expand the RHS of (7.17) and rearrange.

Exercise 7.13. Use the fact that $\Delta_1(O(\boldsymbol{M}) - X, \rho) \le \Delta_3(\boldsymbol{M}, X, \rho)$, problem 7.11, and $\sqrt{x^2 + y^2} \le x + y$ for $x, y \ge 0$.

Exercise 7.14. Choose a 2×2 orthogonal matrix $(a_{i,j})$ such that the two matrices $\tilde{X} \stackrel{\rm def}{=} a_{1,1}X + a_{1,2}Y$ and $\tilde{Y} \stackrel{\rm def}{=} a_{2,1}X + a_{2,2}Y$ satisfy $\mathrm{Cov}_\rho(\tilde{X}, \tilde{Y}) = 0$. Show (7.36) for $\tilde{X}, \tilde{Y}$. Finally, use the fact that both sides of (7.36) are invariant under the orthogonal matrix transformation $(X, Y) \mapsto (\tilde{X}, \tilde{Y})$.

Exercise 7.15. **g**: Note that both sides of (7.38) become their $\det(b_{i,j})$ times value when we perform the transformation $(\boldsymbol{x}, \boldsymbol{y}) \mapsto (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})$.

Exercise 7.16. Show that $\Delta_3(\boldsymbol{M}_{X,Y,\rho}, O^1(\boldsymbol{M}_{X,Y,\rho}), \rho) = \frac{1-p}{p}\Delta_1(X, \rho)$, $\Delta_3(\boldsymbol{M}_{X,Y,\rho}, O^2(\boldsymbol{M}_{X,Y,\rho}), \rho) = \frac{1-q}{q}\Delta_1(Y, \rho)$, and note that $\frac{1-p}{p}\frac{1-q}{q} = 1$.

Exercise 7.19. Use Exercise 6.21.

Exercise 7.20. Note that the case of $p = 1$ has been shown in the main body. Use Exercise 7.19.

Exercise 8.2. Note that d_1 is monotone concerning the partial trace.

Exercise 8.3. Let u and v be purifications of ρ and σ such that $F(\rho, \sigma) = F(|u\rangle\langle u|, |v\rangle\langle v|) = \mathrm{Tr}\sqrt{|u\rangle\langle u|}\sqrt{|v\rangle\langle v|} = |\langle u|v\rangle|^2 = F^2(|u\rangle\langle u|, |v\rangle\langle v|)$. Using the monotonicity of $\phi(1/2, \rho, \sigma)$, we have $F^2(\rho, \sigma) \le \mathrm{Tr}\sqrt{\rho}\sqrt{\sigma}$.

In addition, $F^2(\rho, \sigma) = \mathrm{Tr}\,|\sqrt{\rho}\sqrt{\sigma}| \ge \mathrm{Tr}\sqrt{\rho}\sqrt{\sigma}$.

Exercise 8.4. $F_e^2(\rho, \kappa) = \sum_i \langle x|E_i \otimes I|x\rangle\langle x|E_i \otimes I|x\rangle$.

Exercise 8.5. **a**: Consider the singular value decomposition of the matrix $\{\mathrm{Tr}\, E_i A_j \rho\}_{i,j}$, and retake the Choi–Kraus representation. Use Exercise 5.2.

b: Applying **a** and defining $p_i \stackrel{\rm def}{=} \mathrm{Tr}\, A_i \rho A_i$, $A'_i \stackrel{\rm def}{=} A_i/\sqrt{p_i}$, we have $F_e^2(\rho, \kappa \circ$

$\kappa') = \sum_i p_i |\operatorname{Tr} E_i A_i' \rho|^2$. **c**: Note that $EA = (U|E|^{1/2})(|E|^{1/2}A)$, and use the Schwarz inequality for the inner product $\operatorname{Tr} X^* Y \rho$. **d**: Note that $F_e^2(\rho, \kappa \circ \kappa_U) \geq |\operatorname{Tr} E_1 U^* \rho|^2$.

Exercise 8.6. Show that $|\operatorname{Tr}(A_1 + A_2 i)\rho|^2 = (\operatorname{Tr} A_1 \rho)^2 + (\operatorname{Tr} A_2 \rho)^2$, where A_1 and A_2 are Hermitian matrices.

Exercise 8.12. Use ② or (8.19).

Exercise 8.14. Denote the input and output system of κ by $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Let $|x\rangle\langle x|$ be a purification of ρ_{mix}. Choose the probalistic decomposition as $\kappa \otimes \iota_R(|x\rangle\langle x|) = \sum_i p_i |y_i\rangle\langle y_i|$. Then, the Schmidt rank of $|y_i\rangle$ is less than d'. That is, the rank of $\operatorname{Tr}_B |y_i\rangle\langle y_i|$ is less than d' Therefore, $\operatorname{Tr}_R \sqrt{\operatorname{Tr}_A |y_i\rangle\langle y_i|} \leq \sqrt{d'}$. Thus,

$$\begin{aligned}
&\langle x|(\kappa' \circ \kappa) \otimes \iota_R(|x\rangle\langle x|)|x\rangle = \langle x| \sum_i p_i (\kappa' \otimes \iota_R)(|y_i\rangle\langle y_i|)|x\rangle \\
\leq &\sum_i p_i F^2(\operatorname{Tr}_A(\kappa' \otimes \iota_R)(|y_i\rangle\langle y_i|), \operatorname{Tr}_A |x\rangle\langle x|) \\
= &\sum_i p_i F^2(\operatorname{Tr}_A |y_i\rangle\langle y_i|, \rho_{\mathrm{mix},R}) = \sum_i p_i (\operatorname{Tr}_R |\sqrt{\operatorname{Tr}_A |y_i\rangle\langle y_i|}\sqrt{\rho_{\mathrm{mix},R}}|)^2 \\
= &\sum_i p_i \frac{(\operatorname{Tr}_R \sqrt{\operatorname{Tr}_A |y_i\rangle\langle y_i|})^2}{d} \leq \frac{d'}{d}.
\end{aligned}$$

Exercise 8.15. Since the final state on $\mathcal{H}_R \otimes \mathcal{H}_E \otimes \mathcal{H}_B$ is a pure state, $H(\rho)$ is equal to the entropy of the final state on the reference system $\mathcal{H}_R$. $H(\kappa(\rho))$ is equal to the entropy of the final state on $\mathcal{H}_R \otimes \mathcal{H}_E$. $H_e(\rho, \kappa)$ is therefore equal to the entropy of the final state on the environment $\mathcal{H}_E$.

Exercise 8.16. Note the second inequality in (5.42) and the monotonicity of the trace norm concerning the partial trace on the reference system.

Exercise 8.17. Note that the entropy of $U(\rho \otimes |u\rangle\langle u|)U$ is equal to the entropy of ρ.

Exercise 8.20. Consider the Stinespring represntation of κ', and consider the partial trace with respect to the environment of κ' after the state evolution.

Exercise 8.22. Since x' is a pure state on $\mathcal{H}_{A'} \otimes \mathcal{H}_{E'} \otimes \mathcal{H}_R$, $H_{x'}(A'R) = H_{x'}(E'), H_{x'}(R) = H_{x'}(A'E')$.

Exercise 8.25. **b**: Use (5.52) for the last inequality in (8.54).

Exercise 8.28. Use the concavity of the entropy and (5.52) for the pinching of a PVM that commutes with $|u\rangle\langle u|$.

Exercise 8.29. Consider the unitary matrix

$$\begin{pmatrix} S_0 & 0 & 0 & 0 \\ 0 & S_1 & 0 & 0 \\ 0 & 0 & S_2 & 0 \\ 0 & 0 & 0 & S_3 \end{pmatrix} \begin{pmatrix} \sqrt{p_0} I & * & * & * \\ \sqrt{p_1} I & * & * & * \\ \sqrt{p_2} I & * & * & * \\ \sqrt{p_3} I & * & * & * \end{pmatrix}$$

as a Stinespring representation in $\mathbb{C}^2 \otimes \mathbb{C}^4$, where the elements $*$ of the second matrix are choosen appropriately to preserve unitarity.

Exercise 8.30. Note that this map is a double stochastic matrix.

Exercise 8.32. This follows immediately from Exercise 8.31.

Exercise 8.33. Use Theorem 2.3.

Exercise 8.35. Use inequality (2.140). Note that $\frac{-1}{n} \log L(\kappa_n) = R$ in (8.78).

Exercise 8.36. Put $1 - t = s$ and $\sigma = \sum_i \lambda_i |u_i^A \otimes u_i^B\rangle\langle u_i^A \otimes u_i^B|$ in the proof of Theorem 8.7.

Exercise 8.38. First, consider the case of a pure state in $\{v \otimes u - u \otimes v | u, v \in \mathbb{C}^3\}$.

Exercise 8.39. Show the pure-state case using Theorem 8.4. Next, extend this fact to the mixed-state case.

Exercise 8.40. Show that the RHS of (8.85) in Theorem 8.8 approaches 0 exponentially when $R > E(\rho)$, $L = [e^{nR}]$.

Exercise 8.41. Let $\rho = \sum_i p_i |x_i\rangle\langle x_i|$, and consider a separable state σ_i, where $E_f(|x_i\rangle\langle x_i|) = D(|x_i\rangle\langle x_i| \| \sigma_i)$. Use the joint convexity of the relative entropy.

Exercise 8.42. **a:** First, show that $\| |x_n\rangle\langle x_n| - |y_n\rangle\langle y_n| \|_1 \to 0$ when $\|\rho_n - \sigma_n\|_1 \to 0$. Next, give a probabilistic decomposition using a POVM $\boldsymbol{M}^n = \{M_i^n\}$ on the reference system $\mathcal{H}_R$ (Lemma 8.3). Let $\boldsymbol{M}$ be the POVM giving the decomposition minimizing the average entropy of σ_n. Then, the average entropy $\sum_x p_x^i H(\mathrm{Tr}_B \rho_x^i)$ on $\mathcal{H}_A$ is equal to $H_{\hat{\kappa}_{\boldsymbol{M}^n} \otimes \iota_A(\mathrm{Tr}_B |x_n\rangle\langle x_n|)}(A|R)$. From monotonicity, show that $\|\hat{\kappa}_{\boldsymbol{M}^n} \otimes \iota_A(\mathrm{Tr}_B |x_n\rangle\langle x_n|) - \hat{\kappa}_{\boldsymbol{M}^n} \otimes \iota_A(\mathrm{Tr}_B |y_n\rangle\langle y_n|)\|_1 \le \|\rho_n - \sigma_n\|_1$. Finally, use (5.71).

Exercise 8.43. **a:** First, show that $\| |x_n\rangle\langle x_n| - |y_n\rangle\langle y_n| \|_1 \to 0$ when $\|\rho_n - \sigma_n\|_1 \to 0$. Next, choose a system $\mathcal{H}_E$ as a subsystem of an extended space of the reference system $\mathcal{H}_R$. Note that extenstions of ρ_n and σ_n can be given as the reduced densities on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E$. Now, we choose the subsystem $\mathcal{H}_E$ giving $E_{sq}(\sigma_n)$. Finally, use (5.73).

Exercise 8.46.

$$\begin{aligned} I(AA' : BB'|E) &= I(A : BB'|E) + I(A : BB'|EA) \\ &= I(A : B|E) + I(A : B'|BE) + I(A : B'|EA) + I(A : B|EAB') \\ &\ge I(A : B|E) + I(A : B'|BE). \end{aligned}$$

Exercise 8.52. First, show that $I^{A \to B}(|u\rangle\langle u|) \le H(\mathrm{Tr}_A |u\rangle\langle u|)$. Next, show its equality.

Exercise 8.57. **a**: Use Theorem 5.8. **b**: $I_{(\kappa_{\boldsymbol{M}} \otimes \iota_{AB})(\rho^{ABE})}(AB : E) \le I_{\rho^{ABE}}(AB : E)$.

Exercise 9.3. Refer to the proof of Theorem 4.2, and use Fano's inequality for the converse part of the theorem.

Exercise 9.4. Using the monotonicity of the quantum relative entropy and (5.58), we obtain

$$\begin{aligned}
I(X:Y) &\le \frac{1}{N_n}\sum_{i=1}^{N_n} D\Big(\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n})\Big\|\frac{1}{N_n}\sum_{i=1}^{N_n}\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n})\Big)\\
&= H\Big(\frac{1}{N_n}\sum_{i=1}^{N_n}\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n})\Big) - \frac{1}{N_n}\sum_{i=1}^{N_n} H\Big(\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n})\Big)\\
&\le H\Big(\operatorname{Tr}_{A_n'}\frac{1}{N_n}\sum_{i=1}^{N_n}\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n})\Big)\\
&\quad + H\Big(\operatorname{Tr}_B\frac{1}{N_n}\sum_{i=1}^{N_n}\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n})\Big) - \min_{\kappa} H((\kappa\otimes\iota_B^{\otimes n})(\rho_{A,B}^{\otimes n})).
\end{aligned}$$

Using

$$\begin{aligned}
\operatorname{Tr}_{A_n'}\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n}) &= \operatorname{Tr}_A\rho_{A,B}^{\otimes n}\\
H\Big(\operatorname{Tr}_B\frac{1}{N_n}\sum_{i=1}^{N_n}\Big(\varphi_e^{(n)}(i)\otimes\iota_B^{\otimes n}\Big)(\rho_{A,B}^{\otimes n})\Big) &\le \log\dim\mathcal{H}_{A_n'},
\end{aligned}$$

we obtain

$$\begin{aligned}
I(X:Y) \le& H\big(\operatorname{Tr}_A\rho_{A,B}^{\otimes n}\big) + \log\dim\mathcal{H}_{A_n'} - \min_{\kappa} H((\kappa\otimes\iota_B^{\otimes n})(\rho_{A,B}^{\otimes n}))\\
=& nH\left(\operatorname{Tr}_A\rho_{A,B}\right) + \log\dim\mathcal{H}_{A_n'} - \min_{\kappa} H((\kappa\otimes\iota_B^{\otimes n})(\rho_{A,B}^{\otimes n})).
\end{aligned}$$

Exercise 9.5. The LHS of (9.42) can be rewritten as

$$\begin{aligned}
&\sum_j p_j D((\kappa\circ\varphi_e(j)\otimes\iota_R)(\rho_{A',R})\|\sum_j p_j(\kappa\circ\varphi_e(j)\otimes\iota_R)(\rho_{A',R}))\\
&\quad\le \sum_j p_j D((\kappa\circ\varphi_e(j)\otimes\iota_R)(\rho_{A',R})\|(\sum_j p_j(\kappa\circ\varphi_e(j))\rho_{A'})\otimes\rho_R)\\
&\quad= H(\sum_j p_j(\kappa\circ\varphi_e(j))\rho_{A'}) + H(\rho_R) - \sum_j p_j H((\kappa\circ\varphi_e(j)\otimes\iota_R)\rho_{A',R})\\
&\quad= H(\kappa(\sum_j p_j\varphi_e(j)(\rho_{A'})) + \sum_j p_j\tilde{I}_c(\rho_{A'},\kappa\circ\varphi_e(j))\\
&\quad\le H(\kappa(\sum_j p_j\varphi_e(j)(\rho_{A'}))) + \tilde{I}_c(\sum_j p_j\varphi_e(j)(\rho_{A'}),\kappa)\\
&\quad= I(\sum_j p_j\varphi_e(j)(\rho_{A'}),\kappa),
\end{aligned}$$

from (4.6) and (8.50). Since $\rho_{A',R}$ is a pure state, we may write $H(\rho_{A'}) = H(\rho_R)$.

Exercise 9.8. First, show that

$$J(\rho,\sigma,\kappa) \le J\Big(\int_\theta U_\theta \rho U_\theta^* d\theta, \sigma, \kappa\Big) = J(P_1\rho P_1 + P_2\rho P_2, \sigma, \kappa),$$

where $U_\theta = P_1 + e^{i\theta}P_2$.

Next, show that

$$J(\lambda\rho_1 \oplus (1-\lambda)\rho_2, \sigma, \kappa) = \lambda J(\rho_1,\sigma,\kappa) + (1-\lambda)J(\rho_2,\sigma,\kappa).$$

Finaly, using (9.40), we obtain

$$\begin{aligned}
\max_\rho I(\rho,\kappa) &= \max_\rho \min_\sigma J(\rho,\sigma,\kappa) = \min_\sigma \max_\rho J(\rho,\sigma,\kappa) \\
&= \min_\sigma \max_{\lambda,\rho_1,\rho_2} J(\lambda\rho_1 \oplus (1-\lambda)\rho_2, \sigma, \kappa) \\
&= \min_\sigma \max_{\lambda,\rho_1,\rho_2} \lambda J(\rho_1,\sigma,\kappa) + (1-\lambda)J(\rho_2,\sigma,\kappa) \\
&\le \max_{\lambda,\rho_1,\rho_2} \lambda J(\rho_1,\kappa_1(\rho_{\max,1}),\kappa) + (1-\lambda)J(\rho_2,\kappa_1(\rho_{\max,1}),\kappa) \\
&= \max_\lambda \lambda C^e_{c,e}(\kappa_1) + (1-\lambda)C^e_{c,e}(\kappa_2).
\end{aligned}$$

Exercise 9.9. Applying (2.63) to the two-valued POVM $\{\{\kappa_{W_p}(W_x) - CW_p \ge 0\}, \{\kappa_{W_p}(W_x) - CW_p < 0\}\}$, we obtain

$$\begin{aligned}
&\big(\operatorname{Tr} W_x\{\kappa_{W_p}(W_x) - CW_p \ge 0\}\big)^{1-s}\big(\operatorname{Tr} W_p\{\kappa_{W_p}(W_x) - CW_p \ge 0\}\big)^s \\
&\quad + \big(\operatorname{Tr} W_x\{\kappa_{W_p}(W_x) - CW_p < 0\}\big)^{1-s}\big(\operatorname{Tr} W_p\{\kappa_{W_p}(W_x) - CW_p < 0\}\big)^s \\
&\le \operatorname{Tr} W_x^{1-s}W_p^s
\end{aligned}$$

for $0 \ge s$. Since $\operatorname{Tr} W_x\{\kappa_{W_p}(W_x) - CW_p \ge 0\} = \operatorname{Tr} \kappa_{W_p}(W_x)\{\kappa_{W_p}(W_x) - CW_p \ge 0\} \ge C \operatorname{Tr} W_p\{\kappa_{W_p}(W_x) - CW_p \ge 0\}$, we have

$$\begin{aligned}
&\operatorname{Tr} W_x\{\kappa_{W_p}(W_x) - CW_p \ge 0\} \\
&\quad \le C^s\big(\operatorname{Tr} W_x\{\kappa_{W_p}(W_x) - CW_p \ge 0\}\big)^{1-s}\big(\operatorname{Tr} W_p\{\kappa_{W_p}(W_x) - CW_p \ge 0\}\big)^s,
\end{aligned}$$

and thus we obtain (9.53).

Exercise 9.10. Solve $\frac{-\phi(s)+s(R-r)}{2} = \frac{r}{2}$ with respect to r.

Exercise 9.11. In this case, we can replace (9.56) by $\|P_x W_x\|_1 \le \operatorname{Tr} W_x P_x$.

Exercise 9.13. Use the fact that

$$\begin{aligned}
&I(p, W_B) + I(p', W'_B) - I(q, W_B \otimes W'_B) \\
&\quad = D\left(\sum_{x,x'} q(x,x') W_{B,x} \otimes W'_{B,x'} \,\middle\|\, \left(\sum_x p(x) W_{B,x}\right) \otimes \left(\sum_{x'} p'(x') W'_{B,x'}\right)\right).
\end{aligned}$$

Exercise 9.17. First, note that

$$\begin{aligned}\varepsilon_{E,a}[\Phi] &= \sum_i \sum_{j\neq i} \frac{1}{M(M-1)} d_1((W^E Q)_i, (W^E Q)_j) \\ &\geq \frac{1}{M}\sum_i d_1((W^E Q)_i, \frac{1}{M}\sum_{j=1}(W^E Q)_j).\end{aligned}$$

From Fannes' inequality (5.64) and the concavity and monotonicity (Exercise 5.35) of η_0, we have

$$\begin{aligned}I_E(\Phi) &= \frac{1}{M}\sum_{i=1}^{M} H(\frac{1}{M}\sum_{j=1}(W^E Q)_j)) - H(W^E Q)_i) \\ \leq &\frac{1}{M}\sum_{i=1}^{M} |H(\frac{1}{M}\sum_{j=1}(W^E Q)_j)) - H(W^E Q)_i)| \\ \leq &\frac{1}{M}\sum_{i=1}^{M} d_1\left((W^E Q)_i, \frac{1}{M}\sum_{j=1}(W^E Q)_j\right)\log d \\ &+ \eta_0\left(d_1\left((W^E Q)_i, \frac{1}{M}\sum_{j=1}(W^E Q)_j\right)\right) \\ \leq &\varepsilon_{E,a}[\Phi]\log d + \eta_0(\varepsilon_{E,a}[\Phi]).\end{aligned}$$

Exercise 9.18. First, show that there exists a code such that $\sum_i \frac{1}{M} d_1((W^E Q)_i, \frac{1}{M}\sum_{j=1}(W^E Q)_j)$ converges to 0. Next, choose the smallest $M/2$ values $i_1, \ldots, i_{M/2}$ concerning $d_1((W^E Q)_i, \frac{1}{M}\sum_{j=1}(W^E Q)_j)$. Finally, show this exercise based on the fact that

$$\sup_i \sup_j d_1((W^E Q)_i, (W^E Q)_j) \leq 2 \sup_i d_1((W^E Q)_i, \frac{1}{M}\sum_{j=1}(W^E Q)_j).$$

Exercise 9.20. Consider, for example, $a = 1/2$.

Exercise A.4. Compare the maximizations in definition (A.13) of $\|X\|_1$ and $\|\operatorname{Tr}_B X\|_1$.

Exercise A.6. If A possesses the inverse, use Exercise A.5. If A does not possess the inverse, choose an invertible matrix approximating A.

Exercise A.7. **a**: Use $B^{-1/2}AB^{-1/2} = (A^{1/2}B^{-1/2})^*(A^{1/2}B^{-1/2})$ and (1.29). **d**: Consider $B + \epsilon I$ and take the limit $\epsilon \to 0$.

Exercise A.10. Consider the $(d-k+1)$-dimensional subspace $\mathcal{K}'$ spanned by the eigenvectors according to the eigenvalues $a_k, \ldots, a_d$. Evaluate $\langle x|A|x\rangle$ when $\|x\| = 1$ and $x \in \mathcal{K} \cap \mathcal{K}'$.

Exercise A.11. Apply Exercise A.10.

Exercise A.12. Apply Exercise A.11.

Postface to Japanese Version

My research on quantum information theory started in October of 1994, when I was a first year master's student. At that time, although Shor's paper on factorization had already been published, I was still unaware of his work. Nor was the field of quantum information theory very well known at that time. The following is a brief summary of how I began working in the field of quantum information theory. Although this is merely a personal account of my experiences, it might be of some interest to those thinking of starting postgraduate studies and pursuing a career in research.

I began my university studies at Kyoto University by studying both mathematics and physics, thanks to their policy that allows students to complete their graduation without electing their major. Although I was mainly interested in physics, I decided to study both physics and mathematics because I was not entirely comfortable with the common thinking manner in physics; I was more naturally inclined towards a mathematical way of thought. As a result, during my undergraduate years although I had a reasonable understanding of mathematics, I could not understand physics sufficiently. Particularly, I could not catch the essence of statistical mechanics, in which "physics thinking" appears most prominently, more seriously. In my fourth year of my undergraduate degree, I realized that based on my understanding of physics at the time, I would probably not pass the entrance exams for a postgraduate course in physics. Therefore, I decided to apply for a postgraduate course in mathematics (which I just barely passed). In particular, while I elected the early universe in the cosmology group to my main research in the undergraduate course, its outcome was rather hopeless due to my poor knowledge of statistical mechanics. In fact, when I told a professor of physics that I would work as a casual teacher of cram school for high school physics in the next year, he said to me "I would never let you teach physics to anyone." Indeed, my physics knowledge was lacking in such a extent at that time. This was a particularly depressing event for me. Although I was still able to graduate, it had hardly felt like a time for celebration.

The following April, I began my postgraduate studies in twistor theory [411][1] under Professor Ueno, who is a professor in mathematics in Kyoto University. I chose this topic because it is related to relativity theory, which I was interested in at that time. However, as is the case with many topics in mathematical physics, although its basic is rooted in physics, it was essentially mathematical. I also found it very difficult to understand the physics behind this mathematical concepts. Ultimately, I realized that it did not suit my interests, and therefore I searched for another topic. Although I was capable of thinking in a mathematical way, I was not interested in mathematics itself. Therefore it was impossible for me to concentrate the research only of pure mathematics. Meanwhile, by teaching high school physics as a casual teacher in cram school during my postgraduate years, I felt that I could truly understand physics for the first time. Since I usually concerned difficult mathematical structure in physics, I realized for the first time that it is important to understand physics based on fundamental concepts.

When I searched for my new research topic, I met Dr. Akio Fujiwara, who came to Osaka University as an assistant professor at that time. He advised me to study Holevo's textbook [216], and I decided that I would start research in quantum information theory. Until this point, I had mainly studied abstract mathematics with little physical connection. I was particularly impressed by the way Holevo's textbook expressed the fundamental concepts of quantum mechanics without high levels of abstraction. although Holevo's textbook is not a particularly easy book to read from the current viewpoint, it was not so difficult for me to read because I had read more difficult books on mathematics.

In retrospect, it might be fortunate that I did not proceed to a postgraduate course in physics because physics community had an implicit strong stress to never try the measurement problem in quantum mechanics due to its philosophical aspect in Japan at that time. Therefore, while I appeared to take a rather indirect path during my years for my undergraduate and master courses, my career may have been the most direct path.

However, I faced a problem at starting my research. Since I had only studied physics and mathematics until this point, I was completely ignorant of subjects in information science such as mathematical statistics. In particular, despite having the opportunity to take these subjects, I had not studied these at all. During my undergraduate years, I regarded statistics to be a rather lightweight subject, as compared with physics, which examines the true nature of reality. I considered statistics to be only a convenient subject not an essential subject. This perception has been changed on reading Holevo's text. The reason is that it is impossible to quantitatively evaluate the information obtained by an observer without a statistical viewpoint because the measurement data is inherently probabilistic under the mathematical formulation of quantum mechanics. Ultimately, I was forced to study subjects such as math-

[1] Professor Richard Jozsa also studied twistor theory in his graduate course.

ematical statistics and information theory, which should be studied during the undergraduate course. In the end, the research for my master's thesis has been completed with a rather insufficient knowledge of mathematical statistics.

Further, as another problem, I lacked researchers to discuss my research, nearby, due to my choose of this research area. Hence, I had to arrange the chance to discuss with researchers located long distances away. Moreover, since I was also financially quite unstable during the first half of my doctor course, I had kept my research time only between casual teaching work at high school and cram school at that time. In particular, in the first six months of my doctoral course, my research progress was rather slow due to little opportunity for discussion about my research interest. Henceforth, the Quantum Computation Society in Kansai opened in November 1996, and it gave me a chance to talk about topics closely related to my interest. Hence, I could continue my research. During this period, I also had many helpful discussions via telephone with Keiji Matsumoto, who was a research associate at University of Tokyo at that time. Thus, I could learn statistics, and I am deeply indebted to him. I am also grateful to Professor Kenji Ueno, who accepted me as a graduate student until my employment at RIKEN.

In retrospect, in less than 10 years, the situation around quantum information theory has changed completely in Japan. The following are my thoughts and opinions on the future of quantum information theory.

Recently, sophisticated quantum operations have become a reality, and some quantum protocols have been realized. I believe that it is necessary to propose protocols that are relatively easy to implement. This is important not only to motivate further research, but also to have some feedback for the foundations of physics. In particular, I believe that the techniques developed in information theory via quantum information theory will be useful to the foundations of physics.

Thanks to the efforts of many researchers, the field of quantum information theory has become a quite well-known field. However, I feel that many universities in Japan have a trouble to internalize quantum information theory in the current organization of disciplines. Of course, scientific study should have no boundaries in themselves. Hence, It is my presumption that we can construct a more constructive research and educational environment through the treatment for fields such as quantum information theory, which transcend the current framework of disciplines.

My hope is that this book will make those unsatisfied with existing fields to be interested in quantum information theory and inspire them to become active researchers in this or any of its associated fields.

References

1. A. Abeyesinghe, I. Devetak, P. Hayden, A. Winter, "Quantum coherent Slepian-Wolf coding," in preparation.
2. A. Acín, E. Jané, G. Vidal, "Optimal estimation of quantum dynamics," *Phys. Rev. A*, **64**, 050302(R) (2001); quant-ph/0012015 (2000).
3. C. Adami, N. J. Cerf, "On the von Neumann capacity of noisy quantum channels," *Phys. Rev. A*, **56**, 3470 (1997); quant-ph/9610005 (1996).
4. D. Aharonov, A. Kitaev, N. Nisan, "Quantum Circuits with Mixed States," *Proceedings of the 30th Annual ACM Symposium on Theory of Computation (STOC)*, 20–30 (1997); quant-ph/9806029 (1998).
5. R. Ahlswede, G. Dueck, "Identification via channels," *IEEE Trans. Inf. Theory*, **35**, 15–29 (1989).
6. R. Ahlswede, A. Winter, " Strong converse for identification via quantum channels," *IEEE Trans. Inf. Theory*, **48**, 569–579 (2002); quant-ph/0012127 (2000).
7. C. Ahn, A. Doherty, P. Hayden, A. Winter, "On the distributed compression of quantum information," quant-ph/0403042 (2004).
8. R. Alicki, M. Fannes, "Continuity of quantum mutual information," quant-ph/0312081 (2003).
9. R. Alicki, M. Fannes, "Note on multiple additivity of minimal Renyi entropy output of the Werner–Holevo channels," quant-ph/0407033 (2004).
10. S. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics, Vol. 28 (Springer, Berlin Heidelberg New York, 1985).
11. S. Amari, H. Nagaoka, *Methods of Information Geometry*, (AMS & Oxford University Press, Oxford, 2000).
12. H. Araki, E. Lieb, "Entropy inequalities," *Comm. Math. Phys.*, **18**, 160–170 (1970).
13. S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inf. Theory*, **18**, 14–20 (1972).
14. E. Arthurs, M. S. Goodman, "Quantum correlations: a generalized Heisenberg uncertainty relation," *Phys. Rev. Lett.*, **60**, 2447–2449 (1988).
15. E. Arthurs, J. L. Kelly, Jr., "On the simultaneous measurement of a pair of conjugate observables," *Bell Syst. Tech.*, **44**, 725–729 (1965).

16. L. M. Artiles, R. D. Gill, M. I. Guţă, "An invitation to quantum tomography (II)," *J. R. Stat. Soc. Ser. B*, (Stat. Methodol.), **67**, 109–134 (2005); math.ST/0405595 (2004).
17. K. M. R. Audenaert, S. L. Braunstein, "On strong superadditivity of the entanglement of formation," *Comm. Math. Phys.*, **246**(3), 443–452 (2004); quant-ph/0303045 (2003).
18. K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, B. De Moor, "The asymptotic relative entropy of entanglement," *Phys. Rev. Lett.*, **87**, 217902 (2001); quant-ph/0103096 (2001).
19. K. Audenaert, M. B. Plenio, J. Eisert, "Entanglement cost under positive-partial-transpose-preserving operations," *Phys. Rev. Lett.*, **90**, 027901 (2003); quant-ph/0207146 (2002).
20. E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, R. Muñoz-Tapia, O. Romero-Isart, "Parameter estimation of qubit mixed states," *Proc. ERATO Conference on Quantum Information Science (EQIS) 2005*, 35–36 (2005).
21. E. Bagan, M. Baig, R. Muñoz-Tapia, "Entanglement-assisted alignment of reference frames using a dense covariant coding," *Phys. Rev. A*, **69**, 050303(R) (2004); quant-ph/0303019 (2003).
22. E. Bagan, M. Baig, R. Muñoz-Tapia, "Quantum reverse-engineering and reference-frame alignment without nonlocal correlations," *Phys. Rev. A*, **70**, 030301(R) (2004); quant-ph/0405082 (2004).
23. R. R. Bahadur, "On the asymptotic efficiency of tests and estimates," *Sankhyā*, **22**, 229 (1960).
24. R. R. Bahadur, *Ann. Math. Stat.*, **38**, 303 (1967).
25. R. R. Bahadur, *Some limit theorems in statistics,* Regional Conference Series in Applied Mathematics, no. 4, SIAM, Philadelphia (1971).
26. M. A. Ballester, "Estimation of unitary quantum operations," *Phys. Rev. A*, **69**, 022303 (2004); quant-ph/0305104 (2003).
27. M. Ban, K. Kurokawa, R. Momose, O. Hirota, "Optimum measurements for discrimination among symmetric quantum states and parameter estimation," *Int. J. Theor. Phys.*, **36**, 1269 (1997).
28. A. Barenco, A. K. Ekert, "Dense coding based on quantum entanglement," *J. Mod. Opt.*, **42**, 1253 (1995).
29. H. Barnum, C. A. Fuchs, R. Jozsa, B. Schumacher, "A general fidelity limit for quantum channels," *Phys. Rev. A*, **54**, 4707–4711 (1996); quant-ph/9603014 (1996).
30. H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, B. Schumacher, "On quantum coding for ensembles of mixed states," *J. Phys. A Math. Gen.*, **34**, 6767–6785 (2001); quant-ph/0008024.
31. H. Barnum, E. Knill, M. A. Nielsen, "On quantum fidelities and channel capacities," *IEEE Trans. Inf. Theory*, **46**, 1317–1329 (2000); quant-ph/9809010 (1998).
32. H. Barnum, M. A. Nielsen, B. Schumacher, "Information transmission through a noisy quantum channel," *Phys. Rev. A*, **57**, 4153–4175 (1997); quant-ph/9702049 (1997).
33. H. Bechmann-Pasquinucci, N. Gisin, "Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography," *Phys. Rev. A*, **59**, 4238–4248 (1999).

34. V. P. Belavkin, "Generalized uncertainty relations and efficient measurements in quantum systems," *Teor. Mat. Fiz.*, **26**, 3, 316–329 (1976); quant-ph/0412030 (2004).
35. C. H. Bennett, G. Brassard, "Quantum cryptography: public key distribution and coin tossing," *Proc. IEEE International Conference on Computers, Systems and Signal Processing*, (Bangalore, India), pp. 175–179 (1984).
36. C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, **70**, 1895 (1993).
37. C. H. Bennett, C. A. Fuchs, J. A. Smolin, "Entanglement-enhanced classical communication on a noisy quantum channel," in *Quantum Communication, Computing, and Measurement*, O. Hirota, A. S. Holevo, C. M. Caves (eds.), pp. 79–88 (Plenum, New York, 1997).
38. C. H. Bennett, H. J. Bernstein, S. Popescu, B. Schumacher, "Concentrating partial entanglement by local operations," *Phys. Rev. A*, **53**, 2046 (1996); quant-ph/9511030 (1995).
39. C. H. Bennett, "Quantum cryptography using any two nonorthogonal states," *Phys. Rev. Lett.*, **68**, 3121–3124 (1992).
40. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, W. K. Wootters, "Mixed state entanglement and quantum error correction," *Phys. Rev. A*, **54**, 3824–3851 (1996); quant-ph/9604024 (1996).
41. C. H. Bennett, P. W. Shor, J. A. Smolin, A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, **83**, 3081 (1999); quant-ph/9904023 (1999).
42. C. H. Bennett, P. W. Shor, J. A. Smolin, A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Trans. Inf. Theory*, **48**, (10), 2637–2655 (2002); quant-ph/0106052.
43. C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Phys. Rev. Lett.*, **69**, 2881 (1992).
44. C. H. Bennett, I. Devetak, P. W. Shor, J. A. Smolin, "Inequalities and separations among assisted capacities of quantum channels," quant-ph/0406086 (2004).
45. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, "Capacities of quantum erasure channels," quant-ph/9701015 (1997).
46. C. H. Bennett, A. W. Harrow, S. Lloyd, "Universal quantum data compression via gentle tomography," quant-ph/0403078 (2000).
47. C. H. Bennett, A. Winter, Private Communication (2005).
48. A. Ben-Tal, A. Nemirovski, *Lectures on Modern Convex Optimization*, (SIAM/MPS, Philadelphia, 2001).
49. R. Bhatia, *Matrix analysis*. (Springer, Berlin Heidelberg New York, 1997).
50. I. Bjelaković, T. Kruger, R. Siegmund-Schultze, A. Szkoła, "The Shannon-McMillan theorem for ergodic quantum lattice systems," Invent. Math. **155**, 203–222 (2004); math.DS/0207121 (2002).
51. I. Bjelaković, A. Szkoła, "The data compression theorem for ergodic quantum information sources," *Quant. Inf. Process.*, **4**, 49–63 (2005); quant-ph/0301043 (2003).
52. D. Blackwell, L. Breiman, A.J. Thomasian, "The capacity of a class of channels," *Ann. Math. Stat.*, **30**, 1229–1241 (1959).

53. R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inf. Theory*, **18**, 460–473 (1972).
54. G. Blakely, "Safeguarding cryptographic keys," *Proc. AFIPS*, **48**, 313 (1979).
55. N. A. Bogomolov, "Minimax measurements in a general statistical decision theory," *Teor. Veroyatnost. Primenen.*, **26**, 798–807 (1981); (English translation: *Theory Probab. Appl.*, **26**, 4, 787–795 (1981))
56. S. Bose, M. B. Plenio, B. Vedral, "Mixed state dense coding and its relation to entanglement measures," *J. Mod. Opt.*, **47**, 291, (2000); quant-ph/9810025 (1998).
57. D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, A. Zeilinger, "Experimental quantum teleportation," *Nature*, **390**, 575–579 (1997).
58. G. Bowen, "Classical information capacity of superdense coding," *Phys. Rev. A*, **63**, 022302 (2001); quant-ph/0101117 (2001).
59. D. C. Brody, L. P. Hughston, *R. Soc. Lond. Proc. A*, **454**, 2445–2475 (1998).
60. D. Bruß, "Optimal eavesdropping in quantum cryptography with six states," *Phys. Rev. Lett.*, 81, 3018–3021 (1998).
61. D. Bruß, A. Ekert, C. Machiavello, "Optimal universal quantum cloning and state estimation," *Phys. Rev. Lett.*, **81**, 2598–2601, (1998). (also appeared as Chap. 24 of *Asymptotic Theory of Quantum Statistical Inference,* M. Hayashi eds.)
62. J. A. Bucklew, *Large Deviation Techniques in Decision, Simulation, and Estimation*, (Wiley, New York, 1990).
63. M. V. Burnashev, A. S. Holevo, "On reliability function of quantum communication channel," *Prob. Inf. Trans.*, **34**, 97–107 (1998); quant-ph/9703013 (1997).
64. P. Busch, M. Grabowski, P. J. Lahti, *Operational quantum physics*, Lecture Notes in Physics, vol. 31, (Springer, Berlin Heidelberg New York, 1997).
65. V. Bužek, R. Derka, and S. Massar, "Optimal quantum clocks," *Phys. Rev. Lett.*, **82**, 2207 (1999); quant-ph/9808042 (1998).
66. A. R. Calderbank, E. M. Rains, P. W. Shor, N. J. A. Sloane, "Quantum error correction and orthogonal geometry," *Phys. Rev. Lett.*, **78**, 405 (1996).
67. A.R. Calderbank, P.W. Shor: "Good quantum error-correcting codes exist," *Phys. Rev. A*, **54**, 1098 (1996); quant-ph/9512032 (1995).
68. N. J. Cerf, C. Adami: "Negative entropy and information in quantum mechanics," *Phys. Rev. Lett.*, **79**, 5194 (1997); quant-ph/9512022.
69. N. J. Cerf, C. Adami, R. M. Gingrich, "Reduction criterion for separability," *Phys. Rev. A*, **60**, 898 (1999).
70. A. Chefles, "Condition for unambiguous state discrimination using local operations and classical communication," *Phys. Rev. A*, **69**, 050307(R) (2004).
71. Y.-X. Chen, D. Yang, "Distillable entanglement of multiple copies of Bell states," *Phys. Rev. A*, **66**, 014303 (2002).
72. H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Stat.*, **23**, 493–507 (1952).
73. G. Chiribella, G. M. D'Ariano, P. Perinotti, M. F. Sacchi, "Efficient use of quantum resources for the transmission of a reference frame," *Phys. Rev. Lett.*, **93**, 180503 (2004); quant-ph/0405095 (2004).
74. G. Chiribella, G. M. D'Ariano, P. Perinotti, M. F. Sacchi, Covariant quantum measurements which maximize the likelihood *Phys. Rev. A*, **70**, 061205 (2004); quant-ph/0403083 (2004).

75. G. Chiribella, G. M. D'Ariano, M. F. Sacchi, "Optimal estimation of group transformations using entanglement," *Phys. Rev. A*, **72**, 042338 (2005); quant-ph/0506267 (2005).
76. G. Chiribella, G. M. D'Ariano, P. Perinotti, M. F. Sacchi, "Maximum likelihood estimation for a group of physical transformations," quant-ph/0507007 (2005).
77. M.-D. Choi, "Completely positive linear maps on complex matrices," *Lin. Alg. Appl.*, **10**, 285–290 (1975).
78. M. Christandl, A. Winter, "Squashed entanglement"–an additive entanglement measure," *J. Math. Phys.*, **45**, 829–840 (2004); quant-ph/0308088 (2003).
79. M. Christandl, A. Winter, "Uncertainty, monogamy, and locking of quantum correlations," *IEEE Trans Inf Theory*, **51**, 3159–3165 (2005); quant-ph/0501090 (2005).
80. J. I. Cirac, W. Dür, B. Kraus, M. Lewenstein, "Entangling operations and their implementation using a small amount of entanglement," *Phys. Rev. Lett.*, **86**, 544 (2001); quant-ph/0007057 (2000).
81. R. Cleve, D. Gottesman, H.-K. Lo, "How to share a quantum secret," *Phys. Rev. Lett.*, **82**, 648 (1999); quant-ph/9901025 (1999).
82. T. Cover, J. Thomas, *Elements of Information Theory*, (Wiley, New York, 1991).
83. H. Cramér, "Sur un nouveaux theoorème-limite de la théorie des probabilités," in *Actualités Scientifiques et Industrielles*, no. 736 in *Colloque consacré à la thèorie des probabilités*, 5–23, Hermann, Paris, 1938.
84. I. Csiszár, "Information type measures of difference of probability distribution and indirect observations," *Studia Scient. Math. Hungar.*, **2**, 299–318 (1967).
85. I. Csiszár, J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, (Academic, 1981).
86. G. M. D'Ariano, "Homodyning as universal detection," in *Quantum Communication, Computing, and Measurement*, O. Hirota, A. S. Holevo, C. A. Caves (eds.), Plenum, New York, pp. 253–264 (1997); quant-ph/9701011 (1997). (also appeared as Chap. 33 of *Asymptotic Theory of Quantum Statistical Inference,* M. Hayashi eds.)
87. G. M. D'Ariano, P. L. Presti, "Imprinting Complete Information about a Quantum Channel on its Output State," *Phys. Rev. Lett.*, **91**, 047902 (2003).
88. G. M. D'Ariano, P. Lo Presti, M. F. Sacchi, "Bell measurements and observables," *Phys. Lett. A*, **272**, 32 (2000).
89. N. Datta, Y. Suhov, "Data compression limit for an information source of interacting qubits," *Quant. Inf. Process.*, **1**, (4), 257–281 (2002); quant-ph/0207069 (2002).
90. N. Datta, A. S. Holevo, Y. Suhov, "Additivity for transpose depolarizing channels," quant-ph/0412034 (2004).
91. N. Datta, M. B. Ruskai, "Maximal output purity and capacity for asymmetric unital qudit channels," quant-ph/0505048 (2005).
92. E. B. Davies, J. T. Lewis, "An operational approach to quantum probability," *Comm. Math. Phys.*, **17**, 239 (1970).
93. L. D. Davisson, "Comments on "Sequence time coding for data compression," *Proc. IEEE*, **54**, 2010 (1966).
94. A. Dembo, O. Zeitouni, *Large Deviation Techniques and Applications*, (Springer, Berlin Heidelberg New York, 1997).

95. I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, **51**, 44–55 (2005); quant-ph/0304127 (2003).
96. I. Devetak, P. W. Shor, "The capacity of a quantum channel for simultaneous transmission of classical and quantum information," quant-ph/0311131 (2003).
97. I. Devetak, A. Winter, "Classical data compression with quantum side information," *Phys. Rev. A*, **68**, 042301 (2003); quant-ph/0209029 (2002).
98. I. Devetak, A. Winter, "Distillation of secret key and entanglement from quantum states," *Proc. R. Soc. Lond. A*, **461**, 207–235 (2005); quant-ph/0306078 (2003).
99. D. DiVincenzo, P. Shor, J. Smolin, "Quantum channel capacities of very noisy channels," *Phys. Rev. A*, **57**, 830–839 (1998); quant-ph/9706061 (1997).
100. W. Dür, G. Vidal, J. I. Cirac, "Visible compression of commuting mixed state," *Phys. Rev. A*, **64**, 022308 (2001); quant-ph/0101111 (2001).
101. W. Dür, J. I. Cirac, M. Lewenstein, D. Bruß, "Distillability and partial transposition in bipartite systems," *Phys. Rev. A*, **61**, 062313 (2000).
102. M. J. Donald, M. Horodecki, "Continuity of relative entropy of entanglement," *Phys. Lett. A*, **264**, 257–260 (1999).
103. M. Donald, M. Horodecki, and O. Rudolph, "The uniqueness theorem for entanglement measures," *J. Math. Phys.*, **43**, 4252–4272 (2002).
104. A. Einstein, R. Podolsky, N. Rosen: "Can quantum-mechanical descriptions of physical reality be considered complete?," *Phys. Rev.*, **47**, 777–780 (1935).
105. I. Ekeland, R. Téman, *Convex Analysis and Variational Problems*, (North-Holland, Amsterdam, 1976); (SIAM, Philadelphia, 1999).
106. R. Ellis, "Large deviations for a general class of random vectors," *Ann. Probab.*, **12**, 1, 1–12 (1984). *Entropy, Large Deviations and Statistical Mechanics*, (Springer, Berlin Heidelberg New York, 1985).
107. H. Fan, "Distinguishability and indistinguishability by local operations and classical communication," *Phys. Rev. Lett.*, **92**, 177905 (2004).
108. H. Fan, "A note on quantum entropy inequalities and channel capacities," *J. Phys. A Math. Gen.*, **36**, 12081–12088 (2003).
109. H. Fan, K. Matsumoto, M. Wadati, "Quantum cloning machines of a d-level system," *Phys. Rev. A*, **64**, 064301 (2001); quant-ph/0103053 (2001).
110. H. Fan, K. Matsumoto, X. Wang, M. Wadati, "Quantum cloning machines for equatorial qubits," *Phys. Rev. A*, **65**, 012304 (2002); quant-ph/0101101 (2001).
111. M. Fannes, "A continuity property of the entropy density for spin lattice systems," *Comm. Math. Phys.*, **31**, 291–294 (1973).
112. M. Fannes, B. Haegeman, M. Mosonyi, D. Vanpeteghem, "Additivity of minimal entropy output for a class of covariant channels," quant-ph/0410195 (2004).
113. R. M. Fano, *Transmission of Information: A Statistical Theory of Communication*, (Wiley, New York, 1961).
114. A. Feinstein: "A new basic theorem of information theory," *IRE Trans. PGIT*, **4**, 2–22 (1954).
115. D. G. Fischer, H. Mack, M. A. Cirone, M. Freyberger, "Enhanced estimation of a noisy quantum channel using entanglement," *Phys. Rev. A*, **64**, 022309 (2001); quant-ph/0103160 (2001).

116. G. D. Forney, Jr., S. M. Thesis, MIT 1963 (unpublished).
117. J. C. Fu, "On a theorem of Bahadur on the rate of convergence of point estimators," *Ann. Stat.*, **1**, 745 (1973).
118. C. A. Fuchs, "Distinguishability and accessible information in quantum theory," quant-ph/9601020 (1996).
119. A. Fujiwara, *Statistical Estimation Theory for Quantum States*, master's thesis, Department of Mathematical Engineering and Information Physics, Graduate School of Engineering, University of Tokyo, Japan (1993) (in Japanese).
120. A. Fujiwara, *A Geometrical Study in Quantum Information Systems*, Ph.D. thesis, Department of Mathematical Engineering and Information Physics, Graduate School of Engineering, University of Tokyo, Japan (1995).
121. A. Fujiwara, private communication to H. Nagaoka (1996).
122. A. Fujiwara, "Geometry of quantum information systems," in *Geometry in Present Day Science*, O. E. Barndorff-Nielsen, E. B. V. Jensen (eds.), (World Scientific, Singapore, 1998), pp. 35–48.
123. A. Fujiwara, "Quantum birthday problems: geometrical aspects of quantum randomcoding," *IEEE Trans. Inf. Theory*, **47**, 2644–2649 (2001).
124. A. Fujiwara, "Quantum channel identification problem," *Phys. Rev. A*, **63**, 042304 (2001).
125. A. Fujiwara, "Estimation of SU(2) operation and dense coding: an information geometric approach," *Phys. Rev. A*, **65**, 012316 (2002).
126. A. Fujiwara, H. Nagaoka, "Quantum Fisher metric and estimation for pure state models," *Phys. Lett.*, **201A**, 119–124 (1995).
127. A. Fujiwara, H. Nagaoka, "Coherency in view of quantum estimation theory," in *Quantum Coherence and Decoherence*, K. Fujikawa, Y. A. Ono (eds.), (Elsevier, Amsterdam, 1996), pp. 303–306.
128. A. Fujiwara, H. Nagaoka, "Operational capacity and pseudoclassicality of a quantum channel," *IEEE Trans. Inf. Theory*, **44**, 1071–1086 (1998).
129. A. Fujiwara, H. Nagaoka, "An estimation theoretical characterization of coherent states," *J. Math. Phys.*, **40**, 4227–4239 (1999).
130. A. Fujiwara, P. Algoet, "One-to-one parametrization of quantum channels," *Phys. Rev. A*, **59**, 3290–3294 (1999).
131. A. Fujiwara, "Quantum channel identification problem," *Phys. Rev. A*, **63**, 042304 (2001).
132. A. Fujiwara, "Estimation of SU(2) operation and dense coding: an information geometric approach," *Phys. Rev. A*, **65**, 012316 (2002).
133. A. Fujiwara, T. Hashizume, "Additivity of the capacity of depolarizing channels," *Phys. Lett A*, **299**, 469–475 (2002).
134. A. Fujiwara and H. Imai, "Quantum parameter estimation of a generalized Pauli channel," *J. Phys. A Math. Gen.*, **36**, 8093–8103 (2003).
135. A. Fujiwara, "Estimation of a generalized amplitude-damping channel," *Phys. Rev. A*, **70**, 012317 (2004).
136. A. Fujiwara, "Mathematics of quantum channels," *Suurikagaku*, 474, 28–35 (2002) (in Japanese).
137. M. Fujiwara, M. Takeoka, J. Mizuno, M. Sasaki, "Exceeding classical capacity limit in quantum optical channel," *Phys. Rev. Lett.*, **90**, 167906 (2003); quant-ph/0304037 (2003).
138. M. Fukuda, "Extending additivity from symmetric to asymmetric channels," *J. Phys. A Math. Gen.*, **38**, L753–L758 (2005); quant-ph/0505022 (2005).

139. A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, E. J. Polzik, "Unconditional quantum teleportation," *Science*, **282**, 706 (1998).
140. R. G. Gallager, *Information Theory and Reliable Communication*, (Wiley, New York, 1968).
141. J. Gärtner, "On large deviations from the invariant measure," *Theory Probabil. Appl.*, **22**, 24–39 (1977).
142. R. Gill, S. Massar, "State estimation for large ensembles," *Phys. Rev. A*, **61**, 042312 (2000); quant-ph/9902063 (1999).
143. R. D. Gill, "Asymptotic information bounds in quantum statistics," math.ST/0512443 (2005).
144. R. D. Gill, B. Y. Levit, "Applications of the van Tree inequality: a Bayesian Cramér-Rao bound" *Bernoulli*, **1**, 59–79 (1995).
145. N. Giri, W. von Waldenfels, "An algebraic version of the central limit theorem," *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **42**, 129–134 (1978).
146. N. Gisin, contribution to the Torino Workshop, 1997.
147. S. Ghosh, G. Kar, A. Roy, D. Sarkar, "Distinguishability of maximally entangled states," *Phys. Rev. A*, **70**, 022304 (2004).
148. J. P. Gordon, *Proc. IRE*, **50**, 1898–1908 (1962).
149. J. P. Gordon, "Noise at optical frequencies; information theory," in *Quantum Electronics and Coherent Light*, *Proc. Int. School Phys. "Enrico Fermi," Course XXXI*, P. A. Miles (ed.). (Academic, New York, 1964), pp. 156–181.
150. D. Gottesman, "Class of quantum error-correcting codes saturating the quantum Hamming bound," *Phys. Rev. A*, **54**, 1862 (1996).
151. D. Gottesman, "On the theory of quantum secret sharing," *Phys. Rev. A*, **61**, 042311 (2000); quant-ph/9910067 (1999).
152. J. Gruska, H. Imai, "Power, puzzles and properties of entanglement," in *achines, Computations, and Universality*, *Proc. 3rd International Conference, MCU 2001*, Lecture Notes in Computer Science, vol. 2055, (Springer, Berlin Heidelberg New York, 2001), pp. 25–68.
153. J. Gruska, H. Imai, K. Matsumoto, "Power of quantum entanglement," in *Foundations of Information Technology in the Era of Networking and Mobile Computing*, *IFIP 17th World Computer Congress - TC1 Stream / 2nd IFIP International Conference on Theoretical Computer Science*, vol. 223 of IFIP Conference Proceedings, pp. 3–22, 2001.
154. M. Hamada, "Lower bounds on the quantum capacity and highest error exponent of general memoryless channels," *IEEE Trans. Inf. Theory*, **48**, 2547–2557 (2002); quant-ph/0112103 (2002).
155. M. Hamada, "Notes on the fidelity of symplectic quantum error-correcting codes," *Int. J. Quant. Inf.*, **1**, 443–463 (2003).
156. T. S. Han, "Hypothesis testing with the general source," *IEEE Trans. Inf. Theory*, **46**, 2415–2427 (2000).
157. T. S. Han, "The reliability functions of the general source with fixed-length coding," *IEEE Trans. Inf. Theory*, **46**, 2117–2132 (2000).
158. T. S. Han: *Information-Spectrum Methods in Information Theory*, (Springer, Berlin Heidelberg New York, 2002) (originally appeared in Japanese in 1998).
159. T. S. Han, K. Kobayashi, "The strong converse theorem for hypothesis testing," *IEEE Trans. Inf. Theory*, **35**, 178–180 (1989).
160. T. S. Han, K. Kobayashi, *Mathematics of Information and Encoding*, (American Mathematical Society, 2002) (originally appeared in Japanese in 1999).

161. T. S. Han, S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, **39**, 752–772 (1993).
162. T. S. Han, "Folklore in source coding: information-spectrum approach," *IEEE Trans. Inf. Theory*, **51**, (2), 747–753 (2005).
163. A. Harrow, H. K. Lo, "A tight lower bound on the classical communication cost of entanglement dilution," *IEEE Trans. Inf. Theory*, **50**, 319–327 (2004); quant-ph/0204096 (2002).
164. P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, W. Wooters, "Classical information capacity of a quantum channel," *Phys. Rev. A*, **54**, 1869–1876 (1996).
165. M. Hayashi, *Minimization of deviation under quantum local unbiased measurements,* master's thesis, Department of Mathematics, Graduate School of Science, Kyoto University, Japan (1996).
166. M. Hayashi, "A linear programming approach to attainable cramer-rao type bound and randomness conditions," Kyoto-Math 97–08; quant-ph/9704044 (1997).
167. M. Hayashi, "A linear programming approach to attainable Cramer–Rao type bound," in *Quantum Communication, Computing, and Measurement*, O. Hirota, A. S. Holevo, C. M. Caves (eds.), (Plenum, New York, 1997), pp. 99–108. (Also appeared as Chap. 12 of *Asymptotic Theory of Quantum Statistical Inference,* M. Hayashi eds.)
168. M. Hayashi, "Asymptotic estimation theory for a finite dimensional pure state model," *J. Phys. A Math. Gen.*, **31**, 4633–4655 (1998); quant-ph/9704041 (1997). (Also appeared as Chap. 23 of *Asymptotic Theory of Quantum Statistical Inference*, M. Hayashi eds.).
169. M. Hayashi, "Asymptotics of quantum relative entropy from a representation theoretical viewpoint," *J. Phys. A Math. Gen.*, **34**, 3413–3419 (2001); quant-ph/9704040 (1997).
170. M. Hayashi, "Asymptotic quantum estimation theory for the thermal states family," in *Quantum Communication, Computing, and Measurement 2*, P. Kumar, G. M. D'ariano, O. Hirota (eds.), (Plenum, New York, 2000) pp. 99–104; quant-ph/9809002 (1998). (Also appeared as Chap. 14 of *Asymptotic Theory of Quantum Statistical Inference,* M. Hayashi eds.).
171. M. Hayashi, "Asymptotic large deviation evaluation in quantum estimation theory," *Proc. Symposium on Statistical Inference and Its Information-Theoretical Aspect*, pp. 53–82 (1998) (in Japanese).
172. M. Hayashi, "Simultaneous measurements of non-commuting physical quantities," *RIMS koukyuroku Kyoto Univiversity*, No. 1099, 96–118 (1999) (in Japanese).
173. M. Hayashi, "Quantum hypothesis testing for the general quantum hypotheses," *Proc. 24th Symposium on Information Theory and Its Applications (SITA)*, pp. 591–594 (2001).
174. M. Hayashi, "Optimal sequence of POVMs in the sense of Stein's lemma in quantum hypothesis," *J. Phys. A Math. Gen.*, **35**, 10759–10773 (2002); quant-ph/0107004 (2001).
175. M. Hayashi, "Exponents of quantum fixed-length pure state source coding," *Phys. Rev. A*, **66**, 032321 (2002); quant-ph/0202002 (2002).
176. M. Hayashi, "General formulas for fixed-length quantum entanglement concentration," appear in *IEEE Trans. Infor. Theory*; quant-ph/0206187 (2002).

177. M. Hayashi, "Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation," *J. Phys. A Math. Gen.*, **35**, 7689–7727 (2002); quant-ph/0202003 (2002). (Also appeared as Chap. 28 of *Asymptotic Theory of Quantum Statistical Inference,* M. Hayashi eds.).
178. M. Hayashi, "Quantum estimation and quantum central limit theorem," *Sugaku*, **55**, 4, 368–391 (2003) (in Japanese).
179. M. Hayashi, "Second order asymptotics in fixed-length source coding and intrinsic randomness," cs.IT/0503089 (2005).
180. M. Hayashi, "Characterization of several kinds of quantum analogues of relative entropy," quant-ph/0510181 (2005); *Proc. 9th Quantum Information Technology Symposium (QIT13)*, Touhoku University, Sendai, Miyagi, Japan, 24–25 November 2005, pp. 137–142.
181. M. Hayashi, M. Koashi, K. Matsumoto, F. Morikoshi A. Winter, "Error exponents for entangle concentration," *J. Phys. A Math. Gen.*, **36**, 527–553 (2003); quant-ph/0206097 (2002).
182. M. Hayashi, K. Matsumoto, "Statistical model with measurement degree of freedom and quantum physics," *RIMS koukyuroku Kyoto Univiversity*, **1055**, 96–110 (1998) (in Japanese). (Also appeared as Chap. 13 of *Asymptotic Theory of Quantum Statistical Inference,* M. Hayashi eds.)
183. M. Hayashi, K. Matsumoto, "Variable length universal entanglement concentration by local operations and its application to teleportation and dense coding," quant-ph/0109028 (2001); K. Matsumoto, M. Hayashi, "Universal entanglement concentration," quant-ph/0509140 (2005).
184. M. Hayashi, K. Matsumoto, "Quantum universal variable-length source coding," *Phys. Rev. A*, **66**, 022311 (2002); quant-ph/0202001 (2002).
185. M. Hayashi, K. Matsumoto, " Simple construction of quantum universal variable-length source coding," *Quant. Inf. Comput.*, **2**, Special Issue, 519–529 (2002); quant-ph/0209124, (2002).
186. M. Hayashi, H. Nagaoka: "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inf. Theory*, **49**, 1753–1768 (2003); quant-ph/0206186 (2002).
187. M. Hayashi, F. Sakaguchi, "Subnormal operators regarded as generalized observables and compound-system-type normal extension related to su(1,1)," *J. Phys. A Math. Gen.*, **33**, 7793–7820 (2000); quant-ph/0003079 (2000).
188. M. Hayashi, K. Matsumoto, "Asymptotic performance of optimal state estimation in quantum two level system," quant-ph/0411073 (2004).
189. M. Hayashi, "Parallel Treatment of Estimation of SU(2) and Phase Estimation," appear in *Phys. Lett. A*. quant-ph/0407053 (2004).
190. M. Hayashi, "Estimation of SU(2) action by using entanglement," *Proc. 9th Quantum Information Technology Symposium (QIT9)*, NTT Basic Research Laboratories, Atsugi, Kangawa, Japan, 11–12 December 2003, pp. 9–13 (in Japanese); *ibid*, *Proc. 7th International Conference on Quantum Communication, Measurement and Computing*, Glasgow, UK, 25–29 July 2004 (AIP), pp. 269–272.
191. M. Hayashi, D. Markham, M. Murao, M. Owari, S.Virmani, "Bounds on Multipartite Entangled Orthogonal State Discrimination Using Local Operations and Classical Communication," *Phys. Rev. Lett.*, **96**, 040501 (2006).
192. M. Hayashi, "On the second order asymptotics for pure states family," *IEICE Trans.*, **E88-JA**, 903–916 (2005) (in Japanese).

193. M. Hayashi, H. Imai, K. Matsumoto, M. B. Ruskai, T. Shimono, "Qubit channels which require four inputs to achieve capacity: implications for additivity conjectures," *Quant. Inf. Comput.*, **5**, 13–31 (2005); quant-ph/040317.
194. M. Hayashi (eds.), *Asymptotic Theory of Quantum Statistical Inference: Selected Papers,* (World Scientific, Singapore, 2005).
195. M. Hayashi, "Hypothesis Testing of maximally entangled state" in preparation.
196. P. M. Hayden, M. Horodecki, B. M. Terhal, "The asymptotic entanglement cost of preparing a quantum state," *J. Phys. A Math. Gen.*, **34**, 6891–6898 (2001); quant-ph/0008134 (2000).
197. P. Hayden, R. Jozsa, A. Winter, "Trading quantum for classical resources in quantum data compression," *J. Math. Phys.*, **43**, 4404–4444 (2002); quant-ph/0204038 (2002).
198. P. Hayden, A. Winter, "On the communication cost of entanglement transformations," *Phys. Rev. A*, **67**, 012326 (2003); quant-ph/0204092 (2002).
199. P. Hayden, R. Jozsa, D. Petz, A. Winter, "Structure of states which satisfy strong subadditivity of quantum entropy with equality," *Comm. Math. Phys.*, **246**, 359–374 (2004).
200. W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," *Z. Phys.*, **43**, 172–198 (1927).
201. C. W. Helstrom, "Detection theory and quantum mechanics," *Inf. Contr.*, **10**, 254–291 (1967).
202. C. W. Helstrom, "Minimum mean-square error estimation in quantum statistics," *Phys. Lett.*, **25A**, 101–102 (1976).
203. C. W. Helstrom, *Quantum Detection and Estimation Theory,* (Academic, New York, 1976).
204. C. W. Helstrom, *Int. J. Theor. Phys.*, **11**, 357 (1974).
205. L. Henderson, V. Vedral, "Classical, quantum and total correlations," *J. Phys. A Math. Gen.*, **34**, 6899 (2001); quant-ph/0105028 (2001).
206. F. Hiai, D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Comm. Math. Phys.*, **143**, 99–114 (1991).
207. F. Hiai, D. Petz, "The golden-thompson trace inequality is complemented," *Lin. Alg. Appl.*, **181**, 153–185 (1993).
208. T. Hiroshima, "Majorization criterion for distillability of a bipartite quantum state," *Phys. Rev. Lett.*, **91**, 057902 (2003); quant-ph/0303057 (2003).
209. T. Hiroshima, M. Hayashi, "Finding a maximally correlated state-Simultaneous Schmidt decomposition of bipartite pure states," *Phys. Rev. A*, **70**, 030302(R) (2004).
210. W. Hoeffding, "Asymptotically optimal test for multinomial distributions," *Ann. Math. Stat.*, **36**, 369–400 (1965).
211. A. S. Holevo, "An analog of the theory of statistical decisions in noncommutative theory of probability," *Trudy Moskov. Mat. Obšč.*, **26**, 133–149 (1972) (in Russian). (English translation: *Trans. Moscow Math. Soc.*, **26**, 133–149 (1972)).
212. A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problemy Peredachi Informatsii*, **9**, 3–11 (1973) (in Russian). (English translation: *Probl. Inf. Transm.*, **9**, 177–183 (1975)).
213. A. S. Holevo, "Some statistical problems for quantum Gaussian states," *IEEE Trans. Inf. Theory*, **21**, 533–543 (1975).

214. A. S. Holevo, "On the capacity of quantum communication channel," *Problemly Peredachi Informatsii*, **15**, 4, 3–11 (1979) (in Russian). (English translation: *Probl. Inf. Transm.*, **15**, 247–253 (1979).)
215. A. S. Holevo, "Covariant measurements and uncertainty relations," *Rep. Math. Phys.*, **16**, 385–400 (1979).
216. A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, (North-Holland, Amsterdam, 1982); originally published in Russian (1980).
217. A. S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inf. Theory*, **44**, 269 (1998); quant-ph/9611023 (1996).
218. A. S. Holevo, "On quantum communication channels with constrained inputs," quant-ph/9705054 (1997).
219. A. S. Holevo, "Coding theorems for quantum channels," *Tamagawa University Res. Rev.*, 4, (1998); quant-ph/9809023 (1998).
220. A. S. Holevo: "On entanglement-assisted classical capacity," *J. Math. Phys.*, **43**, 4326–4333 (2002); quant-ph/0106075 (2001).
221. A. S. Holevo, *Statistical structure of quantum theory*, Lecture Notes in Physics, vol. 67, (Springer, Berlin Heidelberg New York, 2001).
222. A. S. Holevo, M. E. Shirokov, "On Shor's channel extension and constrained channels," *Commun. Math. Phys.*, **249**, 417–430, (2004); quant-ph/0306196 (2003).
223. A. S. Holevo, "Covariant Measurements and Optimality," Chap. IV of *Probabilistic and Statistical Aspects of Quantum Theory*, (North-Holland, Amsterdam, 1982). Originally published in Russian (1980): "Bounds for generalized uncertainty of the shift parameter," Lecture Notes in Mathematics, **1021**, 243–251 (1983).
224. R. Horodecki, P. Horodecki, "Quantum redundancies and local realism," *Phys. Lett.*, **A 194**, 147–152 (1994).
225. M. Horodecki, P. Horodecki, R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," *Phys. Lett.*, **A 223**, 1–8 (1996).
226. M. Horodecki, P. Horodecki, R. Horodecki, "Inseparable two-spin 1/2 density matrices can be distilled to a singlet form," *Phys. Rev. Lett.*, **78**, 574 (1997).
227. M. Horodecki, P. Horodecki, R. Horodecki, "Mixed-state entanglement and distillation: is there a "bound" entanglement in nature?" *Phys. Rev. Lett.*, **80**, 5239 (1998); quant-ph/9801069 (1998).
228. M. Horodecki, "Limits for compression of quantum information carried by ensembles of mixed states," *Phys. Rev. A*, **57**, 3364–3369 (1998); quant-ph/9712035 (1997).
229. M. Horodecki, P. Horodecki, "Reduction criterion of separability and limits for a class of distillation protocols," *Phys. Rev. A*, **59**, 4206 (1999); quant-ph/9708015 (1997).
230. M. Horodecki, P. Horodecki, R. Horodecki, "General teleportation channel, singlet fraction and quasi-distillation," *Phys. Rev. A*, **60**, 1888 (1999); quant-ph/9807091 (1998).
231. M. Horodecki, "Optimal compression for mixed signal states," *Phys. Rev. A*, **61**, 052309 (2000); quant-ph/990508 v3 (2000).
232. M. Horodecki, P. Horodecki, R. Horodecki, "Unified approach to quantum capacities: towards quantum noisy coding theorem," *Phys. Rev. Lett.*, **85**, 433–436 (2000).

233. M. Horodecki, P. Horodecki, R. Horodecki, "Mixed-state entanglement and quantum communication," in *Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments*, (Springer Tracts in Modern Physics, 173), G. Alber, T. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rotteler, H. Weinfurter, R. Werner, A. Zeilinger (eds.), (Springer, Berlin Heidelberg New York, 2001).
234. M. Horodecki, P. Horodecki, R. Horodecki, D. W. Leung, B. M. Terhal, "Classical capacity of a noiseless quantum channel assisted by noisy entanglement," *Quant. Inf. Comput.*, **1**, 70–78 (2001); quant-ph/0106080 (2001).
235. M. Horodecki, P. Shor, M. B. Ruskai, "Entanglement breaking channels," *Rev. Math. Phys.*, **15**, 1–13 (2003); quant-ph/0302031 (2003).
236. P. Horodecki, "Separability criterion and inseparable mixed states with positive partial transposition," *Phys. Lett.*, **A232**, 333 (1997); quant-ph/9703004 (1997).
237. T. Hiroshima, "Optimal dense coding with mixed state entanglement," *J. Phys. A Math. Gen.*, **34**, 6907–6912 (2001); quant-ph/0009048 (2000).
238. H. Imai, M. Hachimori, M. Hamada, H. Kobayashi, K. Matsumoto, "Optimization in quantum computation and information," *Proc. 2nd Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications*, Budapest, Hungary (2001).
239. S. Ishikawa, "Uncertainty relations in simultaneous measurements for arbitrary observables," *Rep. Math. Phys.*, **29**, 257–273 (1991).
240. S. Ishizaka, T. Hiroshima, "Maximally entangled mixed states under nonlocal unitary operations in two qubits," *Phys. Rev. A*, **62**, 022310 (2000).
241. S. Ishizaka, "Binegativity and geometry of entangled states in two states," *Phys. Rev. A*, **69**, 020301(R) (2004); quant-ph/0308056 (2003).
242. S. Ishizaka, "Bound entanglement provides convertibility of pure entangled states," *Phys. Rev. Lett.*, **93**, 190501 (2004); quant-ph/0403016 (2004).
243. A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Rep. Math. Phys.*, **3**, 275–278 (1972).
244. R. Jozsa, "Fidelity for mixed quantum states," *J. Mod. Opt.*, **41(12)**, 2315–2323 (1994).
245. R. Jozsa, B. Schumacher: "A new proof of the quantum noiseless coding theorem," *J. Mod. Opt.*, **41(12)**, 2343–2349 (1994).
246. R. Jozsa, "Quantum noiseless coding of mixed states," Talk given at *3rd Santa Fe Workshop on Complexity, Entropy, and the Physics of Information*, May 1994.
247. R. Jozsa, M. Horodecki, P. Horodecki, R. Horodecki, "Universal quantum information compression," *Phys. Rev. Lett.*, **81**, 1714 (1998); quant-ph/9805017 (1998).
248. R. Jozsa, S. Presnell, "Universal quantum information compression and degrees of prior knowledge," quant-ph/0210196 (2002).
249. A. Kaltchenko, E.-H. Yang, "Universal compression of ergodic quantum sources," *Quant. Inf. Comput.*, **3**, 359–375 (2003); quant-ph/0302174 (2003).
250. A. Kent, N. Linden, S. Massar, "Optimal entanglement enhancement for mixed states," *Phys. Rev. Lett.*, **83**, 2656 (1999).
251. M. Keyl, R. F. Werner, "Optimal cloning of pure states, judging single clones," *J. Math. Phys.*, **40**, 3283–3299 (1999); quant-ph/9807010 (1998).
252. M. Keyl, R. F. Werner, "Estimating the spectrum of a density operator," *Phys. Rev. A*, **64**, 052311 (2001); quant-ph/0102027 (2001).

253. C. King, "Additivity for a class of unital qubit channels," *J. Math. Phys.*, **43**, 4641–4653 (2002); quant-ph/0103156 (2001).
254. C. King, "The capacity of the quantum depolarizing channel," *IEEE Trans. Infor. Theory*, **49**, 221–229, (2003); quant-ph/0204172 (2002).
255. E. Klarreich, "Quantum cryptography: Can you keep a secret?" *Nature*, **418**, 270–272 (2002).
256. E. Knill, R. Laflamme, "Theory of quantum error-correcting codes," *Phys. Rev. A*, **55**, 900 (1997).
257. M. Koashi, N. Imoto, "Compressibility of mixed-state signals," *Phys. Rev. Lett.*, **87**, 017902 (2001); quant-ph/0103128, (2001).
258. M. Koashi, N. Imoto, "Quantum information is incompressible without errors," *Phys. Rev. Lett.*, **89**, 097904 (2002); quant-ph/0203045, (2002).
259. M. Koashi, A. Winter, "Monogamy of quantum entanglement and other correlations," *Phys. Rev. A*, **69**, 022309 (2004).
260. H. Kosaka, A. Tomita, Y. Nambu, N. Kimura, K. Nakamura, "Single-photon interference experiment over 100 km for quantum cryptography system using a balanced gated-mode photon detector," *Electron. Lett.*, **39**, 16, 1199–1201 (2003).
261. K. Kraus, *States, Effects, and Operations*, Lecture Notes in Physics, vol. 190, (Springer, Berlin Heidelberg New York, 1983).
262. P. G. Kwiat, A. G. White, et al., *Phys. Rev.* A **60**, R773 (1999), *Phys. Rev. Lett.*, **83**, 3103 (1999), *Sience*, **290**, 498 (2000), *Nature*, **409**, 1014 (2001), quant-ph/0108088 (2001).
263. E. L. Lehman, G. Casella, *Theory of Point Estimation*, (Springer, Berlin Heidelberg New York, 1998).
264. L. B. Levitin, "On quantum measure of information," in *Proc. 4th All-Union Conference on Information Transmission and Coding Theory*, pp. 111–115 (Tashkent, 1969) (in Russian). English translation: *Information, Complexity and Control in Quantum Physics*, A. Blaquiere, S. Diner, G. Lochak (eds.), (Springer, Berlin Heidelberg New York, 1987), pp. 15–47.
265. L. B. Levitin, *Information, Complexity and Control in Quantum Physics*, A. Blaquière, S. Diner, G. Lochak (eds.), (Springer, Vienna), pp. 15–47.
266. E. Lieb, *Bull. Am. Math. Soc.*, **81**, 1–13 (1975).
267. E. Lieb, "Convex trace functions and the Wigner-Yanase-Dyson conjecture," *Adv. Math.*, **11**, 267–288 (1973).
268. E. Lieb, M. B. Ruskai, "A fundamental property of quantum mechanical entropy," *Phys. Rev. Lett.*, **30**, 434–436 (1973).
269. E. Lieb, M. B. Ruskai, "Proof of the strong subadditivity of quantum mechanical entropy," *J. Math. Phys.*, **14**, 1938–1941 (1973).
270. G. Lindblad, "Completely positive maps and entropy inequalities," *Comm. Math. Phys.*, **40**, 147–151 (1975).
271. G. Lindblad, "Expectations and entropy inequalities for finite quantum systems," *Comm. Math. Phys.*, **39**, 111–119 (1974).
272. N. Linden, S. Massar, S. Popescu, "Purifying noisy entanglement requires collective measurements," *Phys. Rev. Lett.*, **81**, 3279 (1998).
273. S. Lloyd, "The capacity of the noisy quantum channel," *Phys. Rev. A*, **56**, 1613 (1997); quant-ph/9604015 (1996).
274. H.-K. Lo, "Proof of unconditional security of six-state quantum key distribution scheme," *Quant. Inf. Comput.*, **1**, 81–94 (2001).

275. H.-K. Lo, S. Popescu, "Classical communication cost of entanglement manipulation: is entanglement an interconvertible resource?" *Phys. Rev. Lett.*, **83**, 1459 (1999).
276. H.-K. Lo, S. Popescu, "Concentrating entanglement by local actions: Beyond mean values," *Phys. Rev. A*, **63**, 022301 (2001); quant-ph/9707038 (1997).
277. T. J. Lynch, "Sequence time coding for data compression," *Proc. IEEE*, **54**, 1490–1491 (1966).
278. S. Massar, S. Popescu, "Optimal extraction of information from finite quantum ensembles," *Phys. Rev. Lett.*, **74**, 1259 (1995).
279. F. De Martini, A. Mazzei, M. Ricci, and G. M. D'Ariano, "Pauli Tomography: complete characterization of a single qubit device," *Fortschr. Phys.*, **51**, 342 (2003); quant-ph/0207143 (2002).
280. L. Masanes, "All entangled states are useful for information processing," quant-ph/0508071.
281. K. Matsumoto, *Geometry of a Quantum State*, master's thesis, Department of Mathematical Engineering and Information Physics, Graduate School of Engineering, University of Tokyo, Japan (1995) (in Japanese).
282. K. Matsumoto: "A new approach to the Cramér–Rao type bound of the pure state model," *J. Phys. A Math. Gen.*, **35**, 3111–3123 (2002); quant-ph/9704044 (1997).
283. K. Matsumoto, "Uhlmann's parallelism in quantum estimation theory," quant-ph/9711027 (1997).
284. K. Matsumoto, *A Geometrical Approach to Quantum Estimation Theory*, Ph.D. thesis, Graduate School of Mathematical Sciences, University of Tokyo (1997).
285. K. Matsumoto, "The asymptotic efficiency of the consistent estimator, Berry-Uhlmann' curvature and quantum information geometry," in *Quantum Communication, Computing, and Measurement 2*, P. Kumar, G. M. D'ariano, O. Hirota (eds.), pp. 105–110 (Plenum, New York, 2000).
286. K. Matsumoto, Seminar notes (1999).
287. K. Matsumoto, private communication (2005).
288. K. Matsumoto, "Yet another additivity conjecture," *Physics Letters A*, **350**, 179–181, (2006); quant-ph/0506052 (2005).
289. K. Matsumoto, "Reverse estimation theory, complementarity between RLD and SLD, and monotone distances," quant-ph/0511170 (2005); *Proc. 9th Quantum Information Technology Symposium (QIT13)*, Touhoku University, Sendai, Miyagi, Japan, 24–25 November 2005, pp. 81–86.
290. K. Matsumoto, T. Shimono, A. Winter, "Remarks on additivity of the Holevo channel capacity and of the entanglement of formation," *Comm. Math. Phys.*, **246(3)**, 427–442 (2004); quant-ph/0206148 (2002).
291. K. Matsumoto, F. Yura, "Entanglement cost of antisymmetric states and additivity of capacity of some quantum channel," *J. Phys. A: Math. Gen.*, **37** L167-L171, (2004); quant-ph/0306009 (2003).
292. Y. Mitsumori, J. A. Vaccaro, S. M. Barnett, E. Andersson, A. Hasegawa, M. Takeoka, M. Sasaki, "Experimental demonstration of quantum source coding," *Phys. Rev. Lett.*, **91**, 217902 (2003); quant-ph/0304036 (2003).
293. A. Miyake, "Classification of multipartite entangled states by multidimensional determinants," *Phys. Rev. A*, **67**, 012108 (2003); quant-ph/0206111 (2002).

294. F. Morikoshi, "Recovery of entanglement lost in entanglement manipulation," *Phys. Rev. Lett.*, **84**, 3189 (2000); quant-ph/9911019 (1999).
295. F. Morikoshi, M. Koashi, "Deterministic entanglement concentration," *Phys. Rev. A*, **64**, 022316 (2001); quant-ph/0107120 (2001).
296. M. A. Morozowa, N. N. Chentsov, *Itogi Nauki i Tekhniki*, **36**, 289–304 (1990) (in Russian).
297. M. Murao, D. Jonathan, M. B. Plenio, V. Vedral, "Quantum telecloning and multiparticle entanglement," *Phys. Rev. A*, **59**, 156–161 (1999); quant-ph/9806082 (1998).
298. H. Nagaoka, "On Fisher information of quantum statistical models," in *Proc. 10th Symposium on Information Theory and Its Applications (SITA)*, Enoshima, Kanagawa, Japan, 19–21 November 1987, pp. 241–246. (Originally in Japanese; also appeared as Chap. 9 of *Asymptotic Theory of Quantum Statistical Inference,* M. Hayashi eds.).
299. H. Nagaoka, "An asymptotically efficient estimator for a one-dimensional parametric model of quantum statistical operators," *Proc. 1988 IEEE International Symposium on Information Theory*, 198 (1988).
300. H. Nagaoka, "On the parameter estimation problem for quantum statistical models," *Proc. 12th Symposium on Information Theory and Its Applications (SITA)*, (1989), pp. 577–582. (Also appeared as Chap. 10 of *Asymptotic Theory of Quantum Statistical Inference*, M. Hayashi eds.).
301. H. Nagaoka, "A generalization of the simultaneous diagonalization of Hermitian matrices and its relation to quantum estimation theory," *Trans. Jpn. Soc. Ind. Appl. Math.*, **1**, 43–56 (1991) (Originally in Japanese; also appeared as Chap. 11 of *Asymptotic Theory of Quantum Statistical Inference*, M. Hayashi eds.).
302. H. Nagaoka: Private communication to A. Fujiwara (1991).
303. H. Nagaoka, "On the relation between Kullback divergence and Fisher information—from classical systems to quantum systems," *Proc. Joint Mini-Workshop for Data Compression Theory and Fundamental Open Problems in Information Theory*, (1992), pp. 63–72. (Originally in Japanese; also appeared as Chap. 27 of *Asymptotic Theory of Quantum Statistical Inference*, M. Hayashi eds.).
304. H. Nagaoka, "Differential geometrical aspects of quantum state estimation and relative entropy," in *Quantum Communications and Measurement*, V. P. Belavkin, O. Hirota, R. L. Hudson (eds.), (Plenum, New York, 1995), pp. 449–452.
305. H. Nagaoka, "Algorithms of Arimoto-Blahut type for computing quantum channel capacity," *Proc. 1998 IEEE International Symposium on Information Theory*, 354 (1998).
306. H. Nagaoka, "Information spectrum theory in quantum hypothesis testing," *Proc. 22th Symposium on Information Theory and Its Applications (SITA)*, (1999), pp. 245–247 (in Japanese).
307. H. Nagaoka, "Strong converse theorems in quantum information theory," *Proc. ERATO Conference on Quantum Information Science (EQIS) 2001*, 33 (2001). (also appeared as Chap. 3 of *Asymptotic Theory of Quantum Statistical Inference*, M. Hayashi eds.).
308. H. Nagaoka, "Limit theorems in quantum information theory," *Suurikagaku*, 456, 47–55 (2001) (in Japanese).

309. H. Nagaoka, "The world of quantum information geometry," *IEICE Trans.*, **J88-A**, 8, 874–885 (2005) (in Japanese).
310. H. Nagaoka, M. Hayashi, "An information-spectrum approach to classical and quantum hypothesis testing," quant-ph/0206185 (2002).
311. H. Nagaoka, S. Osawa, "Theoretical basis and applications of the quantum Arimoto-Blahut algorithms," *Proc. 2nd Quantum Information Technology Symposium (QIT2)*, (1999), pp. 107–112.
312. H. Nagaoka, "A new approach to Cramér–Rao bound for quantum state estimation," *IEICE Tech. Rep.*, **IT 89-42**, 228, 9–14 (1989).
313. M. A. Nielsen, "Conditions for a class of entanglement transformations," *Phys. Rev. Lett.*, **83**, 436 (1999); quant-ph/9811053 (1998).
314. M. A. Nielsen, "Continuity bounds for entanglement," *Phys. Rev. A*, **61**, 064301 (2000).
315. M. A. Naĭmark, *Comptes Rendus (Doklady) de l'Acadenie des Sience de l'URSS*, **41**, 9, 359, (1943).
316. M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, Cambridge, 2000).
317. M. A. Nielsen, J. Kempe, "Separable states are more disordered globally than locally," *Phys. Rev. Lett.*, **86**, 5184–5187 (2001); quant-ph/0011117 (2000).
318. T. Ogawa, *A study on the asymptotic property of the hypothesis testing and the channel coding in quantum mechanical systems*, Ph.D. thesis, Graduate School of Information System, University of Electro-Communication (2000).
319. T. Ogawa, H. Nagaoka, "Strong converse to the quantum channel coding theorem," *IEEE Trans. Inf. Theory*, **45**, 2486–2489 (1999); quant-ph/9808063 (1998).
320. T. Ogawa, H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, **46**, 2428–2433 (2000); quant-ph/9906090 (1999).
321. T. Ogawa, H. Nagaoka, "A new proof of the channel coding theorem via hypothesis testing in quantum information theory," *Proc. 2002 IEEE International Symposium on Information Theory*, 73 (2002); quant-ph/0208139 (2002).
322. T. Ogawa, M. Hayashi, "On error exponents in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, **50**, 1368–1372 (2004); quant-ph/0206151 (2002).
323. M. Ohya, D. Petz, *Quantum Entropy and Its Use,* (Springer, Berlin Heidelberg New York, 1993).
324. M. Ohya, D. Petz, N. Watanabe, "On capacities of quantum channels," *Prob. Math. Stat.*, **17**, 179–196 (1997).
325. S. Osawa, H. Nagaoka, "Numerical experiments on the capacity of quantum channel with entangled input states," *IEICE Trans.*, **E84-A**, 2583–2590 (2001); quant-ph/0007115 (2000).
326. M. Owari, M. Hayashi, "Local copying and local discrimination as a study for non-locality of a set," quant-ph/0509062.
327. M. Ozawa, "Quantum measuring processes of continuous observables," *J. Math. Phys.*, **25**, 79 (1984).
328. M. Ozawa, "Quantum limits of measurements and uncertainty principle, in *Quantum Aspects of Optical Communications*, Lecture Notes in Physics, vol. 378, C. Bendjaballah, O. Hirota, S. Reynaud (eds.), (Springer, Berlin Heidelberg New York, 1991), pp. 3–17.

329. M. Ozawa, "Quantum state reduction and the quantum Bayes principle," in *Quantum Communication, Computing, and Measurement*, O. Hirota, A. S. Holevo, C. M. Caves (eds.), (Plenum, New York, 1997), pp. 233–241.
330. M. Ozawa, "An operational approach to quantum state reduction," *Ann. Phys.*, **259**, 121–137 (1997); quant-ph/9706027 (1997).
331. M. Ozawa, "Quantum state reduction: an operational approach," *Fortschr. Phys.*, **46**, 615–625 (1998); quant-ph/9711006 (1997).
332. M. Ozawa, "Operational characterization of simultaneous measurements in quantum mechanics," *Phys. Lett. A*, **275**, 5–11 (2000); quant-ph/9802039 (1998).
333. M. Ozawa, "Measurements of nondegenerate discrete observables," *Phys. Rev. A*, **63**, 062101 (2000); quant-ph/0003033 (2000).
334. M. Ozawa, "Operations, disturbance, and simultaneous measurability," *Phys. Rev. A*, **62**, 032109 (2001); quant-ph/0005054 (2000).
335. M. Ozawa, "Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement," *Phys. Rev. A*, **67**, 042105 (2003); quant-ph/0207121 (2002).
336. M. Ozawa, "Physical content of Heisenberg's uncertainty relation: limitation and reformulation," *Phys. Lett. A*, **318**, 21–29 (2003); quant-ph/0210044 (2002).
337. M. Ozawa, "Uncertainty principle for quantum instruments and computing," *Int. J. Quant. Inf.*, **1**, 569–588 (2003); quant-ph/0310071 (2003).
338. M. Ozawa, "Uncertainty relations for noise and disturbance in generalized quantum measurements," *Ann. Phys.*, **311**, 350–416 (2004); quant-ph/0307057 (2003).
339. M. Ozawa, "On the noncommutative theory of statistical decisions," *Research Reports on Information Sciences*, **A-74** (1980).
340. M. Owari, K. Matsumoto, M. Murao, "Entanglement convertibility for infinite dimensional pure bipartite states," *Phys. Rev. A*, **70**, 050301 (2004); quant-ph/0406141; "Existence of incomparable pure bipartite states in infinite dimensional systems," quant-ph/0312091 (2003).
341. J.-W. Pan, S. Gasparoni, M. Aspelmeyer, T. Jennewein, A. Zeilinger, "Experimental realization of freely propagating teleported qubits," *Nature*, **421**, 721–725 (2003).
342. J.-W. Pan, S. Gasparoni, R. Ursin, G. Weihs, A. Zeilinger, "Experimental entanglement purification of arbitrary unknown states," *Nature*, **423**, 417–422 (2003).
343. D. Petz, D. Petz, "Quasi-entropies for finite quantum systems," *Rep. Math. Phys.*, **23**, 57–65 (1986).
344. D. Petz, *An Invitation to the Algebra of Canonical Commutation Relations*, Leuven Notes in Mathematical and Theoretical Physics, vol. 2 (1990).
345. D. Petz, "Monotone metrics on matrix spaces," *Lin. Alg. Appl.*, **224**, 81–96 (1996).
346. D. Petz, M. Mosonyi: "Stationary quantum source coding," *J. Math. Phys.*, **42**, 48574864 (2001); quant-ph/9912103 (1999).
347. D. Petz, C. Sudár, "Extending the Fisher metric to density matrices," in *Geometry in Present Day Science*, O. E. Barndorff-Nielsen, E. B. V. Jensen (eds.), 21 (World Scientific, Singapore, 1998).
348. D. Petz, G. Toth, *Lett. Math. Phys.*, **27**, 205 (1993).

349. D. Petz, "Monotonicity of quantum relative entropy revisited," *Rev. Math. Phys.*, **15**, 29–91 (2003).
350. A. A. Pomeransky, "Strong superadditivity of the entanglement of formation follows from its additivity," *Phys. Rev. A*, **68**, 032317 (2003); quant-ph/0305056 (2003).
351. H. P. Robertson, "The uncertainty principle," *Phys. Rev.*, **34**, 163 (1929).
352. E. M. Rains, "Bound on distillable entanglement," *Phys. Rev. A*, **60**, 179–184 (1999).
353. E. M. Rains: "A semidefinite program for distillable entanglement," *IEEE Trans. Inf. Theory*, **47**, 2921–2933 (2001); quant-ph/0008047 (2000).
354. M. B. Ruskai, "Beyond strong subadditivity? improved bounds on the contraction of generalized relative entropy," *Rev. Math. Phys.*, **6**, 1147–1161 (1994).
355. M. B. Ruskai, S. Szarek, E. Werner, "An analysis of completely-positive trace-preserving maps on 2×2 matrices," *Lin. Alg. Appl.*, **347**, 159–187 (2002); quant-ph/0101003 (2001).
356. M. B. Ruskai, "Qubit entanglement breaking channels," *Rev. Math. Phys.*, **15**, 643–662 (2003); quant-ph/0302032 (2003).
357. J. J. Sakurai, *Modern Quantum Mechanics*, (Addison-Wesley, Reading, MA, 1985).
358. I. N. Sanov, "On the probability of large deviations of random variables," *Mat. Sbornik*, **42**, 11–44 (1957) (in Russian). English translation: *Selected Translat. Math. Stat.*, **1**, 213–244 (1961).
359. M. Sasaki, M. Ban, S. M. Barnett, "Optimal parameter estimation of a depolarizing channel," *Phys. Rev. A*, **66**, 022308 (2002); quant-ph/0203113 (2002).
360. B. Schumacher, "Quantum coding," *Phys. Rev. A*, **51**, 2738–2747 (1995).
361. B. Schumacher, "Sending quantum entanglement through noisy channels," *Phys. Rev. A*, **54**, 2614–2628 (1996); quant-ph/9604023 (1996).
362. B. Schumacher, M. A. Nielsen, "Quantum data processing and error correction," *Phys. Rev. A*, **54**, 2629 (1996); quant-ph/9604022 (1996).
363. B. Schumacher, M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, **56**, 131 (1997).
364. B. Schumacher, M. D. Westmoreland, "Optimal signal ensembles," *Phys. Rev. A*, **63**, 022308 (2001).
365. A. Shamir, "How to share a secret," *Commun. ACM*, **22**, 612 (1979).
366. C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, **27**, 623–656 (1948).
367. C. E. Shannon, "Certain results in coding theory for noisy channels," *Inf. Control*, **1**, 6–25 (1957).
368. C. Gobby, Z. L. Yuan, A. J. Shields, "Quantum key distribution over 122 km of standard telecom fiber," *Appl. Phys. Lett.*, **84**, 3762–3764 (2004); quant-ph/0412171 (2004).
369. T. Shimono, H. Fan, *Proc. ERATO Conference on Quantum Information Science (EQIS) 2003*, 119–120 (2003).
370. T. Shimono, "Additivity of entanglement of formation of two three-level-antisymmetric states," *Int. J. Quant. Inf.*, **1**, 259–268 (2003); quant-ph/0301011 (2003).
371. P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Phys. Rev. A*, **52**, 2493 (1995).

372. P. W. Shor, "Additivity of the classical capacity of entanglement-breaking quantum channels," *J. Math. Phys.*, **43**, 4334–4340 (2002); quant-ph/0201149 (2002).
373. P. W. Shor, "The quantum channel capacity and coherent information," *Lecture Notes, MSRI Workshop on Quantum Computation* (2002). Avaliable at http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/
374. P. W. Shor, "Equivalence of additivity questions in quantum information theory," *Comm. Math. Phys.*, **246**, 3, 453–473 (2004); quant-ph/0305035 (2003).
375. P. W. Shor, "Capacities of quantum channels and how to find them," *Math. Programm.*, **97**, 311–335 (2003); quant-ph/0304102 (2003).
376. P. W. Shor, J. Preskill, "Simple proof of security of the BB84 quantum key distribution protocol," *Phys. Rev. Lett.*, **85**, 441–444 (2000); quant-ph/0003004 (2000).
377. D. Slepian, J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, **19**, 471 (1973).
378. A. M. Steane, "Multiple particle interference and quantum error correction," *Proc. R. Soc. Lond. A*, **452**, 2551 (1996); quant-ph/9601029 (1996).
379. W. F. Stinespring, "Positive functions on C-algebras," *Proc. Am. Math. Soc.*, **6**, 211 (1955).
380. R. L. Stratonovich, *Izvest. VUZ Radiofiz.*, **8**, 116–141 (1965).
381. R. L. Stratonovich, "The transmission rate for certain quantum communication channels," *Problemy Peredachi Informatsii*, **2**, 45–57 (1966). (in Russian). English translation: *Probl. Inf. Transm.*, **2**, 35–44 (1966.)
382. R. L. Stratonovich, A. G. Vantsjan, *Probl. Control Inf. Theory*, **7**, 161–174 (1978).
383. D. Stucki, N. Gisin, O. Guinnard, G. Ribordy, H. Zbinden, "Quantum key distribution over 67 km with a plug & play system," *New J. Phys.*, **4**, 41.1–41.8 (2002).
384. F. Tanaka, *Investigation on Fisher metric on the classical statistical models and the quantum ones*, master's thesis, Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Japan (2004) (in Japanese).
385. B. M. Terhal, P. Horodecki, "A Schmidt number for density matrices," *Phys. Rev. A*, **61**, 040301(R)(2000); quant-ph/9911114 (1999).
386. B. M. Terhal, K. G. H. Vollbrecht, "Entanglement of formation for isotropic states," *Phys. Rev. Lett.*, **85**, 2625.
387. B. M. Terhal, M. Horodecki, D. W. Leung, D. P. DiVincenzo, "The entanglement of purification," *J. Math. Phys.*, **43**, 4286 (2002).
388. Y. Tsuda, K. Matsumoto, "Quantum estimation for non-differentiable models," *J. Phys. A Math. Gen.*, **38**, 7, 1593–1613 (2005); quant-ph/0207150 (2002).
389. Y. Tsuda, K. Matsumoto, M. Hayashi, "Hypothesis testing for a maximally entangled state," quant-ph/0504203 (2005).
390. Y. Tsuda, B.S. Shi, A. Tomita, M. Hayashi, K. Matsumoto, Y.K. Jiang, "Hypothesis testing for an entangled state produced by spontaneous parametric down conversion," *ERATO conference on Quantum Information Science 2005 (EQIS 05)*, JST, Tokyo, pp. 57–58 (2005).
391. Y. Tsuda, "Estimation of polynomial of complex amplitude of quantum Gaussian states" *Proc. Annual Meeting of the Japan Statistical Society (2005)* (in Japanese).

392. A. Uhlmann, "The 'transition probability' in the state space of *-algebra," *Rep. Math. Phys.*, **9**, 273–279 (1976).
393. A. Uhlmann, "Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory," *Comm. Math. Phys.*, **54**, 21–32 (1977).
394. A. Uhlmann, "Density operators as an arena for differential geometry," *Rep. Math. Phys.*, **33**, 253–263 (1993).
395. K. Usami, Y. Nambu, Y. Tsuda, K. Matsumoto, K. Nakamura, "Accuracy of quantum-state estimation utilizing Akaike's information criterion," *Phys. Rev. A*, **68**, 022314 (2003); quant-ph/0306083 (2003).
396. A. W. van der Vaart, *Asymptotic Statistics*, (Cambridge University Press, Cambridge, 1998).
397. R. M. Van Slyke, R. J. -B. Wets, "A duality theory for abstract mathematical programs with applications to optimal control theory," *J. Math. Anal. Appl.*, **22**, 679–706 (1968).
398. H. L. van Trees, *Detection, Estimation and Modulation Theory, Part 1*, (Wiley, New York, 1968).
399. V. Vedral, M. B. Plenio, "Entanglement Measures and Purification Procedures," *Phys. Rev. A*, **57**, 822 (1998); quant-ph/9707035 (1997).
400. S. Verdú, T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, **40**, 1147–1157 1994.
401. F. Verstraete, J. Dehaene, B. DeMorr, "Local filtering operations on two qubits," *Phys. Rev. A*, **64**, 010101(R) (2001).
402. G. Vidal, "Entanglement of pure states for a single copy," *Phys. Rev. Lett.*, **83**, 1046–1049 (1999); quant-ph/9902033 (1999).
403. G. Vidal, W. Dür, J. I. Cirac, "Entanglement cost of antisymmetric states," quant-ph/0112131v1 (2001).
404. G. Vidal, D. Jonathan, M. A. Nielsen, "Approximate transformations and robust manipulation of bipartite pure state entanglement," *Phys. Rev. A*, **62**, 012304 (2000); quant-ph/9910099 (1999).
405. S. Virmani, M. Sacchi, M.B. Plenio, D. Markham, "Optimal local discrimination of two multipartite pure states," *Phys. Lett. A*, **288**, 62 (2001).
406. S. Virmani, M. B. Plenio, "Construction of extremal local positive-operator-valued measures under symmetry," *Phys. Rev. A*, **67**, 062308 (2003).
407. K. G. H. Vollbrecht, R. F. Werner, "Entanglement measures under symmetry," quant-ph/0010095 (2000).
408. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, (Princeton University Press, Princeton, NJ, 1955). (Originally appeared in German in 1932).
409. J. Walgate, A. J. Short, L. Hardy, V. Vedral, "Local distinguishability of multipartite orthogonal quantum states," *Phys. Rev. Lett.*, **85**, 4972 (2000).
410. X. Wang, H. Fan, "Non-post-selection entanglement concentration by ordinary linear optical devices," *Phys. Rev. A*, **68**, 060302(R) (2003); quant-ph/0302105 (2003).
411. R. S. Ward, R. O. Wells, Jr., *Twistor Geometry and Field Theory*, (Cambridge University Press, Cambridge, 1991).
412. R. F. Werner, "Optimal cloning of pure states," *Phys. Rev. A*, **58**, 1827 (1998).
413. R.F. Werner and A.S. Holevo, "Counterexample to an additivity conjecture for output purity of quantum channels," *J. Math. Phys.*, **43**, 4353 (2002); quant-ph/0203003 (2002).

414. H. Weyl, *The Classical Groups, Their Invariants and Representations*, (Princeton University Press, Princeton, NJ, 1939).
415. A. Winter, "Schumacher's Quantum Coding Revisited," Preprint 99–034, Sonder forschungsbereich 343 "Diskrete Strukturen in der Mathematik," Universität Bielefeld (1999).
416. A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inf. Theory*, **45**, 2481–2485 (1999).
417. A. Winter, *Coding Theorems of Quantum Information Theory*, Ph.D. dissertation, Universität Bielefeld (2000); quant-ph/9907077 (1999).
418. A. Winter, S. Massar, "Compression of quantum measurement operations," *Phys. Rev. A*, **64**, 012311 (2001); quant-ph/0012128.
419. A. Winter, "Scalable programmable quantum gates and a new aspect of the additivity problem for the classical capacity of quantum channels," *J. Math. Phys.*, **43**, 4341–4352 (2002); quant-ph/0108066 (2001).
420. A. Winter, "Extrinsic" and "intrinsic" data in quantum measurements: asymptotic convex decomposition of positive operator valued measures," *Comm. Math. Phys.*, **244**, (1) 157–185 (2004); quant-ph/0109050.
421. A. Winter, "Secret, public and quantum correlation cost of triples of random variables," *Proc. 2005 IEEE International Symposium on Information Theory*, 2270 (2005).
422. A. Winter, private communication (2005).
423. M. M. Wolf, J. Eisert, "Classical information capacity of a class of quantum channels," *New J. Phys.*, **7**, 93 (2005).
424. W. K. Wootters, "Entanglement of Formation of an Arbitrary State of Two Qubits," *Phys. Rev. Lett.*, **80**, 2245 (1998).
425. A. D. Wyner, "The wire-tap channel," *Bell. Syst. Tech. J.*, **54**, 1355–1387 (1975).
426. A. D. Wyner, "The common information of two dependent random variables," *IEEE Trans. Inf. Theory*, **21**, 163–179 (1975).
427. J. Yard, in preparation.
428. T. Yamamoto, M. Koashi, Ş. Özdemir, N. Imoto, "Experimental extraction of an entangled photon pair from two identically decohered pairs," *Nature*, **421**, 343–346 (2003).
429. T. Y. Young, "Asymptotically efficient approaches to quantum-mechanical parameter estimation," *Inf. Sci.*, **9**, 25–42 (1975).
430. D. Yang, M. Horodecki, R. Horodecki, B. Synak-Radtke: "Irreversibility for all bound entangled state," *Phys. Rev. Lett.*, **95**, 190501 (2005); quant-ph/0506138 (2005).
431. H. P. Yuen, R. S. Kennedy, M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inf. Theory*, 125–134 (1975).
432. H. P. Yuen, M. Lax, "Multiple-parameter quantum estimation and measurement of non-selfadjoint observables," *IEEE Trans. Inf. Theory*, **19**, 740 (1973).
433. F. Yura, "Entanglement cost of three-level antisymmetric states," *J. Phys. A Math. Gen.*, **36**, L237–L242 (2003); quant-ph/0302163 (2003).

Index