Additivity of the capacity of depolarizing channels

Akio Fujiwara* and Takashi Hashizumé
Department of Mathematics, Osaka University
Toyonaka, Osaka 560-0043, Japan

Abstract

We show that for the tensor product of two depolarizing channels, the classical capacity is additive. The key observation is a majorisation relation for the eigenvalues of output states.

PACS numbers: 03.67.-a, 03.65.Ud, 89.70.+c

*fujiwara@math.wani.osaka-u.ac.jp

Given quantum channels Γ_1 and Γ_2, do their (classical) capacities [1, 2] satisfy additivity $C(\Gamma_1 \otimes \Gamma_2) = C(\Gamma_1)+C(\Gamma_2)$? In other words, does it suffice to use tensor product states as codewords to achieve the ultimate communication rate through the channel $\Gamma_1 \otimes \Gamma_2$? This is called the additivity problem of the capacity, and is one of the big open problems in quantum information theory. It is conjectured that the additivity may hold (see, for instance, [3, 4]). However, up to date, the additivity has been proved only for three special classes of quantum channels: Γ_1 is arbitrary and $\Gamma_2 = \mathrm{id}$ [3, 5], Γ_1 is arbitrary and Γ_2 is an entanglement-breaking channel [6], Γ_1 is arbitrary and Γ_2 is a unital binary channel [7]. The purpose of this paper is to add to the above list a new class: both Γ_1 and Γ_2 are depolarizing channels. This gives a substantial generalization of [8].

Let $\mathcal{S}(\mathcal{H})$ denote the set of density operators on a Hilbert space $\mathcal{H}$. A quantum channel $\Gamma : \mathcal{S}(\mathbb{C}^n) \to \mathcal{S}(\mathbb{C}^n)$ is called a *depolarizing channel* if it is represented in the form

$$\Gamma(\rho) = d\rho + (1-d)\frac{I^{(n)}}{n},$$

where $I^{(n)}$ is the identity acting on $\mathbb{C}^n$, and d is the parameter describing the magnitude of depolarization. To ensure that Γ is completely positive, d must lie in the closed interval $[-1/(n^2-1), 1]$. In this paper, we further assume that d is nonnegative, that is, $d \in [0, 1]$. The main result is the following.

Theorem 1 *Let Γ_1 and Γ_2 be depolarizing channels on $\mathcal{S}(\mathbb{C}^m)$ and $\mathcal{S}(\mathbb{C}^n)$, where m, n are arbitrary natural numbers. Then $C(\Gamma_1 \otimes \Gamma_2) = C(\Gamma_1) + C(\Gamma_2)$.*

The proof is based on the theory of majorisation [9]. Given $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, let $x^\downarrow = (x_1^\downarrow, \dots, x_n^\downarrow)$ denote the vector that is obtained by rearranging the components of x in the decreasing order. (We use row and column vector representations interchangeably.) Given $x, y \in \mathbb{R}^n$, we say that x is *majorised* by y (and denote $x \prec y$) if

$$\sum_{j=1}^{k} x_j^\downarrow \le \sum_{j=1}^{k} y_j^\downarrow, \qquad (1 \le k \le n),$$

with equality for $k = n$. For $x, y \in \mathbb{R}^n$, we denote $x \sim y$ if $x^\downarrow = y^\downarrow$. A real valued function f on $\mathbb{R}^n$ is called *Schur-convex* if $x \prec y$ implies $f(x) \le f(y)$. It is called *Schur-concave* if $-f$ is Schur-convex. We list some basic properties of majorisation: (a) For $x, y \in \mathbb{R}^n$, $x \prec y$ if and only if $x = Ty$ for a doubly stochastic matrix T. (b) If $x, y \in \mathbb{R}^n$ and $u, v \in \mathbb{R}^m$ satisfy $x \prec y$ and $u \prec v$, then $(x, u), (y, v) \in \mathbb{R}^{n+m}$ satisfy $(x, u) \prec (y, v)$. (c) Let A be an $n \times n$ Hermitian matrix and let $\lambda(A)$ denotes the vector whose components are eigenvalues of A specified in any order. Then for all $k = 1, \dots, n$

$$\sum_{j=1}^{k} \lambda_j^\downarrow(A) = \max \sum_{j=1}^{k} \langle v_j, Av_j \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product of $\mathbb{C}^n$ and the maximum is taken over all orthonormal k-tuples of vectors $\{v_1, \dots, v_k\}$ in $\mathbb{C}^n$. (This is referred to as Ky Fan's maximum principle.) (d) For a Hermitian matrix A, $\lambda(\mathrm{Diag}(A)) \prec \lambda(A)$ holds, where $\mathrm{Diag}(A)$ denotes the diagonal matrix that corresponds to the diagonal part of A. (This is referred to as Schur's theorem.) The next Lemma is a slight generalization of Schur's theorem.

Lemma 2 *For a Hermitian matrix A and $0 \le \mu \le 1$, it holds that*

$$\lambda(\mu A + (1-\mu)\,\mathrm{Diag}(A)) \prec \lambda(A).$$

Proof For all $k = 1, \ldots, n$, we have

$$
\begin{aligned}
&\sum_{j=1}^{k} \lambda_j^{\downarrow}(\mu A + (1-\mu)\,\mathrm{Diag}(A)) \\
&\quad \leq \sum_{j=1}^{k} \lambda_j^{\downarrow}(\mu A) + \sum_{j=1}^{k} \lambda_j^{\downarrow}((1-\mu)\,\mathrm{Diag}(A)) \\
&\quad = \mu \sum_{j=1}^{k} \lambda_j^{\downarrow}(A) + (1-\mu) \sum_{j=1}^{k} \lambda_j^{\downarrow}(\mathrm{Diag}(A)) \\
&\quad \leq \mu \sum_{j=1}^{k} \lambda_j^{\downarrow}(A) + (1-\mu) \sum_{j=1}^{k} \lambda_j^{\downarrow}(A) \\
&\quad = \sum_{j=1}^{k} \lambda_j^{\downarrow}(A).
\end{aligned}
$$

The first inequality follows easily from Ky Fan's maximum principle and the second inequality is due to Schur's theorem. In addition

$$
\sum_{j=1}^{n} \lambda_j^{\downarrow}(\mu A + (1-\mu)\,\mathrm{Diag}(A)) = \mathrm{tr}\ A = \sum_{j=1}^{n} \lambda_j^{\downarrow}(A).
$$

The claim was verified. $\square$

Proof of Theorem 1 We need only prove that $C(\Gamma_1 \otimes \Gamma_2) \leq C(\Gamma_1) + C(\Gamma_2)$. We first recall that [10]

$$
\begin{aligned}
C(\Gamma_1) &= \log m - \min_{\tau \in \mathcal{S}(\mathbb{C}^m)} H(\Gamma_1(\tau)), \\
C(\Gamma_2) &= \log n - \min_{\tau \in \mathcal{S}(\mathbb{C}^n)} H(\Gamma_2(\tau)).
\end{aligned}
$$

Here $H(\rho) := -\mathrm{Tr}\,\rho \log \rho$ is the von Neumann entropy, and the minimum is attained at the extreme boundary. In what follows we denote by $\partial_e \mathcal{S}(\mathcal{H})$ the extreme boundary of $\mathcal{S}(\mathcal{H})$, that is, the set of pure states on $\mathcal{H}$. On the other hand, since the entropy function is Schur-concave, we see from Lemma 3 below that for all $\sigma \in \partial_e \mathcal{S}(\mathbb{C}^m \otimes \mathbb{C}^n)$, there exist $\tau_1 \in \partial_e \mathcal{S}(\mathbb{C}^m)$ and $\tau_2 \in \partial_e \mathcal{S}(\mathbb{C}^n)$ such that

$$
H(\Gamma_1 \otimes \Gamma_2(\sigma)) \geq H(\Gamma_1 \otimes \Gamma_2(\tau_1 \otimes \tau_2)).
$$

This entailes, in particular, that there exist $\hat{\tau}_1 \in \partial_e \mathcal{S}(\mathbb{C}^m)$ and $\hat{\tau}_2 \in \partial_e \mathcal{S}(\mathbb{C}^n)$ such that

$$
\min_{\sigma \in \mathcal{S}(\mathbb{C}^m \otimes \mathbb{C}^n)} H(\Gamma_1 \otimes \Gamma_2(\sigma)) = H(\Gamma_1 \otimes \Gamma_2(\hat{\tau}_1 \otimes \hat{\tau}_2)).
$$

Then by using the minimax formula for the channel capacity [11, 12]

$$
C(\Gamma) = \inf_{\rho \in \mathcal{S}(\mathcal{H})} \sup_{\sigma \in \mathcal{S}(\mathcal{H})} D(\Gamma(\sigma)||\Gamma(\rho)),
$$

where $D(\sigma||\rho) := \mathrm{Tr}\,\sigma(\log\sigma - \log\rho)$ is the quantum relative entropy, we have

$$\begin{aligned} C(\Gamma_1 \otimes \Gamma_2) &= \inf_\rho \sup_\sigma D(\Gamma_1 \otimes \Gamma_2(\sigma)||\Gamma_1 \otimes \Gamma_2(\rho)) \\ &\leq \sup_\sigma D(\Gamma_1 \otimes \Gamma_2(\sigma)||\Gamma_1 \otimes \Gamma_2(I^{(mn)}/mn)) \\ &= \sup_\sigma D(\Gamma_1 \otimes \Gamma_2(\sigma)||I^{(mn)}/mn) \\ &= \log mn - H(\Gamma_1 \otimes \Gamma_2(\hat{\tau}_1 \otimes \hat{\tau}_2)) \\ &= \log m - H(\Gamma_1(\hat{\tau}_1)) + \log n - H(\Gamma_2(\hat{\tau}_2)) \\ &= C(\Gamma_1) + C(\Gamma_2). \end{aligned}$$

The assertion was verified. $\square$

Let m, n be natural numbers that satisfy $m \geq n$. According to the Schmidt decomposition, given a pure state $\sigma \in \partial_e \mathcal{S}(\mathbb{C}^m \otimes \mathbb{C}^n)$, there exist orthonormal bases $\{\mathbf{e}_i\}_{i=1}^m$, $\{\mathbf{f}_j\}_{j=1}^n$, and an n-dimensional probability vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ such that

$$\sigma = \sigma_\alpha := \sum_{i=1}^n \sum_{j=1}^n \sqrt{\alpha_i \alpha_j}\, |\mathbf{e}_i\rangle\langle\mathbf{e}_j| \otimes |\mathbf{f}_i\rangle\langle\mathbf{f}_j|.$$

We can regard the probability vector α as describing the degree of entanglement of the pure state σ_α. The following lemma is the key to the proof of Theorem 1.

Lemma 3 *Let Γ_1 and Γ_2 be depolarizing channels on $\mathcal{S}(\mathbb{C}^m)$ and $\mathcal{S}(\mathbb{C}^n)$. For any pure state $\sigma_\alpha \in \partial_e \mathcal{S}(\mathbb{C}^m \otimes \mathbb{C}^n)$, it holds that*

$$\lambda(\Gamma_1 \otimes \Gamma_2(\sigma_\alpha)) \prec \lambda(\Gamma_1 \otimes \Gamma_2(\sigma_{(1,0,\ldots,0)})).$$

Proof To demonstrate the basic idea, we first prove the case when $m = 3$ and $n = 2$. Letting

$$\Gamma_1(\rho) = c\rho + (1-c)\frac{I^{(3)}}{3}, \quad \Gamma_2(\sigma) = d\sigma + (1-d)\frac{I^{(2)}}{2},$$

we see that

$$\begin{aligned} \Gamma_1(|\mathbf{e}_i\rangle\langle\mathbf{e}_j|) &= c\,|\mathbf{e}_i\rangle\langle\mathbf{e}_j| + \delta_{ij}(1-c)\frac{I^{(3)}}{3}, \\ \Gamma_2(|\mathbf{f}_i\rangle\langle\mathbf{f}_j|) &= d\,|\mathbf{f}_i\rangle\langle\mathbf{f}_j| + \delta_{ij}(1-d)\frac{I^{(2)}}{2}. \end{aligned}$$

Since both Γ_1 and Γ_2 are unitarily covariant, we can take without loss of generality that $\{\mathbf{e}_i\}_{i=1}^3$ and $\{\mathbf{f}_j\}_{j=1}^2$ are the standard bases of $\mathbb{C}^3$ and $\mathbb{C}^2$. It is then easy to observe that the Kronecker product representation of $\Gamma_1 \otimes \Gamma_2(\sigma_\alpha)$ is partitioned into blocks of 2×2 matrices as follows:

$$\Gamma_1 \otimes \Gamma_2(\sigma_\alpha) = \frac{1}{6} \begin{bmatrix} A_\alpha^{11} & A_\alpha^{12} & O \\ A_\alpha^{21} & A_\alpha^{22} & O \\ O & O & \hat{D}_\alpha \end{bmatrix}.$$

Further, by a suitable rearrangement of the orthonormal basis $\{\mathbf{e}_i \otimes \mathbf{f}_j\}_{ij}$ of $\mathbb{C}^3 \otimes \mathbb{C}^2$, it is represented as

$$\Gamma_1 \otimes \Gamma_2(\sigma_\alpha) = \frac{1}{6}\, A_\alpha \oplus D_\alpha \oplus \hat{D}_\alpha, \tag{1}$$

where

$$\begin{aligned}
A_\alpha &= \begin{bmatrix} \alpha_1(1+2c)(1+d)+\alpha_2(1-c)(1-d) & 6\sqrt{\alpha_1\alpha_2}cd \\ 6\sqrt{\alpha_1\alpha_2}cd & \alpha_1(1-c)(1-d)+\alpha_2(1+2c)(1+d) \end{bmatrix}, \\
D_\alpha &= \begin{bmatrix} \alpha_1(1+2c)(1-d)+\alpha_2(1-c)(1+d) & 0 \\ 0 & \alpha_1(1-c)(1+d)+\alpha_2(1+2c)(1-d) \end{bmatrix}, \\
\hat{D}_\alpha &= \begin{bmatrix} \alpha_1(1-c)(1+d)+\alpha_2(1-c)(1-d) & 0 \\ 0 & \alpha_1(1-c)(1-d)+\alpha_2(1-c)(1+d) \end{bmatrix}.
\end{aligned}$$

According to the basic property (b) of majorisation mentioned above, it suffices to prove a series of relations (i) $\lambda(A_\alpha) \prec \lambda(A_{(1,0)})$, (ii) $\lambda(D_\alpha) \prec \lambda(D_{(1,0)})$, and (iii) $\lambda(\hat{D}_\alpha) \prec \lambda(\hat{D}_{(1,0)})$ separately. If $c = d = 0$, these relations are obvious. We thus assume that either c or d is nonzero.

We first prove (i). Choose real numbers p_{21}, p_{22} appropreately so that

$$P = \begin{bmatrix} \sqrt{\alpha_1} & \sqrt{\alpha_2} \\ p_{21} & p_{22} \end{bmatrix}$$

becomes an orthogonal matrix. (For example, let $p_{21} = -\sqrt{\alpha_2}$ and $p_{22} = \sqrt{\alpha_1}$.) Then

$${}^tPA_{(1,0)}P = \begin{bmatrix} \alpha_1(1+2c)(1+d)+\alpha_2(1-c)(1-d) & \sqrt{\alpha_1\alpha_2}\,(3c+2d+cd) \\ \sqrt{\alpha_1\alpha_2}\,(3c+2d+cd) & \alpha_1(1-c)(1-d)+\alpha_2(1+2c)(1+d) \end{bmatrix}.$$

Observe that the diagonal entries of A_α and ${}^tPA_{(1,0)}P$ are identical, and that the off-diagonal entries are in the constant ratio

$$\mu := \frac{6cd}{3c+2d+cd}.$$

We thus have

$$A_\alpha = \mu\, {}^tPA_{(1,0)}P + (1-\mu)\,\mathrm{Diag}({}^tPA_{(1,0)}P).$$

Since $0 \le \mu \le 1$, it follows from Lemma 2 that

$$\lambda(A_\alpha) \prec \lambda({}^tPA_{(1,0)}P) \sim \lambda(A_{(1,0)}).$$

We next prove (ii). Let T be the doubly stochastic matrix defined by

$$T := \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_1 \end{bmatrix}.$$

Then it follows from the basic property (a) that

$$\lambda(D_\alpha) \sim T\lambda(D_{(1,0)}) \prec \lambda(D_{(1,0)}).$$

The relation (iii) is proved in the same way:

$$\lambda(\hat{D}_\alpha) \sim T\lambda(\hat{D}_{(1,0)}) \prec \lambda(\hat{D}_{(1,0)}).$$

This completes the proof of Lemma 3 for $m = 3$ and $n = 2$.

We can generalize the above argument for any natural numbers m and n. Assume without loss of generality that $m \geq n$. Let c, d be the parameters of Γ_1, Γ_2 as before. By a straightforward computation, we have a direct sum representation of $\Gamma_1 \otimes \Gamma_2(\sigma_\alpha)$ analogous to Eq. (1) as follows:

$$\begin{aligned}&\Gamma_1 \otimes \Gamma_2(\sigma_\alpha)\\ &= \frac{1}{mn} A_\alpha \oplus D_\alpha^{(1)} \oplus \cdots \oplus D_\alpha^{(n-1)} \oplus \overbrace{\hat{D}_\alpha \oplus \cdots \oplus \hat{D}_\alpha}^{m-n}.\end{aligned}$$

Here A_α is the $n \times n$ matrix whose ith diagonal entry is

$$\alpha_i(mc + 1 - c)(nd + 1 - d) + (1 - \alpha_i)(1 - c)(1 - d),$$

and the (i, j)th off-diagonal entry is $\sqrt{\alpha_i \alpha_j} mncd$. On the other hand, $D_\alpha^{(j)}$ is the $n \times n$ diagonal matrix whose ith diagonal entry is

$$\begin{aligned}&\alpha_i(mc + 1 - c)(1 - d) + \alpha_{i+j}(1 - c)(nd + 1 - d)\\ &+(1 - \alpha_i - \alpha_{i+j})(1 - c)(1 - d),\end{aligned}$$

where the subscript of α should be understood modulo n. Further, $\hat{D}_\alpha$ is the $n \times n$ diagonal matrix whose ith diagonal entry is

$$\alpha_i(1 - c)(nd + 1 - d) + (1 - \alpha_i)(1 - c)(1 - d).$$

We show (i) $\lambda(A_\alpha) \prec \lambda(A_{(1,0,\ldots,0)})$, (ii) $\lambda(D_\alpha^{(j)}) \prec \lambda(D_{(1,0,\ldots,0)}^{(j)})$ for $j = 1, \ldots, n-1$, and (iii) $\lambda(\hat{D}_\alpha) \prec \lambda(\hat{D}_{(1,0,\ldots,0)})$.

Let us first prove (i). Choose real numbers p_{ij} $(2 \leq i \leq n, 1 \leq j \leq n)$ appropreately so that

$$P = \begin{bmatrix} \sqrt{\alpha_1} & \cdots & \sqrt{\alpha_n} \\ p_{21} & \cdots & p_{2n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}$$

becomes an orthogonal matrix. Then it is shown that

$$A_\alpha = \mu\, {}^tPA_{(1,0,\ldots,0)}P + (1 - \mu)\,\mathrm{Diag}({}^tPA_{(1,0,\ldots,0)}P),$$

where

$$\mu := \frac{mncd}{mc + nd + (mn - m - n)cd}.$$

Since

$$\begin{aligned} mncd &= mcd + ncd + (mn - m - n)cd \\ &\leq mc + nd + (mn - m - n)cd, \end{aligned}$$

it follows that $0 \le \mu \le 1$. Then by Lemma 2

$$\lambda(A_\alpha) \prec \lambda({}^tPA_{(1,0,\dots,0)}P) \sim \lambda(A_{(1,0,\dots,0)}).$$

We next prove (ii). Let T be the doubly stochastic matrix defined by

$$T := \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_2 & \alpha_3 & \cdots & \alpha_1 \\ \vdots & \vdots & & \vdots \\ \alpha_n & \alpha_1 & \cdots & \alpha_{n-1} \end{bmatrix},$$

and let the vector $v^{(j)} \in \mathbb{R}^n$ ($j = 1, \dots, n-1$) be such that the first and the $(j+1)$th components are $(mc+1-c)(1-d)$ and $(1-c)(nd+1-d)$, and the rest are all $(1-c)(1-d)$. Then it is easy to see that $\lambda(D_\alpha^{(j)}) \sim Tv^{(j)}$ and $\lambda(D_{(1,0,\dots,0)}^{(j)}) \sim v^{(j)}$. This proves (ii).

Finally (iii) is proved as follows.

$$\lambda(\hat{D}_\alpha) \sim T\lambda(\hat{D}_{(1,0,\dots,0)}) \prec \lambda(\hat{D}_{(1,0,\dots,0)}).$$

This completes the proof. □

Now we proceed to a question inspired by Lemma 3: does $\alpha \prec \beta$ imply $\lambda(\Gamma_1 \otimes \Gamma_2(\sigma_\alpha)) \prec \lambda(\Gamma_1 \otimes \Gamma_2(\sigma_\beta))$? The next theorem gives a partial answer.

Theorem 4 *Let Γ_1 and Γ_2 be depolarizing channels on $\mathcal{S}(\mathbb{C}^m)$ and $\mathcal{S}(\mathbb{C}^2)$. If $\alpha \prec \beta$, then*

$$\lambda(\Gamma_1 \otimes \Gamma_2(\sigma_\alpha)) \prec \lambda(\Gamma_1 \otimes \Gamma_2(\sigma_\beta)).$$

Proof According to the proof of Lemma 3

$$\Gamma_1 \otimes \Gamma_2(\sigma_\alpha) = \frac{1}{2m} A_\alpha \oplus D_\alpha \oplus \overbrace{\hat{D}_\alpha \oplus \cdots \oplus \hat{D}_\alpha}^{m-2}.$$

We show (i) $\lambda(A_\alpha) \prec \lambda(A_\beta)$, (ii) $\lambda(D_\alpha) \prec \lambda(D_\beta)$, and (iii) $\lambda(\hat{D}_\alpha) \prec \lambda(\hat{D}_\beta)$.

Let $\alpha^\downarrow = (a, 1-a)$ and $\beta^\downarrow = (b, 1-b)$. Then $\alpha \prec \beta$ implies $1/2 \le a \le b \le 1$. The maximum eigenvalue of A_α is

$$\lambda_1^\downarrow(A_\alpha) = \frac{1}{2}\left(\operatorname{tr} A_\alpha + \sqrt{(\operatorname{tr} A_\alpha)^2 - 4\det A_\alpha}\right).$$

Here we observe that

$$\operatorname{tr} A_\alpha = (mc+1-c)(1+d) + (1-c)(1-d)$$

is independent of a, and that

$$\begin{aligned} \det A_\alpha \quad = \quad & a(1-a)\{mc(1-d) + 2d(1-c)\} \\ & \qquad \times\{mc(1+3d) + 2d(1-c)\} \\ & +(\text{terms independent of } a). \end{aligned}$$

As a consequence, $\lambda_1^{\downarrow}(A_\alpha)$ is monotone increasing in $a(\in [1/2, 1])$. In particular, the inequality $\lambda_1^{\downarrow}(A_\alpha) \leq \lambda_1^{\downarrow}(A_\beta)$ follows from the fact that $1/2 \leq a \leq b \leq 1$. On the other hand

$$\lambda_1^{\downarrow}(A_\alpha) + \lambda_2^{\downarrow}(A_\alpha) = \operatorname{tr} A_\alpha = \lambda_1^{\downarrow}(A_\beta) + \lambda_2^{\downarrow}(A_\beta).$$

The relation (i) is thus verified.

We next prove (ii). Since $\alpha \prec \beta$, there is a doubly stochastic matrix T such that $\alpha = T\beta$. Then by a direct computation, we see that $\lambda(D_\alpha) \sim T\lambda(D_\beta)$, which implies (ii). The relation (iii) is also proved in the same way. $\square$

It is an open question whether Theorem 4 can be generalized for a depolarizing channel Γ_2 acting on $\mathcal{S}(\mathbb{C}^n)$ of arbitrary dimension n.

We have shown that for the tensor product of two depolarizing channels having nonnegative depolarization parameters, the classical capacity is additive (Theorem 1). The key observation was a majorisation relation for the eigenvalues of output states (Lemma 3). Unfortunately this argument does not apply to those depolarizing channels which have negative depolarization parameters. We also studied a generalization of Lemma 3 and obtained a partial answer (Theorem 4).

References

[1] A. S. Holevo, IEEE Trans. Inform. Theory, **44**, 269 (1998).

[2] B. Schumacher and M. D. Westmoreland, Phys. Rev. A, **56**, 131 (1997).

[3] G. G. Amosov, A. S. Holevo, and R. F. Werner, Probl. Inform. Transm., **36**, 25 (2000).

[4] S. Osawa and H. Nagaoka, IEICE Trans. Fundamentals, **E84-A**, 2583 (2001); eprint quant-ph/0007115.

[5] B. Schumacher and M. D. Westmoreland, eprint quant-ph/0004045.

[6] P. W. Shor, eprint quant-ph/0201149.

[7] C. King, eprint quant-ph/0103156.

[8] D. Bruss, L. Faoro, C. Macchiavello, and G. M. Palma, J. Mod. Opt., **47**, 325 (2000).

[9] R. Bhatia, *Matrix Analysis* (Springer, New York, 1997).

[10] A. Fujiwara and H. Nagaoka, IEEE Trans. Inform. Theory, **44**, 1071 (1998).

[11] M. Ohya, D. Petz, and N. Watanabe, Prob. Math. Stat., **17**, 170 (1997).

[12] B. Schumacher and M. D. Westmoreland, Phys. Rev. A, **63**, 022308 (2001).