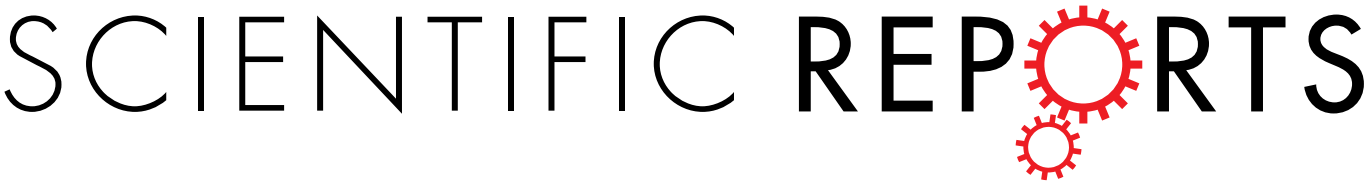


OPEN

Holevo Capacity of Discrete Weyl Channels

Junaid ur Rehman[1], Youngmin Jeong[1], Jeong San Kim[2] & Hyundong Shin[1]

Received: 10 May 2018
Accepted: 12 November 2018
Published online: 29 November 2018

Holevo capacity is the maximum rate at which a quantum channel can reliably transmit classical information without entanglement. However, calculating the Holevo capacity of arbitrary quantum channels is a nontrivial and computationally expensive task since it requires the numerical optimization over all possible input quantum states. In this paper, we consider discrete Weyl channels (DWCs) and exploit their symmetry properties to model DWC as a classical symmetric channel. We characterize lower and upper bounds on the Holevo capacity of DWCs using simple computational formulae. Then, we provide a sufficient and necessary condition where the upper and lower bounds coincide. The framework in this paper enables us to characterize the exact Holevo capacity for most of the known special cases of DWCs.

One of the fundamental tasks in the context of information theory is to compute the maximum rate at which information can be reliably transmitted[1,2]. Classical channels have the capability of transmitting classical information only. On the contrary, quantum channels are more rich in terms of communication tasks[3,4]. Trivially, quantum channels are capable of transmitting quantum information. However, due to the versatile nature and unique features of quantum mechanics, it is possible to associate multiple communication tasks with a quantum channel[5]. Thus, we have classical capacity, quantum capacity, private classical capacity, and entanglement-assisted classical capacity of a quantum channel. All of theses correspond to different information communication tasks[6–9].

The calculation of various capacities involves an optimization task that is not easy to perform. For example, the capacity of a classical channel is given by a single letter formula—the mutual information between input and output of the channel—maximized over the probability distribution of the input random variable[10]. Efficient methods exist that can perform this maximization[11,12]. On the contrary, capacities (except the entanglement-assisted classical capacity) of a quantum channel are given in terms of regularization of asymptotically many channel uses. These regularized formulae are mathematically intractable in general and put forth an unsolvable optimization problem[13]. Simplification of these formulae is not possible due to the nonadditive and nonconvex natures of capacities of quantum channels[14–16]. The need of regularization, however, can be removed either (1) if the capacity of the channel is additive, or (2) if we restrict the optimization to be on the individual channel use. For example, unital qubit channels[17] and entanglement breaking channels[18] are known to be additive and thus their classical capacity can be computed without the need of regularization. Similarly, for the task of classical communication over a quantum channel, one can prohibit the use of inputs states correlated over multiple uses of the channel–effectively allowing optimization on the individual channel use only–to obtain a lower bound on the classical capacity of a quantum channel. This notion of capacity is known as the Holevo capacity. Even with such a simplification of the problem, the calculation remains considerably demanding. As a matter of fact, calculation of the Holevo capacity falls in the category of NP-complete problems[15,19].

This multilayer difficulty has stimulated a good amount of research in the field of quantum information theory. Different researchers have taken different routes to accomplish this seemingly impossible task. For example, different definitions of capacities have been proposed[20], analytical expressions for the special channels have been found[21], and some bounds that are additive and easier to calculate have been computed[22] to solve the problem of regularization. While for solving the difficulty of calculation, exploiting special properties of a given channel[23], and methods that can approximate the capacity upto a fixed a posteriori error have been proposed[24].

In this work we give easy to compute lower and upper bounds on the Holevo capacity of discrete Weyl channels (DWCs). Our employed approach involves modeling the DWC as a classical symmetric channel and using the existing results from the classical information theory to lower bound the Holevo capacity of a DWC. The upper bound is based on the majorization relation of any possible output state of a DWC with the most ordered

[1]Department of Electronic Engineering, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do, 17104, Korea. [2]Department of Applied Mathematics and Institute of Natural Sciences, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do, 17104, Korea. Correspondence and requests for materials should be addressed to Y.J. (email: yjeong@khu.ac.kr) or H.S. (email: hshin@khu.ac.kr)

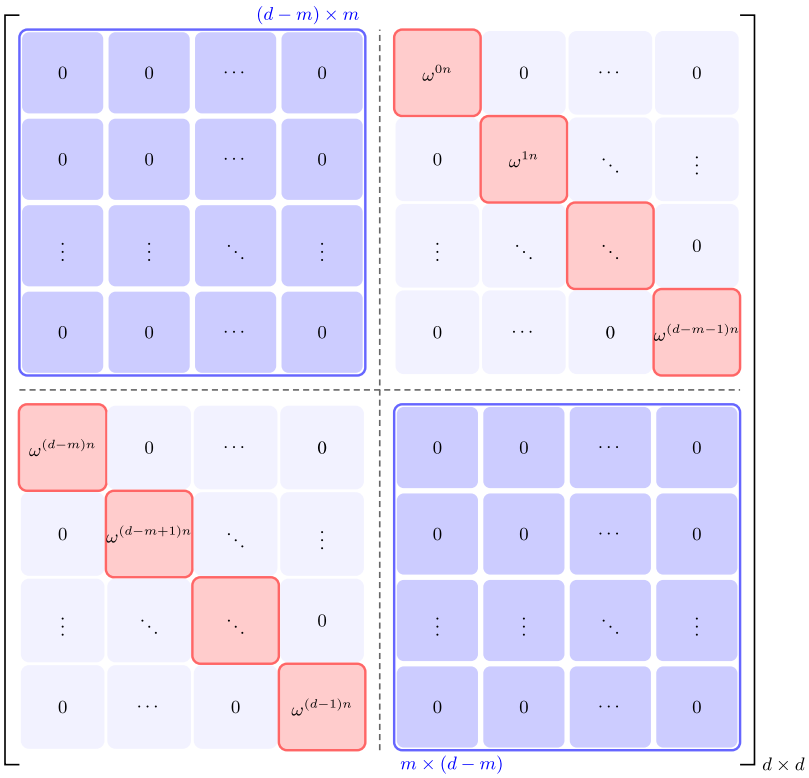


Figure 1. The general structure of a Weyl operator $\boldsymbol{W}_{nm}$ in an arbitrary dimension d.

state based on the channel parameters. We give a necessary and sufficient condition for which the two bounds coincide. We find that this condition is met for the known special cases (Pauli qubit channel, and the qudit depolarizing channel) of DWC and hence we can recover the exact capacity expression for these cases. Through numerical examples we show that the coincidence of two bounds is sufficient but not necessary for the lower bound to give exact capacity.

Discrete Weyl Channel

A quantum state $\boldsymbol{\rho}$ on the Hilbert space is a positive operator with unit trace (i.e., a density operator). We consider the Hilbert space of finite dimension d. The state is said to be *pure* if it has the form $\boldsymbol{\rho} = |\psi\rangle\langle\psi|$. We usually denote a pure state simply by a ket e.g., $|\psi\rangle$, which is a column vector in the Hilbert space. A quantum channel $\mathcal{N}$: $\boldsymbol{\rho} \to \mathcal{N}(\boldsymbol{\rho})$ is a completely positive trace preserving (CPTP) map transforming the input state $\boldsymbol{\rho}$ to an output state $\mathcal{N}(\boldsymbol{\rho})$. The map can be specified in terms of Kraus operators $\{\boldsymbol{A}_i\}$ as $\mathcal{N}(\boldsymbol{\rho}) = \sum_i \boldsymbol{A}_i \boldsymbol{\rho} \boldsymbol{A}_i^\dagger$ where $\sum_i \boldsymbol{A}_i^\dagger \boldsymbol{A}_i = \boldsymbol{I}_d$ and $\boldsymbol{I}_d$ is the identity operator on the d-dimensional Hilbert space. For a random unitary channel, it is possible to represent Kraus operators as $\boldsymbol{A}_i = \sqrt{p_i}\boldsymbol{B}_i$, such that the channel applies an operator $\boldsymbol{B}_i$ on the input state with the probability p_i[25].

Let $\boldsymbol{\sigma}_0 = \boldsymbol{I}_2$ be the 2×2 identity matrix, and

$$\boldsymbol{\sigma}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{\sigma}_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \boldsymbol{\sigma}_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1}$$

be the Pauli matrices. The Pauli qubit channel, denoted by $\mathcal{N}_\mathrm{p}(\boldsymbol{\rho})$, is then defined as

$$\mathcal{N}_\mathrm{p}(\boldsymbol{\rho}) = \sum_{i=0}^{3} p_i \boldsymbol{\sigma}_i \boldsymbol{\rho} \boldsymbol{\sigma}_i^\dagger \tag{2}$$

which is a random unitary channel.

Discrete Weyl operators are a non-Hermitian generalization of Pauli operators for dimension d[26]. A Weyl operator $\boldsymbol{W}_\mathrm{nm}$ on the d-dimensional Hilbert space is defined as[27]

$$\boldsymbol{W}_{nm} = \sum_{k=0}^{d-1} \omega^{kn} |k\rangle \langle (k+m) \mod d| \tag{3}$$

for $n,\ m = 0, 1, \cdots,\ d-1$; $\omega = \exp(2\pi\iota/d)$; and $|k\rangle$ is the kth basis vector in the computational basis (for notational convenience, the indexing of entries of vectors and matrices start from 0). A general structure of a d-dimensional Weyl operator $\boldsymbol{W}_{nm}$ is shown in Fig. 1.

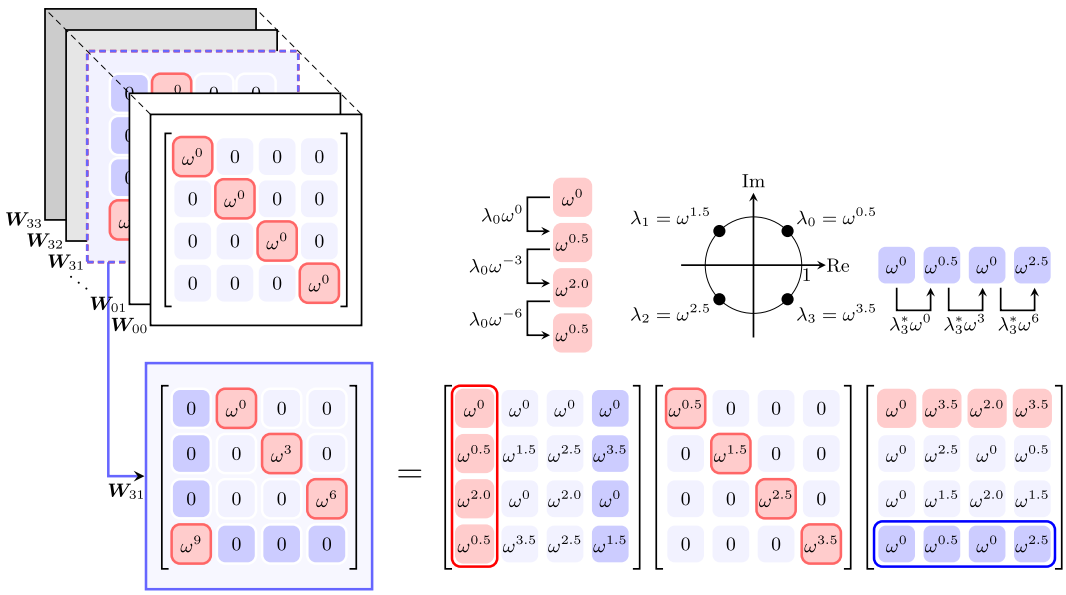


Figure 2. A schematic illustration for the structure of discrete Weyl operator $\boldsymbol{W}_{31}$ on a 4-dimensional Hilbert space. Each eigenvalue λ_s and eigenvector $|\lambda_s\rangle$ can be found using (4) and (30), respectively.

Property 1. *A Weyl operator $\boldsymbol{W}_{nm}$, when applied on a d-dimensional vector α, up-shifts the entries of α by m locations and rotates ith entry (according to new indexing) by a phase of ω^{in}. We refer to this property as shift and phase operation of Weyl operators.*

Eigenvalues of a Weyl operator $\boldsymbol{W}_{nm}$ are given by (see supplementary material),

$$\lambda_s = \omega^{mn\frac{(d-1)}{2}+s} \tag{4}$$

where $s \in \{(mk - nj) \mod d\}$ for $j,\ k = 0, \cdots,\ d-1$. A schematic illustration for the Weyl operator $\boldsymbol{W}_{31}$ on a 4-dimensional Hilbert space is given in Fig. 2. Note that Weyl operators operating on a prime dimensional Hilbert space have d distinct eigenvalues (and we can simply state that $s = 0, 1, \cdots,\ d-1$) except for $\boldsymbol{W}_{00}$. On the other hand, some Weyl operators of a composite dimension may have repeated eigenvalues. This repetition of eigenvalues restrains us from deriving general forms of our results directly. We circumvent this problem by first presenting our results for the Hilbert space of a prime dimension, and then show that an alternate formulation of our results can be applied to the case of a composite dimensional Hilbert space as well.

A DWC, denoted by $\mathcal{N}_{\mathrm{dw}}(\boldsymbol{\rho})$, is a generalization of the Pauli qubit channel[1], defined in terms of discrete Weyl operators as

$$\mathcal{N}_{\mathrm{dw}}(\boldsymbol{\rho}) = \sum_{n=0}^{d-1}\sum_{m=0}^{d-1} p_{nm}\boldsymbol{W}_{nm}\boldsymbol{\rho}\boldsymbol{W}_{nm}^{\dagger} \tag{5}$$

where $\boldsymbol{W}_{nm}$ acts on the input state $\boldsymbol{\rho}$ with probability p_{nm}.

The Holevo capacity of a quantum channel is defined as[6,28]

$$\chi(\mathcal{N}) = \sup_{\{p_i,\boldsymbol{\rho}_i\}}\left[S\left(\sum_i p_i\mathcal{N}(\boldsymbol{\rho}_i)\right) - \sum_i p_i S(\mathcal{N}(\boldsymbol{\rho}_i))\right] \tag{6}$$

where p_i is the *a priori* probability of input state $\boldsymbol{\rho}_i$; $S(\boldsymbol{\rho}) = -\mathrm{Tr}(\boldsymbol{\rho}\log\boldsymbol{\rho})$ is the von Neumann entropy, and $\mathcal{N}(\boldsymbol{\rho})$ is the output state produced by the action of channel $\mathcal{N}$ on the input state $\boldsymbol{\rho}$. The Holevo capacity corresponds to the maximum rate of classical information when input states are restricted to be separable, i.e., the inputs of the channel are not entangled over multiple uses.

Lemma 1. *If the input state of a DWC operating on a d-dimensional Hilbert space is an eigenstate of a d-dimensional Weyl operator $\boldsymbol{W}_{nm}$, then the output state is diagonal in the eigenbasis of $\boldsymbol{W}_{nm}$.*

Proof . See Methods section. □

As a consequence of the above Lemma, we can choose the set of input states to be d orthogonal eigenvectors of some Weyl operator $\boldsymbol{W}_{nm}$, and measure the output in the eigenbasis of $\boldsymbol{W}_{nm}$. The uncertainty at the output of the channel in this case is purely classical in nature. In this sense, a DWC is behaving as a classical channel, transitioning a distinguishable state into an unknown but perfectly distinguishable state. We completely characterize the simulated classical channel in terms of channel transition matrix in the following Proposition.

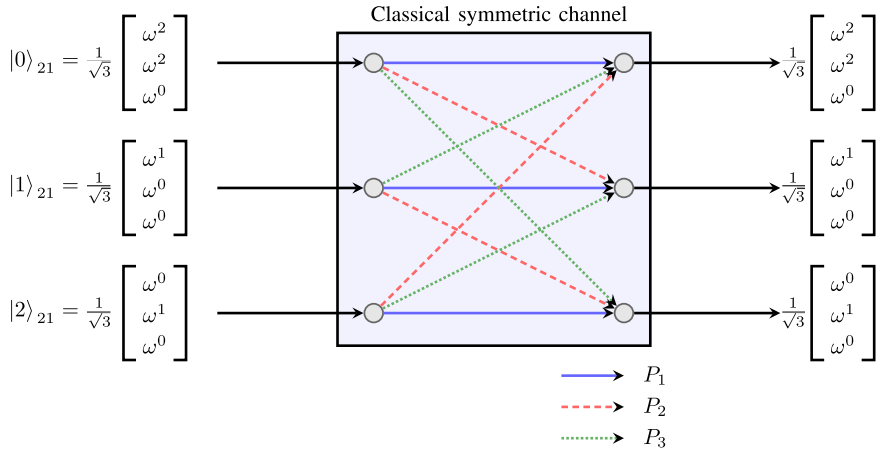


Figure 3. An example DWC for $d=3$ driven by the eigenstates of $\boldsymbol{W}_{21}$.

Proposition 1. *A DWC of a prime dimension d with orthonormal eigenstates of* $\boldsymbol{W}_{nm}$ *as the input states behaves as a classical symmetric channel with the following transition matrix*

$$\boldsymbol{T}_{nm} = \begin{bmatrix} P_1 & P_2 & \cdots & P_d \\ P_d & P_1 & \cdots & P_{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ P_2 & P_3 & \cdots & P_1 \end{bmatrix}, \qquad (n,\ m) \neq (0,0) \tag{7}$$

where

$$P_k = \sum_{ij:\omega^{mi-nj}=\omega^{k-1}} p_{ij}. \tag{8}$$

Proof. See Methods section. □

As an example, a DWC driven by the eigenstates of $\boldsymbol{W}_{21}$ with $d=3$ is shown in Fig. 3. In this example, we have $P_1 = p_{00} + p_{21} + p_{12}$, $P_2 = p_{20} + p_{11} + p_{02}$, and $P_3 = p_{10} + p_{01} + p_{22}$.

Results

Based on the proposition 1, we give the following simple and natural lower bound on the Holevo capacity of a DWC:

Theorem 1. *The Holevo capacity* $\chi(\mathcal{N}_{\mathrm{dw}})$ *of the channel in* (5) *with a prime d is bounded as*

$$\chi(\mathcal{N}_{\mathrm{dw}}) \geq \log_2(d) - \min_{n,m} H(\text{row of } \boldsymbol{T}_{nm}), \qquad (n,\ m) \neq (0,0) \tag{9}$$

where $\boldsymbol{T}_{nm}$ *is the channel transition matrix of the* (n, m) *th symmetric channel obtained by fixing the eigenstates of* $\boldsymbol{W}_{nm}$ *as the signal states and* $H(\cdot)$ *is the Shannon entropy.*

Proof . See Methods section. □

The restriction on d to be a prime number is primarily because the repetition of eigenvalues of $\boldsymbol{W}_{nm}$ of a composite d does not allow us to construct the channel transition matrix $\boldsymbol{T}_{nm}$. The following remark provides us an alternative approach to lower bound the Holevo capacity of DWC of any d.

Remark 1. *It is straightforward to show that* $H(\text{row of } \boldsymbol{T}_{nm}) = S(\mathcal{N}_{\mathrm{dw}}(|\lambda\rangle\ \langle\lambda|_{nm}))$ *when d is prime, where* $|\lambda\rangle\langle\lambda|_{nm}$ *is the density matrix of any eigenstate of* $\boldsymbol{W}_{nm}$*. Therefore, we can equivalently calculate*

$$\chi(\mathcal{N}_{\mathrm{dw}}) \geq \log_2(d) - \min_{n,m} S\left(\mathcal{N}_{\mathrm{dw}}(|\lambda\rangle\langle\lambda|_{nm})\right) \tag{10}$$

for prime d. Then, we can extend (10) *to any d by replacing the optimization on any* $\boldsymbol{\rho}$ *in* (20) *with the optimization on the eigenstates of* $\boldsymbol{W}_{nm}$ *only.*

Theorem 2. *Let us define a vector* $\boldsymbol{\zeta}(\boldsymbol{p}) \in \mathbb{R}^d$ *such that*

$$\boldsymbol{\zeta}(\boldsymbol{p}) = \boldsymbol{S}\boldsymbol{p}^{\downarrow} \tag{11}$$

where the elements of $\boldsymbol{p}^{\downarrow}$ *are the elements of vector* $\boldsymbol{p} \in \mathbb{R}^{d^2}$ *in descending order; the matrix* $\boldsymbol{S} \in \mathbb{R}^{d\times d^2}$ *is given by*

$$S = \begin{bmatrix} \mathbf{1}_d^T & \mathbf{0}_d^T & \mathbf{0}_d^T & \cdots & \mathbf{0}_d^T \\ \mathbf{0}_d^T & \mathbf{1}_d^T & \mathbf{0}_d^T & \cdots & \mathbf{0}_d^T \\ \mathbf{0}_d^T & \mathbf{0}_d^T & \mathbf{1}_d^T & \cdots & \mathbf{0}_d^T \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_d^T & \mathbf{0}_d^T & \mathbf{0}_d^T & \cdots & \mathbf{1}_d^T \end{bmatrix} \tag{12}$$

where $(\cdot)^T$ *denotes the transpose operation, and* $\mathbf{1}_d$ *and* $\mathbf{0}_d$ *are all-one and all-zero vectors of d elements, respectively. Then, the Holevo capacity of a DWC is*

$$\chi(\mathcal{N}_{\mathrm{dw}}) \leq \log_2(d) - H(\boldsymbol{\zeta}(\boldsymbol{p})), \tag{13}$$

where $\boldsymbol{p} = [p_{00}\ p_{01}\ \cdots\ p_{nm}]^T$*, whose elements are probabilities associated with respective Weyl operators* $\boldsymbol{W}_{nm}$*.*

Proof. See Methods section. □

In a d-dimensional Hilbert space, d^2 Weyl operators are defined whose indices are given in the form of 2-tuples, e.g., (i, j). We define a set $\mathcal{W}$ that contains all the d^2 indices of defined Weyl operators. We call a set $\mathcal{D}$ a d-set if all its elements $\mathcal{D}_i$ for $i = 0, \cdots, \ d-1$ are non-overlapping d element subsets of $\mathcal{W}$

$$\mathcal{D} = \left\{\mathcal{D}_i | \mathcal{D}_i \subset_d \mathcal{W},\ \mathcal{D}_i \bigcap \mathcal{D}_j = \varnothing \text{ for } i \neq j,\ i,\ j = 0, \cdots, \ d-1\right\} \tag{14}$$

where $\mathcal{A} \subset_d \mathcal{B}$ means that $\mathcal{A}$ is a d-element subset of $\mathcal{B}$, $\varnothing$ is the empty set, and $\mathcal{A} \cap \mathcal{B}$ gives a set whose elements are the common elements of $\mathcal{A}$ and $\mathcal{B}$. In the d dimensional Hilbert space, there are

$$\frac{1}{d!}\prod_{i=0}^{d-1}\binom{d^2 - id}{d} \tag{15}$$

different possible d-sets, where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

are the binomial coefficients.

A d-set $\mathcal{D}$ whose all elements $\mathcal{D}_t$ satisfy the property

$$mi - nj \bmod d = k_t, \ \ \forall (i,\ j) \in \mathcal{D}_t \tag{16}$$

for some n, m, and some constants k_t is called an achievable d-set. For example

$$\mathcal{D} = \{\{(0,0),(2,1),(1,2)\},\ \{(2,0),(1,1),(0,2)\},\ \{(1,0),(0,1),(2,2)\}\} \tag{17}$$

is an achievable d-set for $(n, m) = (2, 1)$ but

$$\mathcal{D} = \{\{(0,0),(0,1),(1,2)\},\ \{(2,0),(1,1),(2,2)\},\ \{(1,0),(2,1),(0,2)\}\} \tag{18}$$

is a d-set which is not achievable.

Theorem 3. *We arrange the elements of p in nonincreasing order and collect the indices of* p_{nm} *while preserving the order to form a d-set. The bounds of Theorem 1, and Theorem 2 coincide if and only if (resp. only if) the obtained d-set is achievable and d is a prime number (resp. a composite number).*

Proof. See Methods section. □

Remark 2. *If the two bounds coincide, we have*

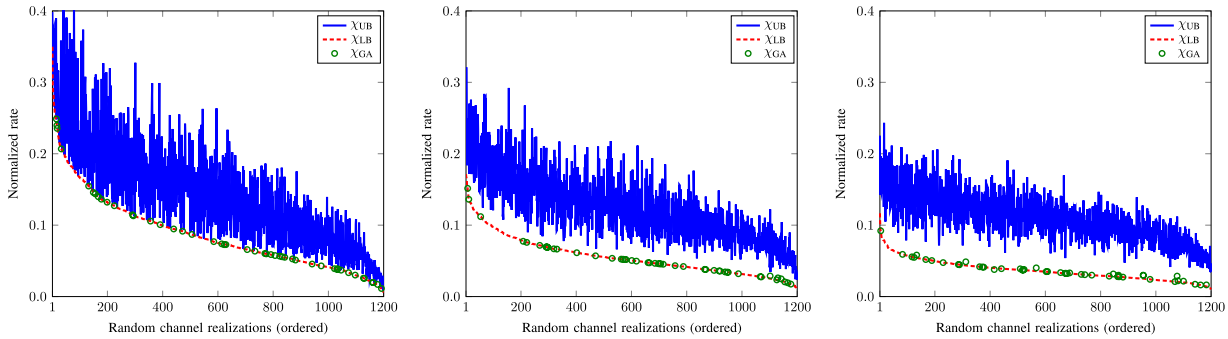
$$\chi(\mathcal{N}_{\mathrm{dw}}) = \log_2(d) - \min_{n,m} H(\text{row of } \boldsymbol{T}_{nm}), \qquad (n,\ m) \neq (0,0). \tag{19}$$

However, the converse is not true as will be shown by the numerical examples in the next section.

Discussion

An efficient approximation for the capacity of classical-quantum channels has previously been discussed without exploiting any special properties of a given channel[24]. For example, it takes 40,154 seconds in order to approximate the Holevo capacity of a Pauli qubit channel with a posteriori error of 1.940×10^{-3}. In contrast to existing methods, the average time to calculate the (lower) bound in this paper is of the order 10^{-4} seconds even for large d by virtue of the use of special properties of DWCs.

We have strong numerical evidence that the lower bound is tighter and is saturated more often even when the two bounds do not coincide, as shown in the Fig. 4(a–c) where the upper (χ_{UB}) and the lower (χ_{LB}) bounds



(a) $d = 3$ (b) $d = 4$ (c) $d = 5$

Figure 4. χ_{UB}, χ_{LB}, and χ_{GA} of random channel realizations (in decreasing order of χ_{UB}) when $d = 3, 4, 5$.

(normalized by $\log_2(d)$) are plotted for 1200 random channel realizations for $d = 3$, 4, and 5, respectively. In these figures, Holevo capacity by using[23]

$$\chi(\mathcal{N}_{dw}) = \log_2(d) - \min_{\boldsymbol{\rho}} S\left(\mathcal{N}_{dw}(\boldsymbol{\rho})\right) \tag{20}$$

with the optimization performed via genetic algorithm (χ_{GA}) is also presented. Comparison of χ_{LB}, χ_{UB}, and χ_{GA} shows that the frequency of coincidence of two bounds as well as the frequency of the saturation of the lower bound is higher for the case of $d = 3$.

Our bounds not only ease the requirement of optimization for the calculation of tight bounds for a general DWC, but also allows to recover the analytic expressions for the special cases of DWC. For example, here we recover the analytic expression for the classical capacity of a qudit depolarizing channel using the approach developed above. A quantum depolarizing channel transforms an input state to the output state according to the following map

$$\mathcal{N}_d(\boldsymbol{\rho}) = (1 - \mu)\boldsymbol{\rho} + \mu\boldsymbol{\pi} \tag{21}$$

where $\boldsymbol{\pi} = \boldsymbol{I}_d/d$ is the maximally mixed state on the output Hilbert space. In terms of Weyl operators,

$$\boldsymbol{\pi} = \frac{1}{d^2}\sum_{n,m=0}^{d-1} \boldsymbol{W}_{nm}\boldsymbol{\rho}\boldsymbol{W}_{nm}^{\dagger}. \tag{22}$$

Thus, we can rewrite equation (21) as

$$\mathcal{N}_d(\boldsymbol{\rho}) = \left(1 - \mu + \frac{\mu}{d^2}\right)\boldsymbol{\rho} + \frac{\mu}{d^2}\sum_{n,m=0(n,m)\neq(0,0)}^{d-1} \boldsymbol{W}_{nm}\boldsymbol{\rho}\boldsymbol{W}_{nm}^{\dagger}. \tag{23}$$

Therefore

$$p_{00} = 1 - \mu + \frac{\mu}{d^2}, \quad p_{nm} = \frac{\mu}{d^2} \quad \forall\ (n,\ m) \neq (0,\ 0) \tag{24}$$

which shows that all d-sets (whether achievable or not) are equivalent in terms of summation of p_{nm} over the elements $\mathcal{D}_i$. Therefore, we can choose an ordering of p_{nm} such that the condition of Theorem 3 is satisfied and we can use equation (13) to calculate the Holevo capacity. From equation (21) and the output vector of $\boldsymbol{\zeta}(\boldsymbol{p}) = (r_0,\ r_1,\ \cdots,\ r_d)$, we see that

$$r_0 = 1 - \mu + \frac{\mu}{d}, \quad r_i = \frac{\mu}{d} \quad \text{for} \quad i = 1, \cdots,\ d - 1. \tag{25}$$

Thus, the Holevo capacity $\chi(\mathcal{N}_d)$ of this channel is

$$\chi(\mathcal{N}_d) = \log_2(d) + \left(1 - \mu + \frac{\mu}{d}\right)\log_2\left(1 - \mu + \frac{\mu}{d}\right) + (d - 1)\frac{\mu}{d}\log_2\left(\frac{\mu}{d}\right) \tag{26}$$

which is equal to the classical capacity of the quantum depolarizing channel[21].

Additionally, it is easy to see that for a Pauli qubit channel ($d = 2$), there are 3 possible d-sets which are all achievable. Therefore, both bounds are exact for the Pauli qubit (and *all its special cases*) channel. With simple algebraic manipulations one can obtain the analytic expressions for the capacities of any of the special cases of the Pauli qubit channel[24].

From Theorem 3, we can also define special channels for which the two bounds always coincide. This approach gives us a class of quantum channels whose exact Holevo capacity can readily be calculated. We define two such

channels here and call them one-parameter depolarizing-like, and two-parameter depolarizing-like channels, respectively.

The one-parameter depolarizing-like channel is defined as

$$\mathcal{N}_{\mathrm{d1}}(\boldsymbol{\rho}) = (1-\xi)\boldsymbol{W}_{ij}\boldsymbol{\rho}\boldsymbol{W}_{ij}^{\dagger} + \xi\boldsymbol{\pi}, \tag{27}$$

whose exact Holevo capacity is same as (26) with the depolarizing parameter ξ.

The two-parameter depolarizing-like channel is

$$\mathcal{N}_{\mathrm{d2}}(\boldsymbol{\rho}) = (1-\eta)\boldsymbol{W}_{ij}\boldsymbol{\rho}\boldsymbol{W}_{ij}^{\dagger} + (1-\kappa)\boldsymbol{W}_{nm}\boldsymbol{\rho}\boldsymbol{W}_{nm}^{\dagger} + (\eta+\kappa-1)\boldsymbol{\pi} \tag{28}$$

where $0 \leq \eta,\ \kappa \leq 1$, and $1 \leq \eta + \kappa \leq 2$. This channel is a further generalization of the one-parameter depolarizing-like channel. The exact Holevo capacity of this channel can readily be calculated by Theorem 3.

In this work we modeled a DWC as a classical symmetric channel for the task of classical communication. Through this modeling, we presented a simple to compute lower bound on the Holevo capacity of a given DWC of an arbitrary dimension. We also gave an intuitive upper bound which coincides with the lower bound under a certain condition. This (sufficient and necessary for a prime d, and necessary for a composite d) condition, however, is not frequently met despite the frequent convergence of the lower bound to the actual Holevo capacity as shown by the numerical examples. The lower bound was derived by noting the similarity of a quantum channel with a classical channel. An interesting future direction is to find similar cases where the results of classical information theory (which is more mature despite being a special case of quantum information theory) can be applied on the problems of quantum information theory with a little or no modification. Similarly, based on the equality of upper and lower bounds, one can define special channels for which these bounds always coincide. Such characterization of quantum channels can give us a class of channels whose exact Holevo capacity can readily be calculated.

Methods

Proof of Lemma 1. Since the DWC is a random unitary channel, the output of the channel is merely the state obtained by randomly applying one of the d^2 Weyl operators on the input. Thus, we need to show that operation of $\boldsymbol{W}_{ij}$ on any eigenstate of $\boldsymbol{W}_{nm}$ results into an eigenstate of $\boldsymbol{W}_{nm}$.

Let

$$|\lambda\rangle = [\alpha_0,\ \alpha_1,\ \cdots,\ \alpha_{d-1}]^T \tag{29}$$

be a normalized eigenvector of $\boldsymbol{W}_{nm}$ with the corresponding eigenvalue λ. From the eigenvalue relation $\boldsymbol{W}_{nm}|\lambda\rangle = \lambda|\lambda\rangle$, and due to the property 1, we get the following relation among the entries of vector of (29)

$$\alpha_{(m+k) \bmod d} = \lambda\omega^{-nk}\alpha_k, \tag{30}$$

where the eigenvalues λ are equidistant points on the unit circle (see Fig. 2). Since we have obtained this relation from the condition of eigenvector, any vector satisfying above relation will be an eigenvector of $\boldsymbol{W}_{nm}$.

Now let us consider the effect of any $\boldsymbol{W}_{ij}$ on the vector of (29). To this end, we let $\boldsymbol{W}_{ij}|\lambda\rangle = |\beta\rangle$, and recall property 1 again to write

$$\boldsymbol{W}_{ij}|\lambda\rangle = \begin{bmatrix} \alpha_j \\ \omega^i\alpha_{(j+1) \bmod d} \\ \vdots \\ \omega^{ki}\alpha_{(j+k) \bmod d} \\ \vdots \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \\ \vdots \end{bmatrix} = |\beta\rangle. \tag{31}$$

i.e., the kth entry of $|\beta\rangle$ is $\omega^{ki}\alpha_{(j+k) \bmod d}$.

If the elements of $|\beta\rangle$ exhibit a similar relation as (30), $|\beta\rangle$ is also an eigenvector of $\boldsymbol{W}_{nm}$. Repeated use of (30) gives the following relation between the entries of $|\beta\rangle$

$$\beta_{(m+k) \bmod d} = \lambda\omega^{mi-nj}\omega^{-nk}\beta_k \tag{32}$$

which essentially bears the same form as (30); because $\lambda\omega^{mi-nj}$ is another eigenvalue of $\boldsymbol{W}_{nm}$. Hence the vector $|\beta\rangle = \boldsymbol{W}_{ij}|\lambda\rangle$ is an eigenvector of $\boldsymbol{W}_{nm}$. Since the output state is a statistical mixture of orthonormal eigenstates of $\boldsymbol{W}_{nm}$, it is diagonal in the same basis, i.e., in the eigenbasis of $\boldsymbol{W}_{nm}$.

Proof of Proposition 1. Let the input state be an eigestate $|\lambda\rangle$ of $\boldsymbol{W}_{nm}$ corresponding to the eigenvalue λ. From the proof of Lemma 1, the application of $\boldsymbol{W}_{ij}$ transforms the input state to the eigenstate of $\boldsymbol{W}_{nm}$ corresponding to the eigenvalue $\lambda\omega^{mi-nj}$. Since $\omega = \exp(2\pi\iota/d)$, ω^{mi-nj} is always from the set $\{\omega^0,\ \omega^1,\ \cdots,\ \omega^{d-1}\}$. Therefore, we can define,

$$P_k = \sum_{ij:\omega^{mi-nj}=\omega^{k-1}} p_{ij} \tag{33}$$

as the transition probability of $|\lambda\rangle$ to the orthogonal state $|\lambda\omega^{k-1}\rangle$. We can define the complete set of transition probabilities P_k, for $k = 1, 2, \cdots, d$ only if $\boldsymbol{W}_{nm}$ does not have any repeated eigenvalues which is guaranteed only if d is prime and $(n, m) \neq (0, 0)$ (note the similarity between ω^{mi-nj} and the expression for s in the definition of eigenvalues).

Furthermore, we notice that the rows of $\boldsymbol{T}_{nm}$ are permutations of each other and its columns are permutation of each other. Therefore, $\boldsymbol{T}_{nm}$ in (7) defines a classical symmetric channel.

Proof of Theorem 1. From proposition 1 we know that in this setting DWC acts as a classical symmetric channel. Since the capacity of a symmetric channel with d inputs and outputs is given by[2]

$$C_{\text{Symmetric}} = \log_2(d) - H(\text{row of transition matrix}), \tag{34}$$

and we have restricted our input states to be from the eigenstates of Weyl operators, thus

$$\chi(\mathcal{N}_{\text{dw}}) \geq \log_2(d) - \min_{n,m} H \text{ (row of } \boldsymbol{T}_{nm}), \qquad (n, m) \neq (0, 0)$$

where the condition $(n, m) \neq (0, 0)$ along with the condition on d to be prime ensures that we can model the given DWC as a classical symmetric channel with the channel transition matrix $\boldsymbol{T}_{nm}$ by virtue of Proposition 1.

Proof of Theorem 2. For a vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, we denote x_i in non-increasing order as

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}, \tag{35}$$

and denote the vector $\boldsymbol{x}^{\downarrow} = (x_{[1]}, x_{[2]}, \cdots, x_{[n]})$ of elements of $\boldsymbol{x}$ rearranged in nonincreasing order. We denote by $\boldsymbol{x} \prec \boldsymbol{y}$ and say $\boldsymbol{x}$ is majorized by $\boldsymbol{y}$ if

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]} \qquad \text{for } k = 1, \cdots, n \tag{36}$$

with strict equality when $k = n$. For two Hermitian operators $\boldsymbol{A}$ and $\boldsymbol{B}$, we denote $\boldsymbol{A} \prec \boldsymbol{B}$ if $\boldsymbol{\lambda}(\boldsymbol{A}) \prec \boldsymbol{\lambda}(\boldsymbol{B})$, where $\boldsymbol{\lambda}(\boldsymbol{A})$ is the vector of eigenvalues of $\boldsymbol{A}$.

Let γ be the optimal input state, then the Holevo capacity of a DWC is[23]

$$\chi(\mathcal{N}_{\text{dw}}) = \log_2(d) - S(\mathcal{N}_{\text{dw}}(\gamma)). \tag{37}$$

We can rewrite (13) as

$$\chi(\mathcal{N}_{\text{dw}}) \leq \log_2(d) - S(\boldsymbol{\rho}), \tag{38}$$

where $\boldsymbol{\rho}$ is some state with the eigenvalues q_i given by the elements of $\boldsymbol{\zeta}(\boldsymbol{p})$. Comparing (37) and (38), our claim simplifies to

$$S(\boldsymbol{\rho}) \leq S(\mathcal{N}_{\text{dw}}(\gamma)), \tag{39}$$

or from the Schur concavity of von Neumann entropy[29]

$$\mathcal{N}_{\text{dw}}(\gamma) \prec \boldsymbol{\rho}, \tag{40}$$

where

$$\mathcal{N}_{\text{dw}}(\gamma) = \sum_{n=0}^{d-1} \sum_{m=0}^{d-1} p_{nm} \boldsymbol{W}_{nm} \gamma \boldsymbol{W}_{nm}^{\dagger}. \tag{41}$$

Eigendecomposition of $\boldsymbol{\rho}$ can be written as

$$\boldsymbol{\rho} = \sum_{k=0}^{d-1} q_k \boldsymbol{\rho}_k \tag{42}$$

$$= \sum_{k=0}^{d-1} q_k \boldsymbol{S}_k \boldsymbol{\sigma} \boldsymbol{S}_k^{\dagger} \tag{43}$$

where $\boldsymbol{\sigma}$, and $\boldsymbol{\rho}_k$ are some pure states; $\text{Tr}\{\boldsymbol{\rho}_i \boldsymbol{\rho}_j\} = 1$ if $i = j$, and 0 otherwise; and $\boldsymbol{S}_k$ are some unitary operators defined by the relation $\boldsymbol{S}_k \boldsymbol{\sigma} \boldsymbol{S}_k^{\dagger} = \boldsymbol{\rho}_k$. We note that we are free to choose *any* $\boldsymbol{\rho}$ as long it has eigenvalues q_k. This freedom translates to the choice of $\boldsymbol{\rho}_k$, and hence to $\boldsymbol{S}_k$.

Equation (40) is true if and only if[30], [Theorem 5]

$$\sum_{n=0}^{d-1} \sum_{m=0}^{d-1} p_{nm} \boldsymbol{W}_{nm} \gamma \boldsymbol{W}_{nm}^{\dagger} = \sum_i s_i \boldsymbol{U}_i \boldsymbol{\rho} \boldsymbol{U}_i^{\dagger} \tag{44}$$

for some probability vector $\mathbf{s}$ with elements s_i and some unitary matrices $\boldsymbol{U}_i$. We write

$$\sum_i s_i \boldsymbol{U}_i \boldsymbol{\rho} \boldsymbol{U}_i^\dagger = \sum_{i=0}^{d-1} s_i \boldsymbol{U}_i \left(\sum_{k=0}^{d-1} q_k \boldsymbol{S}_k \boldsymbol{\sigma} \boldsymbol{S}_k^\dagger \right) \boldsymbol{U}_i^\dagger \tag{45}$$

$$= \sum_{i=0}^{d-1} \sum_{k=0}^{d-1} r_{ik} \boldsymbol{U}_i \boldsymbol{S}_k \boldsymbol{\sigma} \boldsymbol{S}_k^\dagger \boldsymbol{U}_i^\dagger \tag{46}$$

$$= \sum_{i=0}^{d-1} \sum_{k=0}^{d-1} r_{ik} \boldsymbol{U}_i \boldsymbol{S}_k \boldsymbol{V} \boldsymbol{\gamma} \boldsymbol{V}^\dagger \boldsymbol{S}_k^\dagger \boldsymbol{U}_i^\dagger \tag{47}$$

where we can write $\boldsymbol{\sigma} = \boldsymbol{V} \boldsymbol{\gamma} \boldsymbol{V}^\dagger$ because both $\boldsymbol{\sigma}$ and $\boldsymbol{\gamma}$ are pure states, and we can obtain $r_{ik} = p_{ik}$ due to[30], [Theorem 4]. Without a loss of generality we can assume both $\boldsymbol{\sigma}$ and $\boldsymbol{\gamma}$ to be the basis states of a basis set each, i.e., $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 \in \mathcal{B}_{\boldsymbol{\sigma}}$, and $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \in \mathcal{B}_{\boldsymbol{\gamma}}$. There is also no loss of generality in assuming $\mathcal{B}_{\boldsymbol{\sigma}}$ to be the computational basis. Under these assumptions, the unitary $\boldsymbol{V}$ is the change of basis unitary from $\mathcal{B}_{\boldsymbol{\gamma}}$ to the computational basis, i.e.,

$$\boldsymbol{V} = \sum_{j=0}^{d-1} |j\rangle \langle \boldsymbol{\gamma}|_j . \tag{48}$$

We need to find $\mathbf{U}_i$, and $\mathbf{S}_k$, such that

$$\boldsymbol{U}_i \boldsymbol{S}_k \boldsymbol{V} = \boldsymbol{W}_{ik} \tag{49}$$

or

$$\boldsymbol{U}_i \boldsymbol{S}_k = \boldsymbol{W}_{ik} \boldsymbol{V}^\dagger \tag{50}$$

$$= \sum_{j=0}^{d-1} \omega^{ji} |j\rangle \langle (j+k) \mod d| \sum_{j'=0}^{d-1} \left| \boldsymbol{\gamma}_{j'} \right\rangle \langle j'| \tag{51}$$

$$= \sum_{j=0}^{d-1} \sum_{j'=0}^{d-1} \omega^{ji} |j\rangle \langle j'| \ \langle (j+k) \mod d| \left| \boldsymbol{\gamma}_{j'} \right\rangle . \tag{52}$$

Choosing

$$\boldsymbol{S}_k = \sum_{j'=0}^{d-1} \left| \boldsymbol{\gamma}_{(j'-k) \mod d} \right\rangle \langle j'|, \quad \text{and} \quad \boldsymbol{U}_i = \sum_{j=0}^{d-1} \omega^{ji} |j\rangle \langle j| \tag{53}$$

satisfies the above product (the indexing of j and of $\boldsymbol{\gamma}_j$ is arbitrary except for $j = 0$), as well as the orthogonality of $\boldsymbol{\rho}_k = \boldsymbol{S}_k \boldsymbol{\sigma} \boldsymbol{S}_k^\dagger$. Therefore, (13) is an upper bound on the Holevo capacity of a DWC.

Proof of Theorem 3. We first observe that the condition on the summation in (8) for the lower bound, and the condition on a d-set to be achievable (16) are essentially the same and result in the same d-element partitioning and ordering of p_{nm}. Thus, in a prime dimension d, every achievable d-set corresponds to a classical symmetric channel that can be simulated by DWC for some n, m.

On the other hand, the upper bound is obtained by ordering the elements of p_{nm} in a nonincreasing order. Therefore, the achievability of the d-set formed by the indices of p_{nm} when the p_{nm} are arranged in a nonincreasing order is sufficient for the existence of a simulated classical symmetric channel of prime dimension that achieves the upper bound. Similarly, since the correspondence of achievable d-sets to a simulated classical symmetric channel is bijective, therefore the conincidence of two bounds necessarily implies the achievability of the d-set formed above.

For a composite d, the correspondence between the simulated classical symmetric channel to the achievable d-sets is injective-only. Therefore the above condition is necessary but no longer sufficient for the coincidence of two bounds.

References

1. Wilde, M. M. *Quantum Information Theory*, 2 edn. (Cambridge University Press, UK 2017).
2. Cover, T. M. & Thomas, J. A. *Elements of Information Theory*, 2 edn. (John Wiley & Sons, USA 2012)
3. Ur Rehman, J., Qaisar, S., Jeong, Y. & Shin, H. Security of a control key in quantum key distribution. *Mod. Phys. Lett. B* **31**, 1750119, https://doi.org/10.1142/S0217984917501196 (2017).
4. Qaisar, S., Ur Rehman, J., Jeong, Y. & Shin, H. Practical deterministic secure quantum communication in a lossy channel. *Progr. Theor. Exp. Phys.* **2017**, 041A01 (2017).
5. Zaman, F., Jeong, Y. & Shin, H. Counterfactual Bell-state analysis. *Sci. Rep.* **8**, 14641 (2018).
6. Holevo, A. S. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory* **44**, 269–273 (1998).
7. Devetak, I. The private classical capacity and quantum capacity of a quantum channel. *IEEE Trans. Inf. Theory* **51**, 44–55 (2005).
8. Bennett, C. H., Shor, P. W., Smolin, J. A. & Thapliyal, A. V. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Trans. Inf. Theory* **48**, 2637–2655 (2002).

9. Bennett, C. H., Shor, P. W., Smolin, J. A. & Thapliyal, A. V. Entanglement-assisted classical capacity of noisy quantum channels. *Phys. Rev. Lett.* **83**, 3081–3084 (1999).
10. Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–423 (1948).
11. Blahut, R. Computation of channel capacity and rate-distortion functions. *IEEE Trans. Inf. Theory* **18**, 460–473 (1972).
12. Arimoto, S. An algorithm for computing the capacity of arbitrary discrete memoryless channels. *IEEE Trans. Inf. Theory* **18**, 14–20 (1972).
13. Cubitt, T. *et al.* Unbounded number of channel uses may be required to detect quantum capacity. *Nat. Commun.* **6**, 6739 (2015).
14. Smith, G. & Yard, J. Quantum communication with zero-capacity channels. *Science* **321**, 1812–1815 (2008).
15. Elkouss, D. & Strelchuk, S. Nonconvexity of private capacity and classical environment-assisted capacity of a quantum channel. *Phys. Rev. A* **94**, 040301 (2016).
16. Hastings, M. B. Superadditivity of communication capacity using entangled inputs. *Nat. Phys.* **5**, 255–257 (2009).
17. King, C. Additivity for unital qubit channels. *J. Math. Phys.* **43**, 4641–4653 (2002).
18. Shor, P. W. Additivity of the classical capacity of entanglement-breaking quantum channels. *J. Math. Phys.* **43**, 4334–4340 (2002).
19. Beigi, S. & Shor, P. W. On the complexity of computing zero-error and Holevo capacity of quantum channels. *arXiv:0709.2090* (2008).
20. Winter, A. & Yang, D. Potential capacities of quantum channels. *IEEE Trans. Inf. Theory* **62**, 1415–1424 (2016).
21. King, C. The capacity of the quantum depolarizing channel. *IEEE Trans. Inf. Theory* **49**, 221–229 (2003).
22. Fukuda, M. & Gour, G. Additive bounds of minimum output entropies for unital channels and an exact qubit formula. *IEEE Trans. Inf. Theory* **63**, 1818–1828 (2017).
23. Cortese, J. Holevo-Schumacher-Westmoreland channel capacity for a class of qudit unital channels. *Phys. Rev. A* **69**, 022302 (2004).
24. Sutter, D., Sutter, T., Esfahani, P. M. & Renner, R. Efficient approximation of quantum channel capacities. *IEEE Trans. Inf. Theory* **62**, 578–598 (2016).
25. Nielsen, M. A. & Chuang, I. L. Quantum Computation and Quantum Information: 10th Anniversary Edition, 10th edn. (Cambridge University Press, New York, NY, USA 2011).
26. Bertlmann, R. A. & Krammer, P. Bloch vectors for qudits. *J. Phys. A* **41**, 235303 (2008).
27. Weyl, H. Quantenmechanik und gruppentheorie. *Zeitschrift für Physik* **46**, 1–46 (1927).
28. Schumacher, B. & Westmoreland, M. D. Sending classical information via noisy quantum channels. *Phys. Rev. A* **56**, 131–138 (1997).
29. Datta, N. & Ruskai, M. B. Maximal output purity and capacity for asymmetric unital qudit channels. *J. Phys. A Math Gen.* **38**, 9785 (2005).
30. Nielsen, M. A. & Vidal, G. Majorization and the interconversion of bipartite states. *Quantum Information & Computation* **1**, 76–93 (2001).

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1A2B2014462) and ICT R&D program of MSIP/IITP [R0190-15-2030, Reliable crypto-system standards and core technology development for secure quantum key distribution network].

Author Contributions

J.R. contributed the idea. J.R., J.S.K. and Y.J. developed the theory. H.S. improved the manuscript and supervised the research. All the authors contributed in analyzing and discussing the results and improving the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-35777-7.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.