Encoding classical information into quantum resources

Kamil Korzekwa, Zbigniew Puchała, Marco Tomamichel, Karol Życzkowski

Abstract—We introduce and analyse the problem of encoding classical information into different resources of a quantum state. More precisely, we consider a general class of communication scenarios characterised by encoding operations that commute with a unique resource destroying map and leave free states invariant. Our motivating example is given by encoding information into coherences of a quantum system with respect to a fixed basis (with unitaries diagonal in that basis as encodings and the decoherence channel as a resource destroying map), but the generality of the framework allows us to explore applications ranging from super-dense coding to thermodynamics. For any state, we find that the number of messages that can be encoded into it using such operations in a one-shot scenario is upper-bounded in terms of the information spectrum relative entropy between the given state and its version with erased resources. Furthermore, if the resource destroying map is a twirling channel over some unitary group, we find matching one-shot lower-bounds as well. In the asymptotic setting where we encode into many copies of the resource state, our bounds yield an operational interpretation of resource monotones such as the relative entropy of coherence and its corresponding relative entropy variance.

I. INTRODUCTION

Encoding information for storage or transmission is one of the most fundamental tasks in information theory [1]. A typical communication scenario in which a d-dimensional quantum system is employed to transfer classical information between a sender S and a receiver R is composed of three stages. First, S encodes a message $m \in \{1,\ldots,M\}$ by preparing a quantum system in a state ρ_m ; then, she sends it to R via a noisy quantum channel \mathcal{N} ; finally, R decodes the original message by performing a measurement on $\mathcal{N}(\rho_m)$. Crucially, it is assumed that the encoding and decoding steps are unconstrained, so that the problem of optimal information transfer (finding the maximal M for a given average decoding error ϵ) is fully specified by the noisy channel \mathcal{N} . Holevo, Schumacher and Westmoreland [2], [3] analyzed this problem in an asymptotic setting where \mathcal{N} is memoryless and can be used multiple

Kamil Korzekwa is with International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-308 Gdańsk, Poland, and jointly with Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30-348 Kraków, Poland (email: korzekwa.kamil@gmail.com).

Zbigniew Puchała is with Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, 44-100 Gliwice, Poland, and jointly with Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30-348 Kraków, Poland.

Marco Tomamichel is with Centre for Quantum Software and Information, School of Software, University of Technology Sydney, Sydney NSW 2007, Australia.

Karol Życzkowski is with Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30-348 Kraków, Poland, and jointly with Center for Theoretical Physics, Polish Academy of Sciences, 02-668 Warszawa, Poland.

times. They established the Holevo capacity [4] of \mathcal{N} as the maximal transmission rate (in bits per channel use) that still allows for asymptotically vanishing error. More recently, various refinements of the trade-off between the decoding error and the transmission rate have been established [6]–[15] when the number of channel uses is finite.

Here, we analyze an alternative communication scenario, where we assume that S and R are connected via a noiseless channel, $\mathcal{N}=\mathcal{I}$, but the encoding ability of S is constrained. In particular, we assume that S does not have the ability to prepare arbitrary quantum states. Instead, she is given a state ρ that acts as an information carrier, and can encode a message m into it by applying an encoding \mathcal{E}_m from a constrained set of quantum channels. We focus on a family of constraints, each of which is defined via a fixed idempotent channel \mathcal{D} and the following conditions that all allowed encoding maps \mathcal{E}_m have to satisfy,

$$\mathcal{E}_m \circ \mathcal{D} = \mathcal{D},\tag{1a}$$

$$\mathcal{D} \circ \mathcal{E}_m = \mathcal{D}. \tag{1b}$$

Channels \mathcal{E}_m satisfying Eqs. (1a)-(1b) will be called *encodings* into resources destroyed by \mathcal{D} .

To understand the meaning of such constraints, note that one can interpret \mathcal{D} as a resource destroying map [16] that erases information encoded in degrees of freedom that \mathcal{E}_m affects. More precisely, application of \mathcal{D} to any state ρ renders it useless from the perspective of S, as $\mathcal{D}(\rho)$ is invariant under \mathcal{E}_m through Eq. (1a). Similarly, application of \mathcal{D} to an encoded state $\mathcal{E}_m(\rho)$ renders it useless from the perspective of R, as every $\mathcal{E}_m(\rho)$ gets sent to a fixed state $\mathcal{D}(\rho)$ through Eq. (1b). In other words, the above conditions restrict S to encodings satisfying resource non-generating and non-activating conditions [16], which additionally cannot modify free (non-resource) states. Therefore, by focusing on encodings into resources, one investigates the capability of particular degrees of freedom, corresponding to resources destroyed by \mathcal{D} , to carry classical information.

One of the particularly useful choices of \mathcal{D} is a G-twirling channel \mathcal{G} over some unitary subgroup G, i.e.,

$$\mathcal{G}(\rho) := \frac{1}{|G|} \sum_{g \in G} U^{(g)} \rho U^{(g)\dagger}, \tag{2}$$

where for continuous subgroups the above sum can be replaced by an integral $\int dg$ with respect to the Haar measure. When G

¹Here one considers the setting where a joint decoding measurement is allowed, but no entanglement between channel inputs is present, avoiding the issue of super-additivity [5].

is a full unitary group, \mathcal{G} becomes a completely depolarizing channel, i.e., $\mathcal{G}(\rho) = \mathbb{I}/d$ for all ρ . Then, Eq. (1b) is satisfied automatically and Eq. (1a) constrains the encoding to unital channels. Physically, this case corresponds to the sender S being unable to decrease the entropy of the information carrier, and so the information is encoded into the resource of purity [17]. Moreover, if G is a unitary group on a subsystem S_1 of a multipartite system $S_{1...n}$, the encodings are restricted to unital channels acting locally on S_1 . This way one can study encoding information not only into a resource of local purity, but also in entanglement, allowing one to assess the capacity of the system for super-dense coding [18].

When G is a subgroup of unitaries diagonal in a given basis $\{|k\rangle\}$ (so it is a subgroup of commuting unitaries), \mathcal{G} becomes a completely dephasing map Δ with respect to this basis, i.e., $\Delta(\rho) = \sum_{k} \langle k | \rho | k \rangle | k \rangle \langle k |$. Equations (1a)-(1b) then constrain encoding maps \mathcal{E}_m to be Schur-product channels [19], [20], i.e., $\mathcal{E}_m(\rho) = \rho \circ C_m$ with C_m being arbitrary correlation matrices ($C_m \geq 0$ and all diagonal entries being 1), and o denoting Schur (entry-wise) product (not to be confused with composition of quantum channels as in Eqs. (1a)-(1b)). Since such channels do not modify populations (diagonal elements of ρ in the given basis), but only affect coherences (off-diagonal elements), investigating communication scenario with this constraint corresponds to asking how much classical information can be encoded into the resource of quantum coherence [21], [22]. Finally, for a general group G, \mathcal{G} is a projector onto a symmetric subspace, and so \mathcal{G} -constrained encodings correspond to sending information encoded in the resource of asymmetry [23], i.e., into the degrees of freedom that are not invariant under the group action.

Another important choice of \mathcal{D} is given by a completely thermalising map, i.e., $\mathcal{T}(\rho) = \gamma$ for all ρ , with γ denoting the thermal equilibrium Gibbs state. In this case, the considered constraints limit the sender S to only encode information with Gibbs-preserving channels, i.e., satisfying $\mathcal{E}_m(\gamma) = \gamma$. Physically, this constrains S to obey the second law of thermodynamics (in the sense that the encoding channels cannot bring the information carrier farther out of equilibrium), and the information is encoded in the resource of thermodynamic non-equilibrium [24], [25].

In this paper, we derive single-shot lower- and upper-bounds for the number of classical messages that can be encoded in quantum resources and then decoded up to average probability of error ϵ . We then show how, in the limit of large number of information carriers, the two bounds coincide, yielding the optimal encoding rate up to the second order asymptotic expansion. These rates are given by the relative entropy and relative entropy variance between the state of the information carrier ρ and the same state with erased resource content $\mathcal{D}(\rho)$. We thus provide an operational meaning to a number of resource monotones used in various resource theories. In what follows, we first formally state our results in Sec. II. Then, in Sec. III, we discuss their relevance and applications in a variety of quantum communication scenarios. Next, we provide proofs of the main results in Sec. IV, which is finally followed by an outlook for future research in Sec. V.

II. STATEMENT OF THE RESULT

Our main goal is to encode a message m, chosen uniformly at random from the set $\mathcal{M}=\{1,\ldots,M\}$, using a d-dimensional quantum state ρ and a constrained set of quantum channels, described by Eqs. (1a)-(1b), so that this message can be faithfully recovered later up to average probability of error ϵ . We are thus looking for an encoder in terms of the set of quantum channels $\{\mathcal{E}_m\}_{m\in\mathcal{M}}$ that are encodings into resources destroyed by some fixed channel \mathcal{D} ; and a decoder specified by a quantum measurement described by the POVM elements $\{E_m\}_{m\in\mathcal{M}}$, such that

$$\frac{1}{M} \sum_{m=1}^{M} \operatorname{Tr}\left(\mathcal{E}_{m}(\rho) E_{m}\right) \ge 1 - \epsilon, \tag{3}$$

i.e., the average probability of incorrectly decoding the message is smaller then ϵ . We are interested in maximal allowed value of M for given ρ and ϵ . Of special importance is the case when we deal with N independent and identically distributed copies of a state ρ , i.e., when we encode into $\rho^{\otimes N}$. Then, the aim is to find the optimal rate R for encoding information into resources of quantum systems,

$$R(\rho, N, \epsilon) := \sup \left\{ \frac{\log M}{N} \mid \text{Eq. (3) holds} \right\},$$
 (4)

and, in particular, to understand its asymptotic behavior as $N \to \infty$.

In order to present our results, we first need to introduce several entropic quantities. The relative entropy D between density matrices ρ and σ is defined by [26]

$$D(\rho \| \sigma) := \operatorname{tr} \left(\rho(\log \rho - \log \sigma) \right), \tag{5}$$

while the relative entropy variance V is given by [27], [28]

$$V(\rho \| \sigma) := \operatorname{tr} \left(\rho (\log \rho - \log \sigma)^2 \right) - D(\rho \| \sigma)^2. \tag{6}$$

The collision relative entropy D_2 is defined as follows [29]

$$D_2(\rho \| \sigma) := \log \operatorname{Tr} \left(\sigma^{-1/4} \rho \sigma^{-1/4} \right), \tag{7}$$

and the information spectrum relative entropy D_s^δ is given by [30]

$$D_s^{\delta}(\rho \| \sigma) := \sup \left\{ K \mid \operatorname{Tr} \left(\rho \Pi_{\rho < 2^K \sigma} \right) \le \delta \right\}, \tag{8}$$

where $\Pi_{\rho \leq 2^K \sigma}$ is the orthogonal projection onto the union of eigenspaces of $2^K \sigma - \rho$ with non-negative eigenvalues. Finally, the hypothesis testing relative entropy D_H^{ϵ} is defined by [9]

$$D_{H}^{\epsilon}(\rho \| \sigma) := -\log \inf \left\{ \operatorname{Tr} \left(Q \sigma \right) \mid 0 \leq Q \leq \mathbb{1}, \right.$$
$$\operatorname{Tr} \left(Q \rho \right) \geq 1 - \epsilon \right\}. \quad (9)$$

Now, we have the following upper-bound for the number of messages M that can be encoded in resources erased by arbitrary resource destroying map \mathcal{D} .

Lemma 1 (Single-shot upper-bound). The number of messages $M(\rho, \epsilon)$ that can be encoded into resources of a quantum state ρ destroyed by \mathcal{D} , for an average decoding error at most ϵ , is upper-bounded by

$$M(\rho, \epsilon) \le e^{D_H^{\epsilon}(\rho \parallel \mathcal{D}(\rho))}.$$
 (10)

For a resource destroying map given by a G-twirling channel \mathcal{G} , the following theorem bounds M in the single-shot setting from both sides.²

Theorem 2 (Single-shot encoding). The number of messages $M(\rho, \epsilon)$ that can be encoded into resources of a quantum state ρ destroyed by a G-twirling channel G over a unitary subgroup G, for an average decoding error at most ϵ , is bounded by

$$\delta e^{D_s^{\epsilon-\delta}(\rho||\mathcal{G}(\rho))} \le M(\rho, \epsilon) \le \frac{1}{\delta} e^{D_s^{\epsilon+\delta}(\rho||\mathcal{G}(\rho))}, \tag{11}$$

for all $\delta \in (0, \min\{\epsilon, 1 - \epsilon\})$.

Finally, the asymptotic encoding rate into quantum resources destroyed by \mathcal{G} is captured by the following theorem:

Theorem 3. The optimal rate $R(\rho, N, \epsilon)$ for encoding information into resources of a quantum state $\rho^{\otimes N}$ destroyed by a G-twirling channel $\mathcal{G}^{\otimes N}$ over a unitary subgroup $G^{\times N}$, for an average decoding error at most ϵ , is governed by the following second-order asymptotic expansion

$$R = D(\rho \| \mathcal{G}(\rho)) + \frac{\Phi^{-1}(\epsilon)}{\sqrt{N}} \sqrt{V(\rho \| \mathcal{G}(\rho))} + O(\log N), \quad (12)$$

with Φ^{-1} denoting the inverse function of the normal Gaussian cumulative distribution function Φ .

In what follows, we first discuss and interpret the above results in Sec. III, and then in Sec. IV we present their proofs.

III. DISCUSSION

From now on we will drop the explicit dependence of the encoding rate $R(\rho, N, \epsilon)$ on the number of copies N and the acceptable error level ϵ , and will concisely write $R(\rho)$. We will also use the asymptotic notation, with \simeq , \lesssim and \gtrsim denoting equalities and inequalities up to terms of the order $O(\log N)$.

A. Encoding power of unitary subgroups

First, let us note that for a resource destroying map given by a G-twirling channel \mathcal{G} , unitaries $U^{(g)}$ from Eq. (2) satisfy Eqs. (1a)-(1b) due to the rearrangement lemma, i.e., they form encodings into resources destroyed by \mathcal{G} . In fact, as we will show in the proof of Theorem 3 in Sec. IV-B, the optimal rate can be achieved using encodings channels \mathcal{E}_m given precisely by unitaries $U^{(g)}$. Thus, the optimal encoding rate into resources destroyed by \mathcal{G} coincides with the maximal number of messages that the sender can encode while being constrained to only using a subgroup G of all unitary transformations. This way Theorem 3 yields the encoding power of unitary subgroups for a given state.

For G being the full unitary group U, we obtain

$$R_U(\rho) \simeq R_U^{\infty}(\rho) + \frac{\Phi^{-1}(\epsilon)}{\sqrt{N}} \sqrt{V(\rho)},$$
 (13a)

$$R_U^{\infty}(\rho) := \log d - S(\rho), \tag{13b}$$

where $S(\rho) := -\mathrm{Tr}\,(\rho\log\rho)$ is the von Neumann entropy and $V(\rho) := \mathrm{Tr}\,\big(\rho(S(\rho) + \log\rho)^2\big)$ is the entropy variance. As already mentioned in Sec. I, the above not only specifies the encoding power of unitary transformations acting on $\rho^{\otimes N}$, but also the amount of information that can be encoded in the resource of purity of the state $\rho^{\otimes N}$.

Now, if G is a proper subgroup of the full unitary group, we get

$$R_{\rm G}(\rho) \simeq R_{\rm G}^{\infty}(\rho) + \frac{\Phi^{-1}(\epsilon)}{\sqrt{N}} \sqrt{V(\rho \| \mathcal{G}(\rho))},$$
 (14a)

$$R_{\mathcal{G}}^{\infty}(\rho) := D(\rho || \mathcal{G}(\rho)) = S(\mathcal{G}(\rho)) - S(\rho). \tag{14b}$$

A more explicit expression for $V(\rho || \mathcal{G}(\rho))$ can be obtained by using the covariance.

$$cov_{\rho}(A, B) = Tr(\rho AB) - Tr(\rho A) Tr(\rho B).$$
 (15)

Then we have

$$V(\rho || \mathcal{G}(\rho)) = V(\rho) + V(\mathcal{G}(\rho)) + 2\operatorname{cov}_{\rho}(\log \rho, \log \mathcal{G}(\rho)).$$
 (16)

We thus obtain the following additive splitting of the asymptotic encoding rates

$$R_U^{\infty}(\rho) = R_G^{\infty}(\rho) + R_U^{\infty}(\mathcal{G}(\rho)), \tag{17}$$

which can be interpreted as follows. In the asymptotic limit, $N \to \infty$, the number of messages that can be encoded into $\rho^{\otimes N}$ per one copy of ρ using all unitary transformations splits additively into two terms. The first one corresponds to the optimal encoding rate when one is constrained to a subgroup G of all unitary encodings. The second term tells us about the maximal number of messages that can be encoded using all unitaries, but on a state with resources erased by G. In short: the full unitary encoding power for ρ is just a sum of the encoding power of G for ρ and the full unitary encoding power for $G(\rho)$.

Finally, let us note that the above considerations are closely related to the problem of distinguishing between unitary channels $U^{(g)}(\cdot)U^{(g)\dagger}$ using an input state ρ , i.e., distinguishing between states $\sigma^{(g)}:=U^{(g)}\rho U^{(g)\dagger}$. In this problem, the number of messages to be encoded is fixed (and equal to the order of the group), and one wants to maximise the success probability of correctly guessing $\sigma^{(g)}$. The authors of Ref. [31] showed that this success probability is directly related to a resource measure known as robustness of asymmetry.

B. Encoding information in quantum coherence

For a given distinguished orthonormal basis $\{|k\rangle\}_{k=1}^d$, the diagonal elements of ρ , $\langle k|\,\rho\,|k\rangle$, are known as populations, while the off-diagonal elements, $\langle k|\,\rho\,|l\rangle$ with $k\neq l$, are called coherences. The completely dephasing quantum channel Δ with respect to this basis sends all coherences to zero whilst not affecting the populations. More generally, we say that a quantum channel $\mathcal E$ is *population-preserving* if for all k,l we have

$$\langle k|\mathcal{E}(|l\rangle\langle l|)|k\rangle = \delta_{kl},$$
 (18)

with δ_{kl} denoting the Kronecker delta. In other words, such channels process only coherences of a quantum system.

²Lemma 1 gives an alternative, slightly tighter, upper-bound; however, we opted here for a more symmetrical exposition.

As already noted in Sec. I, channels \mathcal{E}_m that are encodings into resources destroyed by Δ are precisely the population-preserving channels. Using Theorems 2 and 3, we can thus study the capacity of coherence to carry information. This way we can provide operational meaning to measures of coherence studied within the resource theory of coherence [22]. In particular, we have that the asymptotic rate of encoding classical information into coherences of $\rho^{\otimes N}$ is given by

$$R_{\Delta}(\rho) \simeq R_{\Delta}^{\infty}(\rho) + \frac{\Phi^{-1}(\epsilon)}{\sqrt{N}} \sqrt{V(\rho \| \Delta(\rho))},$$
 (19a)

$$R_{\Delta}^{\infty}(\rho) := D(\rho \| \Delta(\rho)) = S(\Delta(\rho)) - S(\rho). \tag{19b}$$

Here, $D(\rho\|\Delta(\rho))$ is the well-known relative entropy of coherence that quantifies distillable coherence (and coherence cost) in the asymptotic limit under incoherent operations [32]; while $V(\rho\|\Delta(\rho))$ is the relative entropy variance of coherence which, to the authors' best knowledge, is first introduced here. From the above we see that the states that are asymptotically optimal for encoding information into coherences are pure states (so that $S(\rho)=0$) that get dephased to a maximally mixed state (so that $S(\Delta(\rho))=\log d$ is maximal). The general form of such a state is given by a uniform superposition of all states from the distinguished basis.

Using the additive splitting from Eq. (17), we also have that the amount of information that can be unitarily encoded in states $\rho^{\otimes N}$ is equal to the amount of information that can be encoded in coherences of $\rho^{\otimes N}$ plus the amount of information that can be encoded in a decohered state $\Delta(\rho)^{\otimes N}$. Note, that this splitting is directly related to decomposing uncertainty into classical and quantum parts [33].

Moreover, we want to point out that the problem of encoding information into coherences was studied before in the singleshot and error-free scenario. First, in Ref. [34], the concept of coherifying quantum states was introduced: a state ρ is a coherification of a diagonal state ρ_0 if it dephases to it, i.e., $\mathcal{D}(\rho) = \rho_0$. Then, in Ref. [35], the authors were investigating the number of coherifications of ρ_0 with nonoverlapping support or, in other words, the number of perfectly distinguishable states that are classically indistinguishable (as they are send to the same state by a dephasing map Δ). The number of such perfectly distinguishable states was also related to time-energy uncertainty relation. More precisely, and specialising to the asymptotic scenario captured by Theorem 3, consider N copies of a quantum system, each described by a Hamiltonian $H = \sum_k E_k |E_k\rangle\langle E_k|$ and prepared in a pure state $|\psi\rangle$. Then, the ability of $|\psi\rangle^{\otimes N}$ to act as a clock can be measured by the number T of distinguishable states it passes through during a free evolution generated by the total Hamiltonian. From Eqs. (19a)-(19b), we see that asymptotically this number is upper-bounded as $T \leq 2^{h(p)N}$, where $h(\boldsymbol{p}) = -\sum_k p_k \log p_k$ is the Shannon entropy of the energy distribution $p_k := |\langle E_k | \psi \rangle|^2$. Therefore, the better resolution of the clock we want to get, the higher entropy of energy distribution is required, and so the inequality

$$\frac{1}{\log T} \cdot h(\boldsymbol{p}^{\otimes N}) \ge 1,\tag{20}$$

can be interpreted as time-energy uncertainty relation. Of course, one could also use the second order term from Eq. (19a) to get an even tighter result.

C. Asymptotic super-dense coding

Consider now encoding information into a state ρ_{AB} of a bipartite system AB (with local dimensions d_A and d_B), using encodings into resources destroyed by G-twirling channel $\mathcal G$ over all unitaries on A. Such encodings, according to Eqs. (1a)-(1b), correspond to local unital channels on system A. Using Theorem 3, we can then find the number of approximately orthogonal states that the global state of AB can be steered to by operating only locally on A. We have the encoding rate given by

$$R_{\rm loc}(\rho_{AB}) \simeq R_{\rm loc}^{\infty}(\rho_{AB}) + \frac{\Phi^{-1}(\epsilon)}{\sqrt{N}} \sqrt{V(A|B)},$$
 (21a)

$$R_{loc}^{\infty}(\rho_{AB}) := \log d_A - S(A|B), \tag{21b}$$

with $S(A|B) = D(\rho_{AB} || \mathbb{1}_A \otimes \rho_B) = S(\rho_{AB}) - S(\rho_B)$ being the conditional entropy, and $V(A|B) = V(\rho_{AB} || \mathbb{1}_A \otimes \rho_B)$ being the corresponding variance [27]. We see that the states that are asymptotically optimal for super-dense coding are pure states with a maximally mixed marginal, i.e., maximally entangled states.

In the simplest case of a two-qubit system we can recover the asymptotic rate for super-dense coding [18]: if ρ_{AB} is a pure product state, we can encode only a single bit per a copy of ρ_{AB} by operating on A; but if ρ_{AB} is one of the four Bell states, we can encode 2 bits per copy, since local operations can map between all Bell states. More generally, the above optimal encoding rate recovers the known asymptotic result obtained first for qubit systems in Ref. [36], and then generalised to qudits in Ref. [37], but also yields the second order asymptotic correction term. Also, using again the splitting from Eq. (17), we find that the amount of information that can be encoded via global unitaries in N copies of a state ρ_{AB} is equal to the amount of information that can be encoded unitaries on A plus the amount of information that can be encoded unitarily in ρ_{B} :

$$R_U^{\infty}(\rho_{AB}) = R_{loc}^{\infty}(\rho_{AB}) + R_U^{\infty}(\rho_B). \tag{22}$$

D. Collective encoding and encoding via permutations

Let us now switch to an n-partite system with equal local dimensions d, prepared in a state $\rho_{1...n}$. We want to discuss encodings into resources destroyed by two types of resource destroying maps, $\mathcal{G}_{\rm col}$ and $\mathcal{G}_{\rm per}$, acting on each n-partite system via

$$\mathcal{G}_{\text{col}}(\cdot) = \int_{U(d)} dg \ U(g)^{\otimes n}(\cdot) U^{\dagger}(g)^{\otimes n}, \qquad (23a)$$

$$\mathcal{G}_{\text{per}}(\cdot) = \frac{1}{n!} \sum_{\pi_i \in S_n} \pi_i(\cdot) \pi_i^{\dagger}. \tag{23b}$$

Here, the integral in the first equation goes over all d-dimensional unitaries U(g) (according to the Haar measure), and the sum in the second equation is over all permutations π_i

between n subsystems. In the case of $\mathcal{G}_{\mathrm{col}}$, Theorems 2 and 3 allow us to study the encoding power of collective unitaries, i.e., the amount of information that can be encoded into $\rho_{1...n}$ (or N copies of it) when the allowed operations are given by the same unitary on each subsystem. This can be interpreted as encoding information in the global degrees of freedom, as such unitaries cannot affect relative degrees of freedom between different subsystems. And, in the case of $\mathcal{G}_{\mathrm{per}}$, we can investigate how much information can be encoded by just permuting subsystems between n parties.

In order to find the optimal number of messages that can be encoded, we need to understand how the twirling operations \mathcal{G}_{col} and \mathcal{G}_{per} act on a general state $\rho_{1...n}$. In the simplest case of a bipartite system, n=2, their action takes a particularly simple form:

$$G_{\text{col}}(\rho_{12}) = p_s \frac{\Pi_s}{d_s} + (1 - p_s) \frac{1 - \Pi_s}{d^2 - d_s},$$
 (24a)

$$\mathcal{G}_{per}(\rho_{12}) = \Pi_s \rho_{12} \Pi_s + (\mathbb{1} - \Pi_s) \rho_{12} (\mathbb{1} - \Pi_s),$$
 (24b)

where $p_s = \text{Tr}(\rho_{12}\Pi_s)$, $d_s = \text{Tr}(\Pi_s) = d(d+1)/2$ and Π_s is the projector onto the symmetric subspace, i.e.,

$$\Pi_s = \sum_{k=1}^d |kk\rangle\langle kk| + \sum_{k=1}^d \sum_{l=k+1}^d \frac{(|kl\rangle + |lk\rangle)(\langle kl| + \langle lk|)}{2}.$$
 (25)

We can now ask, which states ρ_{12} are asymptotically optimal for encoding information using collective unitaries and permutations. In the first case these are given by pure states

$$|\psi_{12}^*\rangle = \sqrt{\frac{d+1}{2d}}|\psi_s\rangle + \sqrt{\frac{d-1}{2d}}|\psi_a\rangle,\tag{26}$$

where $|\psi_s\rangle$ and $|\psi_a\rangle$ are arbitrary pure states living in the symmetric and antisymmetric subspaces, respectively. Such states are optimal, as they are pure and twirl to a maximally mixed state. In case of encoding information using permutations, the situation is even simpler, as there are only two allowed channels (identity and transposition between the two systems). Thus, the maximal number of messages we can encode per one copy of the system is upper-bounded by 2. This bound can be attained, even in single-shot scenario, by simply choosing a state $|01\rangle$.

Beyond the above bipartite example, one could further investigate multipartite scenarios for n > 2. In order to find the action of $\mathcal{G}_{\mathrm{col}}$ and $\mathcal{G}_{\mathrm{per}}$ (and so to derive the optimal encoding rate), one could employ the fact that these two resource destroying maps are closely related via a Schur-Weyl duality [23]. That is, the permutation group and the group of collective unitaries commute, and so the tensor space decomposes into a direct sum of tensor products of irreducible modules for these two groups. In particular, for $\mathcal{G}_{\mathrm{col}}$ we could ask whether there always exists a quantum state $\rho_{1...n}$ that under collective unitaries can asymptotically encode the ultimate maximal number of bits per system (which is specified by the dimensionality of $\rho_{1...n}$, i.e., $\log d^n$). Thus, the problem is to find a pure state $|\psi\rangle$ that is mapped to a maximally mixed state by \mathcal{G}_{col} . In Appendix A we briefly discuss how to approach this problem and provide an example for a tripartite system. In the case of \mathcal{G}_{per} we could instead ask,

whether it is always possible to find a state $\rho_{1...n}$ that maps under permutations to a mutually orthogonal set of states, thus encoding the maximal number of messages $\min\{n!,d^n\}$. For example, it is to easy to see that when $d \geq n$, one can simply choose a state $|\psi_\pi\rangle = |1,2,\ldots,d\rangle$, which is mapped by permutations to n! orthogonal states, and so it is optimal for encoding messages with permutation group. On the other hand, when d < n, a plausible ansatz for the optimal state is given by $|\psi_\pi\rangle^{\otimes \lfloor \frac{n}{d} \rfloor} \otimes |\phi_\pi\rangle$ with $|\phi_\pi\rangle = |1,2,\ldots,n-d\lfloor \frac{n}{d} \rfloor\rangle$.

E. Shared reference frames and private communication

We now want to briefly discuss the relation between encoding into resources destroyed by \mathcal{G} and a private classical communication scheme for a decohering superoperator \mathcal{G} [38], as introduced in the studies on quantum reference frames [39]. In this scenario there are three parties: beyond the sender Sand the receiver R, there is also an eavesdropper E. The two communicating parties share a private reference frame for some degree of freedom, i.e., both S and R agree on the form of group representation U(g) given a classical label g describing it. As an example, consider a shared Cartesian frame of reference (given, e.g., by three mutually orthogonal rigid rods defining directions x, y, z). Then, the classical description of an element of the rotation group can be given by three Euler angles with respect to the axes defined by the shared reference frame. Thus, if S tells R that she prepared a spin along the positive z direction, and he wants to rotate it, so that it points in the opposite direction, he knows precisely which U(g) to perform. However, E that does not have access to the reference frame, and thus she does not know what "along positive z direction" means. Therefore, her description of the system is given by a uniform superposition (stemming from no knowledge about the orientation of the reference frame) over all possible rotations of the reference frame. This way a state ρ is described by E as $\mathcal{G}(\rho)$, where \mathcal{G} here is the twirling over SO(3) group, but in general can be given by twirling over arbitrary group G corresponding to a shared reference frame between S and R.

Now, following the definition given in Ref. [38], S and Rhave a private classical communication scheme employing a shared reference frame related to a group G, if S can prepare M orthogonal states σ_m , such that $\mathcal{G}(\sigma_m) = \rho_0$ for all m and some fixed state ρ_0 . This means that S can send one of M perfectly distinguishable messages to R, while at the same time for the eavesdropper E all these messages will be completely indistinguishable, and so the communication will be secure. Through Eq. (1b), it is then clear that if for a state ρ there exists M encodings into resources destroyed by \mathcal{G} , then S and R have a private classical communication scheme: S prepares one of the states $\mathcal{E}_m(\rho)$ that are (almost) perfectly distinguishable by R, but for E they are all described by $\mathcal{G}(\mathcal{E}_m(\rho)) = \mathcal{G}(\rho)$. Note however, that for the asymptotic result to hold for private communication using $\rho^{\otimes N}$, S and R require many copies of uncorrelated reference frames – otherwise E can learn the orientation of the reference frame from the first few copies of ρ . Alternatively, one could use the setup of \mathcal{G}_{col} : namely, instead of performing the asymptotic analysis as in Theorem 3, one could use the one-shot bounds and increase the number of parties that \mathcal{G}_{col} twirls over (i.e. we go $n \to \infty$ keeping N = 1).

F. Thermodynamics

Finally, we want to make a short comment on the case of a resource destroying map given by a completely thermalising map \mathcal{T} . Since it is not a G-twirling channel, we cannot use Theorems 2 or 3, but Lemma 1 still applies in this case. Thus, in the asymptotic limit, the number M of (almost) orthogonal states that can be obtained via Gibbs-preserving operations (that model free thermodynamic transformations with no external access to any sources of work or sinks of entropy) from a state $\rho^{\otimes N}$ is upper-bounded as

$$\log M \lesssim ND(\rho \| \gamma) + \sqrt{NV(\rho \| \gamma)} \Phi^{-1}(\epsilon). \tag{27}$$

Now, the crucial thing is that the entropic quantities appearing on the right hand side of the above inequality have a clear thermodynamic interpretation. Denoting by β the inverse temperature, we have that $D(\rho\|\gamma)/\beta$ is the free energy of the state ρ [24]; while $V(\rho\|\gamma)$ was recently shown to be related to a generalised heat capacity of ρ [40]. Thus, the number of messages that can be thermodynamically encoded for free in a state ρ is bounded by the amount of thermodynamic work that one can perform while thermalising the state ρ . Or, in other words, the volume of the future thermal cone of ρ , i.e., the number of distinguishable states it can evolve to during the evolution generated by interaction with the heat bath, is upper-bounded by the amount of work one can extract from ρ .

IV. DERIVATION OF THE ENCODING BOUNDS

A. Optimality (Proof of Lemma 1)

First, we note that the final state after the encoding, correlated with the chosen message, can be written as

$$\tau_{MQ} = \frac{1}{M} \sum_{m \in \mathcal{M}} |m\rangle\langle m| \otimes \mathcal{E}_m(\rho). \tag{28}$$

We now assume that there exists a suitable decoder $\{E_m\}_{m\in\mathcal{M}}$ for a faithful recovery up to average failure probability ϵ (i.e., Eq. (3) is fulfilled), and will show that this leads to an upper-bound on M given by Lemma 1. To prove this, let us take a closer look at the hypothesis testing relative entropy D_H^ϵ between τ and

$$\zeta := \frac{1}{M} \sum_{m \in \mathcal{M}} |m\rangle\langle m| \otimes \mathcal{D}(\rho). \tag{29}$$

By recalling the definition of D_H^{ϵ} , Eq. (9), we see that

$$D_H^{\epsilon}(\tau_{MQ} \| \zeta) \ge -\log \operatorname{Tr}(Q\zeta) \tag{30}$$

for

$$Q = \sum_{m \in M} |m\rangle\langle m| \otimes E_m. \tag{31}$$

This is because the above (potentially suboptimal) choice of Q clearly satisfies $0 \le Q \le \mathbb{1}$ and also

$$\operatorname{Tr}\left(Q\tau_{MQ}\right) = \frac{1}{M} \sum_{m \in \mathcal{M}} \operatorname{Tr}\left(\mathcal{E}_{m}(\rho)E_{m}\right) \ge 1 - \epsilon, \quad (32)$$

with the final inequality holding, because we assumed that Eq. (3) is fulfilled. At the same time we have

$$\operatorname{Tr}(Q\zeta) = \frac{1}{M} \sum_{m \in \mathcal{M}} \operatorname{Tr}(\mathcal{D}(\rho)E_m) = \frac{1}{M}, \quad (33)$$

so that

$$\log M \le D_H^{\epsilon}(\tau_{MQ} \| \zeta). \tag{34}$$

Next, we employ the data-processing inequality twice. First, for the channel

$$\tilde{\mathcal{E}} := \sum_{m \in \mathcal{M}} |m\rangle\langle m| \otimes \mathcal{E}_m \tag{35}$$

that leaves ζ unchanged, and then for tensoring with the independent message register. This yields the following sequence of inequalities:

$$D_{H}^{\epsilon}(\tau_{MQ} \| \zeta) = D_{H}^{\epsilon} \left(\tilde{\mathcal{E}} \left(\frac{1}{M} \sum_{m \in \mathcal{M}} |m\rangle \langle m| \otimes \rho \right) \| \tilde{\mathcal{E}}(\zeta) \right)$$

$$\leq D_{H}^{\epsilon} \left(\frac{1}{M} \sum_{m \in \mathcal{M}} |m\rangle \langle m| \otimes \rho \| \zeta \right)$$

$$\leq D_{H}^{\epsilon} (\rho \| \mathcal{D}(\rho)). \tag{36}$$

Combining this with Eq. (34), we finally arrive at

$$\log M \le D_H^{\epsilon}(\rho \| \mathcal{D}(\rho)). \tag{37}$$

B. Achieveability (Proof of Theorem 2)

The upper-bound can be proven by employing Lemma 1. We simply note that according to [27, Lemma 12], we have

$$D_H^{\epsilon}(\rho \| \mathcal{D}(\rho)) \le D_s^{\epsilon + \delta}(\rho \| \mathcal{D}(\rho)) + \log \frac{1}{\delta}, \tag{38}$$

for $\delta \in (0, 1-\epsilon)$, which yields the upper-bound in Theorem 2.

In order to prove that the lower-bound from Eq. (11) holds, we will closely follow the approach of Ref. [41]. We will consider a whole family of encoder-decoder pairs, each defined by its codebook C. We will then show that by choosing the encoder-decoder pair from this family uniformly at random, and encoding the number of messages given by the claimed lower-bound, the expected probability of error is below the required threshold. This, in turn, implies that within the introduced family there must exist at least one encoder-decoder pair that allows for encoding that number of messages and decoding with error at most ϵ .

A given codebook $\mathcal C$ is defined by a mapping from the set of messages $\mathcal M$ to the set of integers $\{1,\ldots,|G|\}$. The encoding is then defined as follows. Every message m is first classically encoded into an $g_m \in \{0,\ldots,|G|\}$ (corresponding to the choice of the codebook $\mathcal C$), which is used to encode the message into a quantum system

$$\sigma^{(g_m)} = U^{(g_m)} \rho U^{(g_m)\dagger}, \tag{39}$$

where $U^{(g_m)}$ is a unitary appearing in the definition of a given G-twirling channel in Eq. (2). Note that such encodings satisfy conditions from Eq. (1a)-(1b) due to the rearrangement lemma,

$$\forall g: \quad \mathcal{G}(U^{(g)}(\cdot)U^{(g)\dagger}) = U^{(g)}\mathcal{G}(\cdot)U^{(g)\dagger} = \mathcal{G}(\cdot). \tag{40}$$

The decoder, on the other hand, is given by a pretty good measurement consisting of M POVM elements E_m defined by

$$E_m := S\sigma^{(g_m)}S,\tag{41}$$

with

$$S = \left(\sum_{m=1}^{M} \sigma^{(g_m)}\right)^{-1/2}.$$
 (42)

We now define the joint message-classical encodingquantum encoding system

$$\tau_{MCQ} = \frac{1}{M} \sum_{m=1}^{M} |m\rangle\langle m| \otimes |g_m\rangle\langle g_m| \otimes \sigma^{(g_m)}, \qquad (43)$$

and all its marginals through partial traces, e.g., $\tau_Q = \operatorname{Tr}_{MC}(\tau_{MCQ})$. The probability $p_s(\mathcal{C})$ of successfully decoding the message encoded with the use of a codebook \mathcal{C} is then given by

$$p_s(\mathcal{C}) = \frac{1}{M} \sum_{m=1}^{M} \operatorname{Tr} \left(\sigma^{(g_m)} E_m \right)$$

$$= \frac{1}{M} \operatorname{Tr} \left(\sum_{m=1}^{M} \left(S^{1/2} \sigma^{(g_m)} S^{1/2} \right)^2 \right)$$

$$= \frac{1}{M} \exp D_2(\tau_{MCQ} || \tau_{MC} \otimes \tau_Q), \tag{44}$$

where the final equality is the crucial observation made in Ref. [41], which relates success probability for pretty good measurement scheme to the collision relative entropy defined in Eq. (7).

We now consider a random choice of the codebook \mathcal{C} : each message m is independently encoded into an integer g_m uniformly at random, i.e., all g_m are independent and identically distributed (according to a uniform distribution) random variables. The success probability P_s averaged over all choices of codebooks is then given by

$$P_{S} := \mathbb{E}_{\mathcal{C}} p_{s}(\mathcal{C}) = \frac{1}{M} \mathbb{E}_{\mathcal{C}} \exp D_{2}(\tau_{MCQ} \| \tau_{MC} \otimes \tau_{Q})$$

$$\geq \frac{1}{M} \mathbb{E}_{\mathcal{C}} \exp D_{2}(\tau_{CQ} \| \tau_{C} \otimes \tau_{Q})$$

$$\geq \frac{1}{M} \exp D_{2}(\mathbb{E}_{\mathcal{C}} \tau_{CQ} \| \mathbb{E}_{\mathcal{C}} \tau_{C} \otimes \tau_{Q}), \quad (45)$$

with the first inequality coming from the data processing inequality and the second one from the joint convexity of D_2 . Let us then evaluate $\mathbb{E}_{\mathcal{C}}\tau_{CQ}$ and $\mathbb{E}_{\mathcal{C}}\tau_{C}\otimes\tau_{Q}$. We have

$$\mathbb{E}_{\mathcal{C}}\tau_{CQ} = \frac{1}{M}\mathbb{E}_{\mathcal{C}} \sum_{m=1}^{M} |g_{m}\rangle\langle g_{m}| \otimes \sigma^{(g_{m})}$$

$$= \mathbb{E}_{\mathcal{C}} |g_{1}\rangle\langle g_{1}| \otimes \sigma^{(g_{1})}$$

$$= \frac{1}{|G|} \sum_{g=1}^{|G|} |g\rangle\langle g| \otimes \sigma^{(g)}$$

$$= \mathcal{W}\left(\frac{1}{|G|} \otimes \rho\right), \tag{46}$$

where we introduced a unitary channel \mathcal{W} specified by unitary matrix

$$W = \sum_{g=1}^{|G|} |g\rangle\langle g| \otimes U^{(g)}. \tag{47}$$

Similarly, we can evaluate the following,

$$\mathbb{E}_{\mathcal{C}}\tau_{C} \otimes \tau_{Q} = \frac{1}{M^{2}} \mathbb{E}_{\mathcal{C}} \sum_{m,n=1}^{M} |g_{m}\rangle\langle g_{m}| \otimes \sigma^{(g_{n})} \\
= \frac{1}{M^{2}} \mathbb{E}_{\mathcal{C}} \sum_{m=1}^{M} |g_{m}\rangle\langle g_{m}| \otimes \sigma^{(g_{m})} \\
+ \frac{1}{M^{2}} \mathbb{E}_{\mathcal{C}} \sum_{m\neq n=1}^{M} |g_{m}\rangle\langle g_{m}| \otimes \sigma^{(g_{n})} \\
= \frac{1}{M} \mathcal{W} \left(\frac{1}{|G|} \otimes \rho \right) \\
+ \frac{M-1}{M} \mathbb{E}_{\mathcal{C}} |g_{1}\rangle\langle g_{1}| \otimes \mathbb{E}_{\mathcal{C}}\sigma^{(g_{1})} \\
= \frac{1}{M} \mathcal{W} \left(\frac{1}{|G|} \otimes \rho \right) \\
+ \frac{M-1}{M} \left(\frac{1}{|G|} \otimes \mathcal{G} (\rho) \right) \\
= \mathcal{W} \left(\frac{1}{|G|} \otimes \left(\frac{1}{M} \rho + \frac{M-1}{M} \mathcal{G}(\rho) \right) \right), \quad (48)$$

where we have used the fact that uniformly random application of $U^{(g)}$ acts as a G-twirling channel \mathcal{G} , and that a G-twirled state is invariant under unitaries defining \mathcal{G} .

Employing the unitary invariance of D_2 we thus have

$$P_{s} \geq \frac{1}{M} \exp D_{2} \left(\frac{1}{|G|} \otimes \rho \right)$$

$$\frac{1}{|G|} \otimes \left(\frac{1}{M} \rho + \frac{M-1}{M} \mathcal{G}(\rho) \right)$$

$$= \frac{1}{M} \exp D_{2} \left(\rho \left\| \frac{1}{M} \rho + \frac{M-1}{M} \mathcal{G}(\rho) \right) \right).$$
(49)

Now, we use Theorem 3 of Ref. [41], that allows us to lower-bound the above expression by replacing D_2 with the information spectrum relative entropy D_s^{δ} , defined in Eq. (8). More precisely, for all $0 < \delta < 1$ it holds that

$$\exp D_{2}\left(\rho \left\| \frac{1}{M}\rho + \frac{M-1}{M}\mathcal{G}(\rho) \right) \right.$$

$$\geq (1-\delta)\left(\frac{1}{M} + \frac{M-1}{M}\exp\left[-D_{s}^{\delta}\left(\rho \| \mathcal{G}(\rho)\right)\right]\right)^{-1}$$

$$\geq M(1-\delta)\left(1 - (M-1)\exp\left[-D_{s}^{\delta}\left(\rho \| \mathcal{G}(\rho)\right)\right]\right)$$

$$\geq M(1-\delta)(1 - M\exp\left[-D_{s}^{\delta}\left(\rho \| \mathcal{G}(\rho)\right)\right]\right). \tag{50}$$

We can use the above to bound P_s as follows,

$$P_s \ge (1 - \delta)(1 - M \exp\left[-D_s^{\delta}(\rho \| \mathcal{G}(\rho))\right]), \quad (51)$$

which yields

$$M \ge \frac{\epsilon - \delta}{1 - \delta} \exp D_s^{\delta} \left(\rho \| \mathcal{G}(\rho) \right), \tag{52}$$

with $\epsilon := 1 - P_s$ being the failure probability. Note that, although the above inequality holds for all $0 < \delta < 1$, it is non-trivial only when $\delta < \epsilon$.

The simplified bound displayed in Theorem 2 results after bounding the denominator by 1 and the substitution $\delta \to \epsilon - \delta$.

C. Asymptotics (Proof of Theorem 3)

In order to prove Theorem 3, we simply need to use Theorem 2 and the second order asymptotic expansions of the information spectrum relative entropy. First, it was established in [27] (see also [42]) that for every two density matrices ρ, σ , fixed $0 < \epsilon < 1$ and $\delta = O(1/\sqrt{N})$ we have

$$D_s^{\epsilon \pm \delta} \left(\rho^{\otimes N} \| \sigma^{\otimes N} \right) \simeq ND(\rho \| \sigma) + \sqrt{NV(\rho \| \sigma)} \Phi^{-1}(\epsilon). \tag{53}$$

Recall that \simeq denotes equality up to terms of the order $O(\log N)$. Now, we employ the definition of the encoding rate, Eq. (4). Substituting $\delta = 1/\sqrt{N}$ into the bounds in Theorem 2 and applying the above expansion, we obtain

$$R(\rho, N, \epsilon) \simeq D(\rho \| \mathcal{G}(\rho)) + \frac{\Phi^{-1}(\epsilon)}{\sqrt{N}} \sqrt{V(\rho \| \mathcal{G}(\rho))}.$$
 (54)

V. OUTLOOK

In this paper we have studied the problem of single-shot and asymptotic encoding of classical information in resources of a quantum state ρ destroyed by a decohering quantum channel \mathcal{D} , i.e., in the degrees of freedom that completely decohere under the action of \mathcal{D} . We focused on a particular family of resource destroying maps given by G-twirling operators \mathcal{G} over arbitrary unitary subgroups. In Theorem 2 we found lower- and upper-bounds for the number of messages that can be encoded in resources destroyed by \mathcal{G} with an error probability ϵ ; while in Theorem 3 we found the second order asymptotic expansion for the encoding rate. We then discussed applications of our results to a number of problems in quantum information theory, including quantifying informational capacity of quantum coherence, usefulness of entangled states for super-dense coding and encoding power of unitary subgroups.

We see three clear paths for future research stemming from our results. First, one could look for other unitary subgroups with operational relevance, and thus find second-order asymptotic encoding rates for constrained communication scenarios. Second, we expect that not only upper-bound holds for general resource destroying maps \mathcal{D} (Lemma 1), but also that there should be a lower-bound that asymptotically coincides with the upper one. In particular, it would be interesting to prove the existence of such bounds for \mathcal{D} being given by a completely thermalising map \mathcal{T} . This way one would relate the encoding power of Gibbs-preserving operations in a state ρ with the amount of work that can be extracted from ρ with these operations, thus providing one more strong link between information theory and thermodynamics. Finally, as the optimal encoding rate in resources destroyed by \mathcal{G} is given by $S(\mathcal{G}(\rho)) - S(\rho)$, one could look for states that maximise this quantity for general groups G. More broadly, one could also investigate the generalisation of the concept of coherification, with asymmetrization of a symmetric state $\rho_0 = \mathcal{G}(\rho_0)$ being given by any state ρ such that $\mathcal{G}(\rho) = \rho_0$. In particular, the number of orthogonal asymmetrizations of ρ would be then directly related to the number of classical messages that one can encode in the resources destroyed by \mathcal{G} .

Acknowledgements: The authors would like to thank the anonymous referee of Ref. [35] that suggested asymptotic analysis of the number of orthogonal states with the same diagonal. KK would also like to thank D. Jennings and C. Cîrstoiu for helpful discussions. We acknowledge financial support by the Foundation for Polish Science through IRAP project cofinanced by EU within Smart Growth Operational Programme (contract no. 2018/MAB/5) and through TEAM-NET project (contract no. POIR.04.04.00-00-17C1/18-00). This research was also supported by National Science Center in Poland under the Maestro grant number DEC-2015/18/A/ST2/00274.

APPENDIX

A. Procedure for collective twirling

In this appendix we discuss how to find a pure state $|\psi_{1...n}^*\rangle$ that is mapped to a maximally mixed state by a resource destroying map \mathcal{G}_{col} given by Eq. (23a). For a given operator A, first using the results of Collins and Śniady [43] and then employing the explicit expressions provided by Audenaert [44], we have

$$\mathcal{G}_{\text{col}}(A) = \frac{1}{n!} \sum_{\pi \in S_n} \text{Tr} \left(A P_{\pi} \right) P_{\pi^{-1}} \sum_{\lambda \vdash n} \frac{f^{\lambda}}{s_{\lambda}(1^{\times d})} P^{\lambda}. \tag{55}$$

In the above P_{π} is an operator matrix responsible for permutations of subsystems according to a given permutation π and f^{λ} is the number of standard Young tableaux with shape given by a partition λ of n (denoted $\lambda \vdash n$). Next, $s_{\lambda}(1^{\times d})$ is a Schur polynomial related to partition λ evaluated at a point

$$1^{\times d} := \underbrace{1, 1, \dots, 1}_{d}. \tag{56}$$

Finally, operators P^{λ} are orthogonal projectors indexed by λ , which form an orthogonal set and add up to identity operator. These projectors can be defined using permutation matrices P_{π} and the character $\chi^{\lambda}(\pi)$ of permutation π (in irreducible representation of symmetric group labelled by partition λ),

$$P^{\lambda} = \frac{f^{\lambda}}{n!} \sum_{\pi \in S_n} \chi^{\lambda}(\pi) P_{\pi}.$$
 (57)

Crucially, observe that projectors P^{λ} are invariant under the action of \mathcal{G}_{col} ,

$$\mathcal{G}_{\text{col}}(P^{\lambda}) = P^{\lambda}. \tag{58}$$

For detailed definitions of components and factors above we refer the reader to Ref. [44].

In order to construct $|\psi_{1...n}^*\rangle$ one can look for states $|x^{\lambda}\rangle$, which belong to the non-zero eigenspace of P^{λ} , i.e.

$$P^{\lambda}|x^{\lambda}\rangle = |x^{\lambda}\rangle,\tag{59}$$

and such that under the action of $\mathcal{G}_{\mathrm{col}}$ they are mapped onto the full subspace,

$$\mathcal{G}_{\text{col}}(\left|x^{\lambda}\right\rangle\left\langle x^{\lambda}\right|) = \frac{P^{\lambda}}{\text{Tr}\left(P^{\lambda}\right)}.$$
 (60)

Then, the following linear combination of such states,

$$|x\rangle := \sum_{\lambda \vdash n} \sqrt{\frac{\text{Tr}(P^{\lambda})}{d^n}} |x^{\lambda}\rangle$$
 (61)

would give us the desired state $|\psi_{1...n}^*\rangle$, because

$$\mathcal{G}_{\text{col}}(|x\rangle\langle x|) = \sum_{\lambda \vdash n} \frac{P^{\lambda}}{d^n} = \frac{\mathbb{1}_{d^n}}{d^n}.$$
 (62)

As a particular example consider the case of three qubits, d=2 and n=3. We then have $P^{\{1,1,1\}}=0$ and

Now, if we take

$$|x^{\{2,1\}}\rangle = \frac{1}{2\sqrt{3}}(0, -2, 1, -\sqrt{3}, 1, \sqrt{3}, 0, 0)^{\top},$$
 (64a)

$$|x^{\{3\}}\rangle = \frac{1}{\sqrt{6}}(0, 1, 1, 1, 1, 1, 1, 0)^{\top},$$
 (64b)

and construct

$$|x\rangle = \frac{1}{\sqrt{2}}(|x^{\{2,1\}}\rangle + |x^{\{3\}}\rangle),$$
 (65)

we obtain

$$\mathcal{G}_{\text{col}}(|x\rangle\langle x|) = \frac{\mathbb{1}_8}{8}.$$
 (66)

REFERENCES

- [1] C. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.
- [2] A. Holevo, "The Capacity of the Quantum Channel with General Signal States," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 269–273, 1998. [Online]. Available: https://doi.org/10.1109/18.651037
- [3] B. Schumacher and M. Westmoreland, "Sending Classical Information via Noisy Quantum Channels," *Phys. Rev. A*, vol. 56, no. 1, pp. 131–138, 1997. [Online]. Available: https://doi.org/10.1103/PhysRevA.56.131
- [4] A. S. Holevo, "Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel," *Problems of Information Transmission*, vol. 9, no. 3, pp. 177–183, 1973. [Online]. Available: http://mi.mathnet.ru/eng/ppi903
- [5] M. B. Hastings, "Superadditivity of Communication Capacity Using Entangled Inputs," Nat. Phys., vol. 5, no. 4, pp. 255–257, 2009. [Online]. Available: https://doi.org/10.1038/nphys1224
- [6] T. Ogawa and H. Nagaoka, "Strong Converse to the Quantum Channel Coding Theorem," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2486– 2489, nov 1999. [Online]. Available: https://doi.org/10.1109/18.796386
- [7] A. Winter, "Coding Theorem and Strong Converse for Quantum Channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2481–2485, 1999. [Online]. Available: https://doi.org/10.1109/18.796385

- [8] M. Hayashi and H. Nagaoka, "General Formulas for Capacity of Classical-Quantum Channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1753–1768, 2003. [Online]. Available: https://doi.org/10.1109/TIT.2003.813556
- [9] L. Wang and R. Renner, "One-Shot Classical-Quantum Capacity and Hypothesis Testing," *Phys. Rev. Lett.*, vol. 108, no. 20, p. 200501, 2012. [Online]. Available: https://doi.org/10.1103/PhysRevLett.108.200501
- [10] M. Tomamichel and V. Y. F. Tan, "Second-Order Asymptotics for the Classical Capacity of Image-Additive Quantum Channels," *Commun. Math. Phys.*, vol. 338, no. 1, pp. 103–137, 2015. [Online]. Available: https://doi.org/10.1007/s00220-015-2382-0
- [11] M. Mosonyi and T. Ogawa, "Strong Converse Exponent for Classical-Quantum Channel Coding," 2014. [Online]. Available: http://arxiv.org/abs/1409.3562
- [12] C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, "Moderate Deviation Analysis for Classical Communication over Quantum Channels," *Commun. Math. Phys.*, vol. 355, no. 3, pp. 1283–1315, 2017. [Online]. Available: https://doi.org/10.1007/s00220-017-2971-1
- [13] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1385–1403, 2018. [Online]. Available: https://doi.org/10.1109/TIT.2017.2781254
- [14] M. Dalai, "Lower Bounds on the Probability of Error for Classical and Classical-Quantum Channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8027–8056, 2013. [Online]. Available: https://doi.org/10.1109/TIT.2013.2283794
- [15] H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel, "Quantum Sphere-Packing Bounds With Polynomial Prefactors," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 2872–2898, 2019. [Online]. Available: https://doi.org/10.1109/TIT.2019.2891347
- [16] Z.-W. Liu, X. Hu, and S. Lloyd, "Resource destroying maps," *Phys Rev. Lett.*, vol. 118, no. 6, p. 060502, 2017. [Online]. Available: https://doi.org/10.1103/PhysRevLett.118.060502
- [17] M. Horodecki and J. Oppenheim, "(quantumness in the context of) resource theories," *Int. J. Mod. Phys. B*, vol. 27, no. 01n03, p. 1345019, 2013. [Online]. Available: https://doi.org/10.1142/S0217979213450197
- [18] C. H. Bennett and S. J. Wiesner, "Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states," *Phys. Rev. Lett.*, vol. 69, no. 20, p. 2881, 1992. [Online]. Available: https://doi.org/10.1103/PhysRevLett.69.2881
- [19] S.-H. Kye, "Positive linear maps between matrix algebras which fix diagonals," *Linear Algebra Appl.*, vol. 216, pp. 239–256, 1995. [Online]. Available: https://doi.org/10.1016/0024-3795(93)00140-U
- [20] C.-K. Li and H. J. Woerdeman, "Special classes of positive and completely positive maps," *Linear Algebra Appl.*, vol. 255, no. 1-3, pp. 247–258, 1997. [Online]. Available: https://doi.org/10.1016/S0024-3795(96)00776-8
- [21] J. Åberg, "Quantifying superposition," arXiv:0612146, Dec. 2006. [Online]. Available: https://arxiv.org/abs/quant-ph/0612146
- [22] T. Baumgratz, M. Cramer, and M. B. Plenio, "Quantifying coherence," Phys. Rev. Lett., vol. 113, no. 14, p. 140401, 2014. [Online]. Available: https://doi.org/10.1103/PhysRevLett.113.140401
- [23] I. Marvian, "Symmetry, asymmetry and quantum information," Ph.D. dissertation, University of Waterloo, 2012. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/7088
- [24] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, "Resource theory of quantum states out of thermal equilibrium," *Phys. Rev. Lett.*, vol. 111, p. 250404, 2013. [Online]. Available: https://doi.org/10.1103/PhysRevLett.111.250404
- [25] F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 112, p. 3275, 2015. [Online]. Available: https://doi.org/10.1073/pnas.1411728112
- [26] H. Umegaki, "Conditional Expectation in an Operator Algebra," *Kodai Math. Sem. Rep.*, vol. 14, pp. 59–85, 1962. [Online]. Available: https://doi.org/10.2996/kmj/1138844604
- [27] M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7693–7710, 2013. [Online]. Available: https://doi.org/10.1109/TIT.2013.2276628
- [28] K. Li, "Second-Order Asymptotics for Quantum Hypothesis Testing," Ann. Stat., vol. 42, no. 1, pp. 171–189, feb 2014. [Online]. Available: https://doi.org/10.1214/13-AOS1185
- [29] R. Renner, "Security of quantum key distribution," Int. J. Quantum Inf., vol. 6, no. 01, pp. 1–127, 2008. [Online]. Available: https://doi.org/10.1142/S0219749908003256

- [30] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1753–1768, 2003. [Online]. Available: https://doi.org/10.1109/TIT.2003.813556
- [31] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, "Robustness of asymmetry and coherence of quantum states," *Phys. Rev. A*, vol. 93, no. 4, p. 042107, 2016. [Online]. Available: https://doi.org/10.1103/PhysRevA.93.042107
- [32] A. Winter and D. Yang, "Operational resource theory of coherence," Phys. Rev. Lett., vol. 116, no. 12, p. 120404, 2016. [Online]. Available: https://doi.org/10.1103/PhysRevLett.116.120404
- [33] K. Korzekwa, M. Lostaglio, D. Jennings, and T. Rudolph, "Quantum and classical entropic uncertainty relations," *Phys. Rev. A*, vol. 89, p. 042122, 2014. [Online]. Available: https://doi.org/10.1103/PhysRevA.89.042122
- [34] K. Korzekwa, S. Czachórski, Z. Puchała, and K. Życzkowski, "Coherifying quantum channels," New J. Phys., vol. 20, no. 4, p. 043028, 2018. [Online]. Available: https://doi.org/10.1088/1367-2630/aaaff3
- [35] ——, "Distinguishing classically indistinguishable states and channels," J. Phys. A: Math. Theor., vol. 52, p. 475303, 2019. [Online]. Available: https://doi.org/10.1088/1751-8121/ab30f7
- [36] G. Bowen, "Classical information capacity of superdense coding," Phys. Rev. A, vol. 63, no. 2, p. 022302, 2001. [Online]. Available: https://doi.org/10.1103/PhysRevA.63.022302
- [37] T. Hiroshima, "Optimal dense coding with mixed state entanglement," J. Phys. A, vol. 34, no. 35, p. 6907, 2001. [Online]. Available: https://doi.org/10.1088/0305-4470/34/35/316
- [38] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, "Decoherence-full subsystems and the cryptographic power of a private shared reference frame," *Phys. Rev. A*, vol. 70, no. 3, p. 032307, 2004. [Online]. Available: https://doi.org/10.1103/PhysRevA.70.032307
- [39] —, "Reference frames, superselection rules, and quantum information," Rev. Mod. Phys., vol. 79, no. 2, p. 555, 2007. [Online]. Available: https://doi.org/10.1103/RevModPhys.79.555
- [40] C. T. Chubb, M. Tomamichel, and K. Korzekwa, "Beyond the thermodynamic limit: finite-size corrections to state interconversion rates," *Quantum*, vol. 2, p. 108, 2018. [Online]. Available: https://doi.org/10.22331/q-2018-11-27-108
- [41] S. Beigi and A. Gohari, "Quantum achievability proof via collision relative entropy," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7980–7986, 2014. [Online]. Available: https://doi.org/10.1109/TIT.2014.2361632
- [42] K. Li et al., "Second-order asymptotics for quantum hypothesis testing," Ann. Stat., vol. 42, no. 1, pp. 171–189, 2014. [Online]. Available: https://doi.org/10.1214/13-AOS1185
- [43] B. Collins and P. Śniady, "Integration with respect to the haar measure on unitary, orthogonal and symplectic group," *Commun. Math. Phys.*, vol. 264, no. 3, pp. 773–795, 2006. [Online]. Available: http://dx.doi.org/10.1007/s00220-006-1554-3
- [44] K. Audenaert, "A digest on representation theory of the symmetric group," 2006. [Online]. Available: https://pdfs.semanticscholar.org/78bd/f436a0df9dd59bad52db9572de9c1aae008f.pdf