Gaussian optimizers and the additivity problem in quantum information theory

A. S. Holevo

Abstract. This paper surveys two remarkable analytical problems of quantum information theory. The main part is a detailed report on the recent (partial) solution of the quantum Gaussian optimizer problem which establishes an optimal property of Glauber’s coherent states—a particular case of pure quantum Gaussian states. The notion of a quantum Gaussian channel is developed as a non-commutative generalization of an integral operator with Gaussian kernel, and it is shown that the coherent states, and under certain conditions only they, minimize a broad class of concave functionals of the output of a Gaussian channel. Thus, the output states corresponding to a Gaussian input are the ‘least chaotic’, majorizing all the other outputs. The solution, however, is essentially restricted to the gauge-invariant case where a distinguished complex structure plays a special role. Also discussed is the related well-known additivity conjecture, which was solved in principle in the negative some five years ago. This refers to the additivity or multiplicativity (with respect to tensor products of channels) of information quantities related to the classical capacity of a quantum channel, such as the \((1 \rightarrow p)\)-norms or the minimal von Neumann or Rényi output entropies. A remarkable corollary of the present solution of the quantum Gaussian optimizer problem is that these additivity properties, while not valid in general, do hold in the important and interesting class of gauge-covariant Gaussian channels.

Bibliography: 65 titles.

Keywords: completely positive map, canonical commutation relations, Gaussian state, coherent state, quantum Gaussian channel, gauge covariance, von Neumann entropy, channel capacity, majorization.

Contents

1. Introduction 332
2. The additivity problem for quantum channels 334
   2.1. Definition of a channel 334
   2.2. Stinespring-type representation 335
   2.3. Entropic quantities and additivity 337
   2.4. The channel capacity 338

This work is supported by the Russian Science Foundation under grant 14-21-00162.
AMS 2010 Mathematics Subject Classification. Primary 94A40; Secondary 81P45, 81P68.

© 2015 Russian Academy of Sciences (DoM), London Mathematical Society, Turpion Ltd.
1. Introduction

The quantum Gaussian optimizer problem is an analytical problem that arose in quantum information theory at the end of the past century, and which is of independent mathematical interest. Only recently a solution was found [22], [53] in a quite common situation, though in full generality the problem still remains open. To explain the nature and the difficulty of the problem we start from the related classical problem of Gaussian maximizers, which has been studied rather exhaustively (see Lieb [51] and references therein). Consider an integral operator $G$ from $L_p(\mathbb{R}^s)$ to $L_q(\mathbb{R}^r)$ given by a Gaussian kernel (that is, the exponential of a quadratic form) with the $(q \to p)$-norm

$$\|G\|_{q \to p} = \sup_{f \neq 0} \frac{\|Gf\|_p}{\|f\|_q} = \sup_{\|f\|_q \leq 1} \|Gf\|_p.$$  \hspace{1cm} (1)

Under certain fairly broad assumptions concerning the quadratic form defining the kernel and also $p$ and $q$, this operator is well defined, and the supremum in (1) is attained on a Gaussian function $f$. Moreover, under some additional restrictions any maximizer is Gaussian. As is put in the title of [51], “Gaussian kernels have only Gaussian maximizers”.

Knowledge that the maximizer is Gaussian can be used to compute the exact value of the norm (1). In fact, the classical work on Gaussian maximizers began with results of Babenko [5] and then Beckner [6] which established the best constant in the Hausdorff–Young inequality concerning the $(p \to p')$-norm $(p^{-1} + (p')^{-1} = 1, 1 < p \leq 2)$ of the Fourier transform (which is obviously given by a degenerate imaginary Gaussian kernel).

A difficulty in the optimization problem (1) is that it requires maximization of a convex function, so the general theory of convex optimization is not of great use here (it only implies that a maximizer of $\|Gf\|_p$ belongs to a face of the convex set $\|f\|_q \leq 1$). Instead, the solution is based on substantial use of the classical Minkowski inequality and the related multiplicativity of the classical $(q \to p)$-norms with respect to tensor products of integral operators.

A notable application of these classical results to a problem in quantum mathematical physics was Lieb’s solution [50] of Wehrl’s conjecture [63]. On a separable Hilbert space $\mathcal{H}$ let $\rho$ be a density operator representing a state of a quantum
system. The ‘classical entropy’ of the state $\rho$ is defined as

$$H_{cl}(\rho) = -\int_{C} p_\rho(z) \log p_\rho(z) \, \frac{d^2 z}{\pi},$$

where $p_\rho(z) = \langle z | \rho | z \rangle$ is the diagonal value of the kernel of the operator $\rho$ in the system of Glauber coherent vectors $\{|z]; z \in \mathbb{C} \}$ [48], [33]. The conjecture was that $H_{cl}(\rho)$ has a minimal value if $\rho$ is itself a coherent state, that is, the projection on one of the coherent vectors. Lieb [50] used the exact constants in the Hausdorff–Young inequality for the $L_p$-norms of a Fourier transform [5], [6] and in the Young inequality for a convolution [6] to prove the corresponding maximizer conjecture for $f(x) = x^p$, and considered the limit $\lim_{p \downarrow 1} (1 - p)^{-1}(1 - x^p) = -x \log x$.

Lieb and Solovej [52], using a completely different approach based on a study of spin coherent states, recently strengthened the result in [50] by showing that the Glauber coherent states minimize any functional of the form $\int f(\langle z | \rho | z \rangle) \, \frac{d^2 z}{\pi}$, where $f(x), x \in [0, 1]$, is a non-negative concave function with $f(0) = 0$.

In the language of quantum information theory, the affine map $G: \rho \rightarrow p_\rho(z)$ taking density operators $\rho$ (quantum states) to probability densities $p_\rho(z)$ (classical states) is a ‘quantum-classical channel’ [38]. Moreover, it transforms Gaussian density operators $\rho$ (in the sense defined below in §3.1) into Gaussian probability densities, and in this sense it is a ‘Gaussian channel’. From this point of view, the Wehrl entropy $H_{cl}(\rho)$ is the output entropy of the channel, and Lieb’s result says that it is minimized by pure Gaussian states $\rho$. Moreover, the corresponding result for $f(x) = x^p$ can be interpreted as a statement about the ‘Gaussian maximizer’ for the norm $\|G\|_{1 \rightarrow p}$. We note that the case $q = 1$, which is excluded in the classical problem for obvious reasons, appears and plays a fundamental role in the quantum (non-commutative) case.

The quantum Gaussian optimizer problem described in the present paper relates to bosonic Gaussian channels, which are non-commutative analogues of Gaussian Markov kernels, and it similarly requires maximization of convex functions (or minimization of concave functions, such as entropy) of the output state of the channel, while the argument is the input state. A general conjecture is that the optimizer belongs to the class of pure Gaussian states. This conjecture, first formulated in [42] in the context of quantum information theory, however natural it looks, resisted numerous attacks for several years. Among other results, notable achievements were the exact determination of the classical capacity of a pure loss channel [21] and a proof of additivity of the Rényi entropies of integer orders $p$ [24] for special models of channels. Even restricted to the class of Gaussian input states, the optimization problem remains non-trivial [56], [31]. There was some hope that classical results on ‘Gaussian maximizers’ might help in solving this problem, as in the proof of Wehrl’s conjecture. However, the solution found recently in [22] by Giovannetti,

---

1Throughout this paper, the base of a logarithm is a fixed number $a > 1$. In information theory the natural choice is $a = 2$, when all the entropic quantities are measured in ‘bits’.

2In analysis, they correspond to complex-parametrized Gaussian wavelets. We note that this is the only place in the present article where we formally use Dirac’s notation, which is uncommon among mathematicians.
Holevo, and Garcia-Patron and in [53] by Mari, Giovannetti, and Holevo uses completely different ideas based on a thorough investigation of structural properties of quantum Gaussian channels. As already mentioned, the solution of the classical problem uses the Minkowski inequality and the consequent multiplicativity of the \((q \rightarrow p)\)-norms. However, the non-commutative analogue of Minkowski’s inequality \([12]\) is not powerful enough to guarantee multiplicativity of the norms (or additivity of the corresponding entropic quantities). Moreover, the related long-standing additivity problem in quantum information theory \([37]\) was recently shown to have a negative solution in general \([26]\). We show that, remarkably, the solution of the quantum Gaussian optimizer problem given in \([22]\) implies also a proof of the multiplicativity/additivity property in the restricted class of gauge-covariant or contravariant quantum Gaussian channels.

It would then be interesting to investigate the possible development of such an approach to obtain non-commutative generalizations of the classical ‘Gaussian maximizer’ results for \((q \rightarrow p)\)-norms. Such a generalization could shed new light on the hypercontractivity problem for quantum dynamical semigroups and the related non-commutative analogues of logarithmic Sobolev inequalities (see, for instance, \([62]\)).

2. The additivity problem for quantum channels

2.1. Definition of a channel. Let \(\mathcal{H}\) be a separable complex Hilbert space, \(\mathfrak{L}(\mathcal{H})\) the algebra of all bounded operators on \(\mathcal{H}\), and \(\mathfrak{S}(\mathcal{H})\) the ideal of trace-class operators. The space \(\mathfrak{S}(\mathcal{H})\), equipped with the trace norm \(\|\cdot\|_1\), is a Banach space which it is useful to consider as a non-commutative analogue of the space \(L_1\). The convex subset of \(\mathfrak{S}(\mathcal{H})\) defined by

\[
\mathfrak{S}(\mathcal{H}) = \{\rho: \rho^* = \rho \geq 0, \text{ Tr } \rho = 1\}
\]

is the base of the positive cone in \(\mathfrak{S}(\mathcal{H})\). Operators \(\rho\) in \(\mathfrak{S}(\mathcal{H})\) are called density operators or quantum states. The state space is a convex set with the extreme boundary

\[
\mathfrak{P}(\mathcal{H}) = \{\rho: \rho \geq 0, \text{ Tr } \rho = 1, \rho^2 = \rho\}.
\]

Thus, the extreme points of \(\mathfrak{S}(\mathcal{H})\), called pure states, are the one-dimensional projections \(\rho = P_\psi\) for vectors \(\psi \in \mathcal{H}\) with unit norm (see, for instance, \([55]\)).

The class of maps which will interest us is the class of non-commutative analogues of Markov maps (linear, positive, normalized maps) in classical analysis and probability theory. Let \(\mathcal{H}_A\) and \(\mathcal{H}_B\) be two Hilbert spaces, called the input and output spaces. A map \(\Phi: \mathfrak{S}(\mathcal{H}_A) \rightarrow \mathfrak{S}(\mathcal{H}_B)\) is positive if \(X \geq 0\) implies that \(\Phi[X] \geq 0\), and it is completely positive \([61], [54]\) if the maps \(\Phi \otimes \text{Id}(d)\) are positive for all \(d = 1, 2, \ldots\), where \(\text{Id}(d)\) is the identity map of the algebra \(\mathfrak{L}(d) = \mathfrak{L}(\mathbb{C}^d)\) of complex \(d \times d\) matrices. Equivalently, for every non-negative definite block matrix \([X_{jk}]_{j,k=1,\ldots,d}\) the matrix \([\Phi[X_{jk}]])_{j,k=1,\ldots,d}\) is non-negative definite.

A linear map \(\Phi\) is trace-preserving if \(\text{Tr } \Phi[X] = \text{Tr } X\) for all \(X \in \mathfrak{S}(\mathcal{H}_A)\).

Definition 1. A quantum channel is a linear completely positive trace-preserving map \(\Phi: \mathfrak{S}(\mathcal{H}_A) \rightarrow \mathfrak{S}(\mathcal{H}_B)\). The letter \(A\) will always be associated with the input of a channel, and the letter \(B\) with the output. For brevity we will sometimes simply write \(\Phi: A \rightarrow B\).
Obviously, every channel is a positive map taking states into states: \( \Phi[\mathcal{S}(\mathcal{H}_A)] \subseteq \mathcal{S}(\mathcal{H}_B) \). Since the space \( \mathfrak{T}(\mathcal{H}) \) is base-normed, this implies \(^{[16]}\) that \( \Phi \) is a bounded map from the Banach space \( \mathfrak{T}(\mathcal{H}_A) \) to \( \mathfrak{T}(\mathcal{H}_B) \). The dual map \( \Phi^* \) of \( \Phi \) is uniquely defined by the relation

\[
\text{Tr } \Phi[X]Y = \text{Tr } X \Phi^*[Y], \quad X \in \mathfrak{T}(\mathcal{H}_A), \ Y \in \mathfrak{L}(\mathcal{H}_B),
\]

and is called the dual channel. The dual channel is a linear completely positive \(^*-\)weakly continuous map from the algebra \( \mathfrak{L}(\mathcal{H}_B) \) to \( \mathfrak{L}(\mathcal{H}_A) \) which is unital: \( \Phi[I_{\mathcal{H}_B}] = I_{\mathcal{H}_A} \). Here and in what follows, \( I \) with a possible index denotes the identity operator on the corresponding Hilbert space.

There are positive maps that are not completely positive (a basic example is provided by the matrix transposition \( X \rightarrow X^\top \) in a fixed basis).

From the definition of complete positivity one easily derives \(^{[38]}\) that the composition of channels \( \Phi_2 \circ \Phi_1 \) defined as

\[
\Phi_2 \circ \Phi_1[X] = \Phi_2[\Phi_1[X]]
\]

and the naturally defined tensor product of channels

\[
\Phi_1 \otimes \Phi_2 = (\Phi_1 \otimes \text{Id}_2) \circ (\text{Id}_1 \otimes \Phi_2)
\]

are again channels.

### 2.2. Stinespring-type representation.

The notion of completely positive map was introduced by Stinespring \(^{[61]}\) in the much wider context of \( \mathcal{C}^* \)-algebras. This encompasses also the notion of a hybrid channel, which has a quantum input and classical output, or vice versa. An example of such a channel was mentioned in §1. We will not pursue this topic further here (see \(^{[38]}\)), but only mention that complete positivity reduces to positivity in such cases.

Motivated by the well-known Naimark dilation theorem, Stinespring established a representation for completely positive maps of \( \mathcal{C}^* \)-algebras which in the case of a quantum channel reduces \(^{[38]}\) to the following result.

**Proposition 2.** Let \( \Phi: A \rightarrow B \) be a quantum channel. There exist a Hilbert space \( \mathcal{H}_E \) and an isometric operator \( V: \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E \) such that

\[
\Phi[\rho] = \text{Tr}_{\mathcal{E}} V\rho V^*, \quad \rho \in \mathfrak{T}(\mathcal{H}_A),
\]

where \( \text{Tr}_{\mathcal{E}} \) denotes the partial trace with respect to \( \mathcal{H}_E \). The representation (3) is not unique, but any two representations with operators \( V_1: \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_{E_1} \) and \( V_2: \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_{E_2} \) are related via a partial isometry \( W: \mathcal{H}_{E_1} \rightarrow \mathcal{H}_{E_2} \) such that \( V_2 = (I_B \otimes W)V_1 \) and \( V_1 = (I_B \otimes W^*)V_2 \).

Consider a representation (3) for the channel \( \Phi \). The complementary channel is then defined by

\[
\Phi^*[\rho] = \text{Tr}_{\mathcal{B}} V\rho V^*, \quad \rho \in \mathfrak{T}(\mathcal{H}_A).
\]
complementary to $\Phi$ are isometrically equivalent in the sense that there is a partial isometry $W: \mathcal{H}_E_1 \to \mathcal{H}_E_2$ such that

$$\tilde{\Phi}_2[\rho] = W\tilde{\Phi}_1[\rho]W^*, \quad \tilde{\Phi}_1[\rho] = W^*\tilde{\Phi}_2[\rho]W$$

(5)

for all $\rho$. It follows that the initial projection $W^*W$ satisfies $\tilde{\Phi}_1[\rho] = W^*W\tilde{\Phi}_1[\rho]$, that is, its support contains the support of $\tilde{\Phi}_1[\rho]$, while the final projection $WW^*$ has the analogous property with respect to $\tilde{\Phi}_2[\rho]$.

In general, we will say that two density operators $\rho$ and $\sigma$ (possibly acting in different Hilbert spaces) are isometrically equivalent if there is a partial isometry $W$ such that $\rho = W\sigma W^*$ and $\sigma = W^*\rho W$. Obviously, this is the case if and only if the non-zero eigenvalues (counting multiplicity) of $\rho$ and $\sigma$ coincide. We denote this condition by $\rho \sim \sigma$. We have just shown that $\tilde{\Phi}_1[\rho] \sim \tilde{\Phi}_2[\rho]$ for arbitrary $\rho$.

The complementary channel to a complementary channel can be shown to be isometrically equivalent to the original channel, so that $\Phi$ and $\tilde{\Phi}$ can be said to be mutually complementary channels.

**Lemma 3.** If $\tilde{\Phi}$ is a complementary channel (4), then $\Phi[P_\psi] \sim \tilde{\Phi}[P_\psi]$ for all $\psi \in \mathcal{H}_A$.

**Proof.** Let $V: \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ be the isometry in the representations (3) and (4). Then $\rho_{BE} = VP_\psi V^*$ is a pure state on $\mathcal{H}_B \otimes \mathcal{H}_E$, and the statement follows from a basic result in quantum information theory (‘Schmidt decomposition’): if $\rho_{BE}$ is a pure state on $\mathcal{H}_B \otimes \mathcal{H}_E$ and $\rho_B = \text{Tr}_E \rho_{BE}$ and $\rho_E = \text{Tr}_B \rho_{BE}$ are its partial states, then $\rho_B \sim \rho_E$ (see, for instance, Proposition 3 in [37]). □

Another name for channel is dynamical map. In non-equilibrium quantum statistical mechanics such maps arise as a result of irreversible evolution of an open quantum system interacting with an environment [38]. Assume that there is a composite quantum system $AD = BE$ in the Hilbert space

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_D \simeq \mathcal{H}_B \otimes \mathcal{H}_E$$

(6)

which is initially prepared in the state $\rho_A \otimes \rho_D$ and then evolves according to the unitary operator $U$. Then the output state $\rho_B$ depending on the input state $\rho_A = \rho$ is

$$\Phi_B[\rho] = \text{Tr}_E U(\rho \otimes \rho_D)U^*,$$

(7)

while the output state of the ‘environment’ $E$ is the output for the channel

$$\Phi_E[\rho] = \text{Tr}_B U(\rho \otimes \rho_D)U^*.$$ 

(8)

If the initial state of $D$ is pure, $\rho_D = P_\psi_D$, then by introducing the isometry $V: \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_D$ acting as

$$V\psi = U(\psi \otimes \psi_D), \quad \psi \in \mathcal{H}_A,$$

we see that the relations (7) and (8) become (3) and (4), and $\Phi_E$ is just the complementary channel of $\Phi_B$. We remark also that both the taking of a partial trace and a unitary evolution are completely positive operations, and hence the maps (7)
and (8) are completely positive. Conversely, any quantum channel has a representation of such form (see, for instance, [38]).

A vast literature has been devoted to the study of quantum dynamical semigroups (the non-commutative analogue of Markov semigroups) and quantum Markov processes. The Stinespring-type representation (3) lies at the basis of dilations of quantum dynamical semigroups to the unitary dynamics of open quantum systems interacting with an environment [16], [34].

2.3. Entropic quantities and additivity. Consider the norm of the map $\Phi$ defined like (1):

$$\|\Phi\|_{1\rightarrow p} = \sup_{X \neq 0} \frac{\|\Phi[X]\|_p}{\|X\|_1} = \sup_{\|X\|_1 \leq 1} \|\Phi[X]\|_p,$$

where $\|\cdot\|_p$ is the Schatten $p$-norm [55]. As shown in [4],

$$\|\Phi\|_{1\rightarrow p}^p = \sup_{\rho \in \mathcal{S}(\mathcal{H}_A)} \text{Tr} \Phi[\rho]^p = \sup_{\psi \in \mathcal{H}_A} \text{Tr} \Phi[P\psi]^p,$$

where the second equality follows from the convexity of the function $x^p$, $p > 1$.

The quantum Rényi entropy of order $p > 1$ for a density operator $\rho$ is defined as

$$R_p(\rho) = \frac{1}{1 - p} \log \text{Tr} \rho^p = \frac{p}{1 - p} \log \|\rho\|_p.$$

We define the minimal output Rényi entropy of the channel $\Phi$ as

$$\tilde{R}_p(\Phi) = \inf_{\rho \in \mathcal{S}(\mathcal{H})} R_p(\Phi[\rho]) = \frac{p}{1 - p} \log \|\Phi\|_{1\rightarrow p},$$

and the minimal output von Neumann entropy as

$$\tilde{H}(\Phi) = \inf_{\rho \in \mathcal{S}(\mathcal{H})} H(\Phi[\rho]).$$

In the limit as $p \to 1$ the quantum Rényi entropies are monotonically non-decreasing and converge to the von Neumann entropy:

$$\lim_{p \to 1} R_p(\rho) = - \text{Tr} \rho \log \rho = H(\rho).$$

In finite dimensions the set of quantum states is compact, and hence by Dini’s Lemma the minimal output Rényi entropies converge to the minimal output von Neumann entropy.\(^3\)

Multiplicativity of the norm (9) for some channels $\Phi_1$ and $\Phi_2$, namely,

$$\|\Phi_1 \otimes \Phi_2\|_{1\rightarrow p} = \|\Phi_1\|_{1\rightarrow p} \cdot \|\Phi_2\|_{1\rightarrow p},$$

is equivalent to the additivity of the minimal output Rényi entropies:

$$\tilde{R}_p(\Phi_1 \otimes \Phi_2) = \tilde{R}_p(\Phi_1) + \tilde{R}_p(\Phi_2).$$

\(^3\)The corresponding statement is not valid for infinite-dimensional channels (not even for classical channels with a countable set of states) — M. E. Shirokov, private communication.
Closely related is the analogous property for the minimal output von Neumann entropy:

$$\tilde{H}(\Phi_1 \otimes \Phi_2) = \tilde{H}(\Phi_1) + \tilde{H}(\Phi_2).$$  \hfill (16)

In finite dimensions, the validity of (15) for certain channels $\Phi_1$ and $\Phi_2$ and for values of $p$ close to 1 implies (16) for these channels.

In the last two relations the inequality $\leq$ (like the inequality $\geq$ in (14)) is obvious, because the right-hand side is equal to the infimum over the subset of product states $\rho = \rho_1 \otimes \rho_2$. On the other hand, the existence of ‘entangled’ pure states which are not reducible to product states is the reason for possible violation of the equality for quantum channels.

2.4. The channel capacity. The practical importance of the additivity property (16) is revealed in connection with the notion of channel capacity. To explain it we assume for the moment that $\mathcal{H}_A$ and $\mathcal{H}_B$ are finite-dimensional.

For a quantum channel $\Phi$, a non-commutative analogue of the Shannon capacity, which we call the $\chi$-capacity, is defined by

$$C_{\chi}(\Phi) = \sup_{\{\pi_j, \rho_j\}} \left( H\left( \Phi \left[ \sum_j \pi_j \rho_j \right] \right) - \sum_j \pi_j H(\Phi[\rho_j]) \right),$$  \hfill (17)

where the supremum is over all quantum ensembles, that is, finite collections of states $\{\rho_1, \ldots, \rho_n\}$ with corresponding probabilities $\{\pi_1, \ldots, \pi_n\}$. The quantity (17) is closely related to the capacity $C(\Phi)$ of a quantum channel $\Phi$ for transmitting classical information [37]. The classical capacity of a quantum channel is roughly defined as the maximal transmission rate per use of the channel, with coding and decoding optimally chosen for an increasing number $n$ of independent uses of the channel

$$\Phi^\otimes n = \Phi \otimes \cdots \otimes \Phi,$$

so that the probability of a decoding error goes to zero as $n \to \infty$ (for a precise definition see [38]). The Holevo–Schumacher–Westmoreland theorem [32], which is a basic result of quantum information theory, says that the so-defined capacity $C(\Phi)$ is related to $C_{\chi}(\Phi)$ by the formula

$$C(\Phi) = \lim_{n \to \infty} \frac{1}{n} C_{\chi}(\Phi^\otimes n).$$

Since $C_{\chi}(\Phi)$ is easily seen to be superadditive, that is,

$$C_{\chi}(\Phi_1 \otimes \Phi_2) \geq C_{\chi}(\Phi_1) + C_{\chi}(\Phi_2),$$

one has $C(\Phi) \geq C_{\chi}(\Phi)$. However, if additivity

$$C_{\chi}(\Phi_1 \otimes \Phi_2) = C_{\chi}(\Phi_1) + C_{\chi}(\Phi_2)$$  \hfill (18)

holds for a given channel $\Phi_1 = \Phi$ and an arbitrary channel $\Phi_2$, then

$$C_{\chi}(\Phi^\otimes n) = n C_{\chi}(\Phi),$$  \hfill (19)
implying that
\[ C(\Phi) = C_\chi(\Phi). \]  
(20)

The reason for possible violation of the equality here, as well as in the cases (14), (15), and (16), is the existence of entangled states which are not reducible to product states on the input of the tensor product channel \( \Phi^{\otimes n} \).

2.5. Main conclusions. Thus, it was natural to ask whether the additivity property (16) holds globally, that is, for the tensor product of any pair of quantum channels \( \Phi_1, \Phi_2 \). This problem goes back to [8] (see also [37]). Quite remarkably, Shor [60] (see also [19]) proved that the \( \chi \)-capacity is globally additive if and only if the minimal output entropy has that property.

**Theorem 4** [60]. The properties (18) and (16) are globally equivalent in the sense that if one of them holds for all channels \( \Phi_1, \Phi_2 \), then the other also holds for all channels.

Additivity is rather simple to prove for all classical channels (see, for instance, [15]), but in the quantum case the question remained open for a dozen years, and was ultimately solved in the negative.

A detailed history of the problem up to 2006 can be found in [30], and here we only sketch the basic steps and the final resolution. In [1] it was proposed to approach the additivity property (16) via multiplicativity (14) of the \( (1 \to p) \)-norms (equivalent to additivity (15) of the minimal output Rényi entropies). The first explicit example where this property breaks down for \( d = \text{dim} \mathcal{H} > 3 \) and sufficiently large \( p \) was a transpose-depolarizing channel [64]:

\[ \Phi(\rho) = \frac{1}{d-1} [I \ \text{Tr} \rho - \rho^\top], \]  
(21)

where \( \rho \in \mathcal{L}_d \) is a matrix and \( \rho^\top \) its transpose. In particular, (15) with \( \Phi_1 = \Phi_2 = \Phi \) fails to hold for \( p \geq 4.7823 \) if \( d = 3 \) (nevertheless, the additivity of \( \hat{H}(\Phi) \) and \( C_\chi(\Phi) \) holds for this channel). Five years later came the important findings of Winter [65] and Hayden [28] (see also [29]), who showed the existence of a pair of channels breaking the additivity of the minimal output Rényi entropy for all values of the parameter \( p > 1 \). The method of these and subsequent works is random choice of the channels, which for given dimensions are parametrized by the isometries \( V \) in the representation (3), as well as random choice of the input states of the channels, combined with sufficiently precise probabilistic estimates for the norms of type (10). For finite dimensions the corresponding parametric sets are compact, and one usually takes the uniform distribution. Based on this progress, Hastings [26] gave a proof of the existence of channels breaking the additivity conjecture (16), formally corresponding to \( p = 1 \), in very high dimensions. Moreover, the probability of violation of additivity tends to 1 as the dimension tends to infinity. Hastings gave only a sketch of the argument, a detailed proof following his approach was given by Fukuda, King, and Moser [18], and this was further simplified by Brandao and M. Horodecki [10]. Later Szarek and coauthors [3] proposed a proof based on the Dvoretzky–Mil’man theorem on almost Euclidean sections of high-dimensional convex bodies.
Although when combined with Theorem 4 this gives a definite negative answer to the additivity conjectures, several important issues remain open. All the proofs use the technique of random unitary channels or random states and as such are not constructive: they prove only the existence of counterexamples but do not enable one to actually produce them. Attempts to use Hastings’ approach to get estimates for dimensions in which additivity can fail have led to excessively high values: the detailed estimates made in [18] gave $d \approx 10^{32}$ for breaking the additivity by a quantity of order $10^{-5}$. The best result in this direction, obtained in [7], states that “violations of additivity of the minimal output entropy, using random unitary channels and a maximally entangled state, can occur if and only if the output space has dimension at least 183. Almost surely, the defect of additivity is less than $\log 2$, and it can be made as close as desired to $\log 2$”.

While this does not exclude the possibility of better estimates, based perhaps on different (but yet unknown) models, it casts doubt on the expediency of searching for concrete counterexamples by computer simulation of random channels. From this point of view, the following explicit example given in [25] is of interest. Consider the completely positive map

$$\rho \to \Phi_-[\rho] = \text{Tr}_2 P_- \rho P_-,$$

where $P_-$ is the projection onto the antisymmetric subspace $\mathcal{H}_-$ of $\mathcal{H} \otimes \mathcal{H}$, which has dimension $d(d-1)/2$, and the partial trace is taken with respect to the second copy of $\mathcal{H}$. Its restriction to the operators with support in $\mathcal{H}_-$ is trace preserving and hence is a channel. It can be shown [38] that $\Phi_- = ((d-1)/2)\Phi^*$, where $\Phi^*$ is the dual to the complementary channel of (21). For this simple channel the minimal Rényi entropies are non-additive for all $p > 2$ and sufficiently large $d$, but unfortunately it is not clear whether this can be extended to the most interesting range $p \geq 1$.

Coming back to arbitrary channels, it remains unclear what happens in small dimensions: perhaps additivity still holds generically for some unknown reason, or its violation is so slight that it cannot be revealed by numerical simulations. This is indeed surprising in view of the fact that the physical reason for non-additivity is entanglement between the inputs of parallel quantum channels (see [38] for more details).

On the other hand, these results stress the importance of continuing efforts to find special cases where for some reason additivity holds and can be established analytically.

A survey of the main classes of such ‘additive’ channels acting in finite dimensions was presented in [37]. Below we briefly list the most important classes of channels $\Phi$ for which the additivity properties (16), (18), and (15) for $p > 1$ have been established with $\Phi = \Phi_1$ and arbitrary $\Phi_2$.

- **$q$-bit unital channels**, that is, channels $\Phi : \mathcal{L}_2 \to \mathcal{L}_2$ satisfying the condition $\Phi[I] = I$ [45]. Strikingly, there is still no analytical proof of additivity for non-unital qubit channels, in spite of convincing numerical evidence [27].

- **Depolarizing channel in $\mathcal{L}_d$**:

$$\Phi[\rho] = (1-p)\rho + pI \frac{d}{d^2-1} \text{Tr} \rho, \quad 0 \leq p \leq \frac{d^2}{d^2-1};$$
the only unitarily covariant channel, it can be regarded as a non-commutative
analogue of a completely symmetric channel in classical information theory [15].
The additivity properties (16), (15), and (18) were proved by King [46].

- **Entanglement-breaking channels.** In finite dimensions these are channels of the form
  \[ \Phi[\rho] = \sum_j \rho_B \text{Tr} \rho M_A, \]
  where \( \{M_A\} \) is a resolution of the identity on \( \mathcal{H}_A: M_A \geq 0, \sum_j M_A = I_A \), and \( \rho_B \in \mathcal{S}(\mathcal{H}_B) \) (see [43]).

For finite-dimensional entanglement-breaking channels the additivity of the minimal output von Neumann entropy and of the \( \chi \)-capacity was established by Shor [59] and the additivity of the minimal output Rényi entropies by King [44]. The additivity properties of entanglement-breaking channels were generalized to infinite dimensions by Shirokov [57].

- **Complementary channels.** The additivity of the minimal output entropy is equivalent for a channel \( \Phi \) and its complementary channel \( \hat{\Phi} \) (see Lemma 5 below). The class of channels complementary to entanglement-breaking channels contains the Schur-multiplication maps for matrices \( \rho = [c_{jk}]_{j,k=1,\ldots,d} \) in \( \mathcal{L}_d \):
  \[ \hat{\Phi}[\rho] = [\gamma_{jk}c_{jk}]_{j,k=1,\ldots,d}, \]
  where \( [\gamma_{jk}]_{j,k=1,\ldots,d} \) is a non-negative definite matrix such that \( \gamma_{jj} \equiv 1 \). For these channels, which are also called Hadamard channels, the additivity of the \( \chi \)-capacity has also been established [47].

In the next sections we consider bosonic Gaussian channels which act in infinite-dimensional spaces. It is conjectured that all the additivity properties hold in this class of channels. One of the main goals of the present paper is to show that additivity holds for a wide class of gauge co- or contravariant Gaussian channels, that is, those which respect a fixed complex structure in the underlying symplectic space.

2.6. Majorization for quantum states. From now on we again allow the Hilbert spaces in question to be infinite-dimensional. Denote by \( \mathcal{F} \) the class of real concave functions \( f \) on \( [0,1] \) with \( f(0) = 0 \). For any \( f \in \mathcal{F} \) and any density operator \( \rho \) we can consider the quantity
  \[ \text{Tr} f(\rho) = \sum_j f(\lambda_j), \]
  where the \( \lambda_j \) are the (non-zero) eigenvalues of the density operator \( \rho \) (counting multiplicity). Note that this quantity is defined unambiguously and has values in \( (-\infty, \infty) \). This follows from the fact that \( f(x) \geq cx \), where \( c = f(1) \), and hence \( \text{Tr} f(\rho) \geq c \text{Tr} \rho = c \). We will also use the fact that the functional \( \rho \rightarrow \text{Tr} f(\rho) \) is (strictly) concave on \( \mathcal{S}(\mathcal{H}) \) if \( f \) is (strictly) concave (see, for instance, [11]).

Denote by \( \lambda_j(\rho) \) the eigenvalues of a density operator \( \rho \) (counting multiplicity), arranged in non-increasing order. One says that \( \rho \) majorizes a density operator \( \sigma \)
if

\[
\sum_{j=1}^{k} \lambda_j^1(\rho) \geq \sum_{j=1}^{k} \lambda_j^1(\sigma), \quad k = 1, 2, \ldots.
\]

A consequence of a well-known result (see, for instance, [11]) is that this is the case if and only if \( \text{Tr} f(\rho) \leq \text{Tr} f(\sigma) \) for all \( f \in \mathfrak{F} \).

For a quantum channel \( \Phi \) we introduce the quantity

\[
\bar{f}(\Phi) = \inf_{\rho \in \mathcal{S}(\mathcal{H})} \text{Tr} f(\Phi[\rho]) = \inf_{P_\psi \in \mathcal{P}(\mathcal{H})} \text{Tr} f(\Phi[P_\psi]),
\]

where the last equality follows from the concavity of the functional \( \rho \rightarrow \text{Tr} f(\Phi[\rho]) \) on \( \mathcal{S}(\mathcal{H}) \). Moreover, for strictly concave \( f \) any minimizer has the form \( P_\psi \) for some unit vector \( \psi \in \mathcal{H} \).

In particular, taking \( f(x) = -x \log x \) or \( f(x) = -x^p \), we get that \( \bar{f}(\Phi) = \bar{H}(\Phi) \) or \( \bar{f}(\Phi) = -\|\Phi\|_{1-p} \).

Lemma 5 [36]. For complementary channels, \( \bar{f}(\Phi) = \bar{f}(\overline{\Phi}) \). Hence, \( \|\Phi\|_{1-p} = \|\bar{\Phi}\|_{1-p} \), \( \bar{H}(\Phi) = \bar{H}(\overline{\Phi}) \), \( \bar{R}_p(\Phi) = \bar{R}_p(\overline{\Phi}) \), and the multiplicativity (14) and also the additivity (16), (15) of the minimal output entropies hold simultaneously for the pairs of channels \( \Phi_1, \Phi_2 \) and \( \overline{\Phi}_1, \overline{\Phi}_2 \).

Proof. By Lemma 3, \( \Phi[P_\psi] \) and \( \overline{\Phi}[P_\psi] \) have the same non-zero spectrum (\( \Phi[P_\psi] \sim \overline{\Phi}[P_\psi] \)). Then

\[
\text{Tr} f(\Phi[P_\psi]) = \text{Tr} f(\overline{\Phi}[P_\psi]),
\]

since \( f(0) = 0 \). The second inequality in (22) implies that \( \bar{f}(\Phi) = \bar{f}(\overline{\Phi}) \).

The statement about multiplicativity (additivity) then follows from the fact that the channel \( \bar{\Phi}_1 \otimes \bar{\Phi}_2 \) is complementary to \( \Phi_1 \otimes \Phi_2 \).

3. Quantum Gaussian systems

3.1. Gaussian states and channels. A real vector space \( Z \) equipped with a non-degenerate skew-symmetric form \( \Delta(z, z') \) is called a symplectic space. In what follows, \( Z \) is finite-dimensional, in which case its dimension is necessarily even: \( \dim Z = 2s \) [49]. A basis \( \{e_j, h_j; j = 1, \ldots, s\} \) in which the form \( \Delta(z, z') \) has the matrix

\[
\Delta = \text{diag} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}_{j=1,\ldots,s}
\]

is said to be symplectic. A Weyl system in a Hilbert space \( \mathcal{H} \) is a strongly continuous family \( \{W(z); z \in \mathbb{Z}\} \) of unitary operators satisfying the Weyl–Segal canonical commutation relations (CCR)

\[
W(z)W(z') = \exp \left\{-\frac{i}{2} \Delta(z, z') \right\} W(z + z').
\]

Thus, \( z \rightarrow W(z) \) is a projective representation of the additive group of \( Z \). We always assume that the representation is irreducible. The Stone–von Neumann uniqueness theorem says that such a representation is unique up to unitary equivalence. It is
well known (see, for instance, [55]) that there is a family of selfadjoint operators $z \rightarrow R(z)$ with a common essential domain $\mathcal{D}$ such that

$$W(z) = \exp\{iR(z)\},$$

and moreover, for any symplectic basis \(\{e_j, h_j; j = 1, \ldots, s\}\)

$$R(z) = \sum_{j=1}^{s} (x_j q_j + y_j p_j)$$
on \mathcal{D}, where $R(e_j) = q_j$, $R(h_j) = p_j$, and $[x_1, y_1, \ldots, x_s, y_s]$ are the coordinates of the vector $z$ in this basis. Here the canonical observables $q_j, p_j$ with $j = 1, \ldots, s$ are selfadjoint operators acting in $\mathcal{H}$ and satisfying the Heisenberg CCR on $\mathcal{D}$:

$$[q_j, p_k] \subseteq i\delta_{jk} I,$$

$$[q_j, q_k] = 0, \quad [p_j, p_k] = 0. \quad (26)$$

In physics a symplectic space is the phase space of a classical system (such as the modes of electromagnetic radiation in a cavity), the quantum version of which is described by the CCR. Then $s$ is number of degrees of freedom, or ‘normal modes’, of the classical system.

The state given by a density operator $\rho$ in $\mathcal{H}$ is said to be Gaussian if its quantum characteristic function

$$\phi(z) = \text{Tr} \rho W(z)$$

has the form

$$\phi(z) = \exp\left\{im(z) - \frac{1}{2} \alpha(z, z)\right\}, \quad (27)$$

where $m$ is a real linear form and $\alpha$ is a real bilinear symmetric form on $Z$. A necessary and sufficient condition for (27) to define a state is the non-negative definiteness of the (complex) Hermitian form\(^4\) $\alpha(z, z') - (i/2)\Delta(z, z')$ on $Z$, or briefly,

$$\alpha \geq \frac{i}{2}\Delta. \quad (28)$$

We will agree that the matrix of a bilinear form in a fixed symplectic base is denoted by the same letter, and then the relation (28) can be understood as an inequality for Hermitian matrices, where $\alpha$ is a real symmetric matrix and $\Delta$ is a real skew-symmetric matrix.

A Gaussian state is pure if and only if $\alpha$ is a minimal solution of this inequality (see, for instance, [39]). An operator $J$ in $Z$ is called an operator of complex structure if

$$J^2 = -I, \quad (29)$$

where $I$ is the identity operator on $Z$, and the bilinear form $\Delta(z, Jz')$ is a (Euclidean) inner product in $Z$, that is,

$$\Delta(z, Jz') = \Delta(z', Jz) \quad (= -\Delta(Jz, z')); \quad (30)$$

$$\Delta(z, Jz) \geq 0, \quad z \in Z. \quad (31)$$

The following characterization can be found in [17] and [38].

\(^4\) A complex-valued real-bilinear form $\beta(z, z')$ on $Z$ will be called Hermitian if $\beta(z', z) = \overline{\beta(z, z')}$. 
Proposition 6. The minimal solutions of the inequality (28) are in a one-to-one correspondence with the operators of complex structure $J$ on $Z$ that is given by the relation

$$\alpha(z, z') = \frac{1}{2} \Delta(z, Jz'), \quad z, z' \in Z.$$ 

Thus, to each complex structure there corresponds a family of pure Gaussian states (27) that is parametrized by different values of $m$, called $J$-coherent states. The state with $m = 0$ is called the $J$-vacuum. If $\rho_0$ is a vacuum, then the corresponding coherent states are of the form $W(z')\rho_0 W(z')^*$, as follows from the relation

$$W(z')^* W(z) W(z') = \exp\{-i\Delta(z, z')\} W(z)$$

and from the non-degeneracy of the form $\Delta(z, z')$, due to which any linear form on $Z$ can be expressed as $m(z) = \Delta(z, z_m')$.

An operator $S$ in $Z$ is said to be symplectic if $\Delta(Sz, Sz') = \Delta(z, z')$ for all $z, z' \in Z$. The unitary operators $W(Sz)$ satisfy the CCR (25), and hence by the Stone–von Neumann uniqueness theorem there is a unitary operator $U_S$ on $\mathcal{H}$ such that

$$W(Sz) = U_S^* W(z) U_S, \quad z \in Z.$$ 

The map $S \to U_S$ is a projective representation of the group of all symplectic transformations of $Z$, sometimes called a ‘metaplectic representation’ [2], since it can be extended to a faithful unitary representation of the metaplectic group which is a two-fold covering of the symplectic group.

Similarly, an operator $T$ is antisymplectic if $\Delta(Tz, Tz') = -\Delta(z, z')$ for all $z, z' \in Z$. There is an antiunitary operator $U_T$ on $\mathcal{H}$ such that

$$W(Tz) = U_T^* W(z) U_T, \quad z \in Z.$$ 

Let $Z_A$ and $Z_B$ be two symplectic spaces with the corresponding Weyl systems. A channel $\Phi: A \to B$ is said to be Gaussian if the dual channel satisfies

$$\Phi^*[W_B(z)] = W_A(Kz) \exp\left\{il(z) - \frac{1}{2} \mu(z, z)\right\}, \quad z \in Z_B,$$

where $K: Z_B \to Z_A$ is a linear operator, $l$ a linear form, and $\mu$ a real symmetric form on $Z_B$. In terms of characteristic functions of states,

$$\phi_B(z) = \phi_A(Kz) \exp\left\{il(z) - \frac{1}{2} \mu(z, z)\right\}.$$ 

It follows that a Gaussian channel maps Gaussian states into Gaussian states. The converse statement also holds [17].

A necessary and sufficient condition on the parameters $(K, l, \mu)$ for complete positivity of the map $\Phi$ is the non-negative definiteness of the Hermitian form

$$z, z' \to \mu(z, z') - \frac{i}{2} [\Delta_B(z, z') - \Delta_A(Kz, Kz')]$$

on $Z_B$, or in matrix terms (if some bases are chosen in $Z_A$ and $Z_B$)

$$\mu \geq \frac{i}{2} (\Delta_B - K^t \Delta_A K),$$

(33)
where $^t$ denotes transposition of a matrix. A proof of this using an explicit construction of a representation of type (7) is given in [14] (see also [38]). In Proposition 12 below we give such a construction for an important particular class of Gaussian channels.

We call a Gaussian channel extreme\footnote{In quantum optics one speaks of quantum-limited channels [20].} if $\mu$ is a minimal solution of the inequality (33). This terminology stems from the fact that the minimality of $\mu$ is necessary and sufficient for the channel $\Phi$ to be an extreme point in the convex set of all channels with fixed input and output spaces [39].

**Additivity conjecture for quantum Gaussian channels:** The additivity properties (15) and (16) hold for any pair of Gaussian channels $\Phi_1, \Phi_2$.

**Quantum Gaussian minimizer conjecture:** For any function $f \in \mathcal{F}$ the infimum in (22) is attained on a pure Gaussian state $\rho$.

Any Gaussian channel has the covariance property

$$
\Phi[W_A(z)\rho W_A(z)^*] = W_B(K^s z)\Phi[\rho] W_B(K^s z)^*,
$$

(34)

where $K^s$ is the symplectic adjoint operator defined by the relation

$$
\Delta_B(K^s z_A, z_B) = \Delta_A(z_A, K z_B).
$$

It follows that the value $\text{Tr} f(\Phi[\rho])$ is the same for all coherent states $W(z)\rho_0 W(z)^*$ associated with a given vacuum state $\rho_0$.

These two problems turn out to be closely related. In what follows we describe a positive solution for both of them in a particular and important class of Gaussian channels with gauge symmetry. However, both conjectures remain open for general quantum Gaussian channels.

### 3.2. Complex structures and gauge symmetry

Given an operator of complex structure $J$, one can define in $Z$ a Euclidean inner product $j(z, z') = \Delta(z, Jz')$. Then one can define in $Z$ the structure of an $s$-dimensional unitary space $\mathcal{Z}$ in which $i z$ corresponds to $J z$ and the (Hermitian) inner product\footnote{In accordance with the convention in mathematical physics, the inner product is complex linear with respect to $z'$ and anti-linear with respect to $z$.} is

$$
j(z, z') = \frac{1}{2} [\Delta(z, Jz') + i \Delta(z, z')] = \frac{1}{2} [j(z, z') - ij(z, Jz')].
$$

From (29) and (30) it follows that $J$ is symplectic, that is, $\Delta(J z, J z') = \Delta(z, z')$ for all $z, z' \in Z$. With every complex structure one can associate the cyclic one-parameter group of symplectic transformations $\{e^{i\phi J}; \phi \in [0, 2\pi]\}$, called the \textit{gauge group}. Hence, by the Stone–von Neumann uniqueness theorem the gauge group in $Z$ induces the one-parameter unitary group $\{U_{\phi}; \phi \in [0, 2\pi]\}$ of \textit{gauge transformations} on $\mathcal{H}$ according to the formula

$$
W(e^{i\phi J} z) = U_{\phi} W(z) U_{\phi}.
$$

(35)

For future use it will be convenient to introduce a complex parametrization of the Weyl operators by defining the \textit{displacement operators}

$$
D(z) = W(J z), \quad z \in \mathcal{Z}.
$$

(36)
A state $\rho$ is gauge-invariant if $\rho = U\varphi\rho U^*\varphi$ for all $\varphi$, which is equivalent to the property $\phi(z) = \phi(e^{iJ}z)$ of the characteristic function. In particular, the Gaussian state (27) is gauge invariant if $m(z) \equiv 0$ and $\alpha(z, z') = \alpha(Jz, Jz')$. Introducing the Hermitian inner product

$$\alpha(z, z') = \frac{1}{2}[\alpha(z, z') - i\alpha(z, Jz')]$$

in $\mathbb{Z}$, we have $\alpha(z, z) = \alpha(z, z)/2$ since $\alpha(z, Jz')$ is skew-symmetric. Moreover, the condition (28) is equivalent to non-negative definiteness of the Hermitian form $\alpha(z, z') - j(z, z')/2$ on $\mathbb{Z}$:

$$\alpha \geq \frac{1}{2}j.$$

This follows from the application of Lemma 7 below to the form

$$z, z' \rightarrow \beta(z, z') = \alpha(z, z') - \frac{i}{2}\Delta(z, z').$$

The relation (37) can be regarded as the inequality for the matrices of the forms if a basis is chosen in $\mathbb{Z}$. In an orthonormal basis, $j = I$ is the identity matrix.

**Lemma 7.** Let $\beta(z, z')$ be a bilinear complex-valued Hermitian form on a real vector space $\mathbb{Z}$ satisfying the condition $\beta(Jz, Jz') = \beta(z, z')$, where $J$ is a linear operator such that $J^2 = -I$. Then $\beta(z, z')$ is non-negative definite, that is,

$$\sum_{j,k} \overline{c_j} c_k \beta(z_j, z_k) \geq 0$$

for any finite collection $\{z_j\} \subset \mathbb{Z}$ and any numbers $\{c_j\} \subset \mathbb{C}$ if and only if

$$\text{Re} \beta(z, z) \pm \text{Im} \beta(z, Jz) \geq 0 \quad \text{for all } z \in \mathbb{Z}.$$  

(38)

(39)

**Proof.** (39) $\Rightarrow$ (38). We have $\beta(z, z') = \text{Re} \beta(z, z') + i \text{Im} \beta(z, z')$, where $\text{Im} \beta(z, z')$ is skew-symmetric, and hence $\text{Im} \beta(z, z) = 0$. Using the fact that $\beta(Jz, z') = -\beta(z, Jz')$, we get that $\text{Re} \beta(z, Jz')$ is also skew-symmetric, so $\text{Re} \beta(z, Jz) = 0$. Thus,

$$\text{Re} \beta(z, z) \pm \text{Im} \beta(z, Jz) = \beta(z, z) \mp i\beta(z, Jz).$$

Now we introduce the complexification $z \leftrightarrow z$ by letting $Jz \leftrightarrow iz$, and we define two Hermitian forms on the complexification $\mathbb{Z}$ of $\mathbb{Z}$:

$$\beta^\pm(z, z') = \beta(z, z') \mp i\beta(z, Jz').$$

Then $\beta^-$ is sesquilinear (that is, complex linear with respect to $z'$ and anti-linear with respect to $z$), while $\beta^+$ is antisesquilinear. From (40) and (39),

$$\beta^\pm(z, z) = \text{Re} \beta(z, z) \pm \text{Im} \beta(z, Jz) \geq 0 \quad \text{for all } z \in \mathbb{Z},$$

hence by (anti-)sesquilinearity

$$\sum_{j,k} \overline{c_j} c_k \beta^\pm(z_j, z_k) \geq 0.$$
Adding the two inequalities corresponding to plus and minus, we get (38).

\[ \Rightarrow (39). \] Applying (38) to \{z_j, Jz_j\} \subset Z and \{c_j, \pm ic_j\} \subset \mathbb{C}, we obtain

\[
\sum_{j,k} c_jc_k [\beta(z_j, z_k) \pm i\beta(z_j, Jz_k)] \geq 0,
\]

and hence the forms (40) are non-negative definite. By (anti-)sesquilinearity of these forms, this is equivalent to (41), that is, to (39). \( \square \)

Assume that on \( Z_A \) and \( Z_B \) the operators of complex structure \( J_A \) and \( J_B \) are fixed, and let \( U^A_\phi \) and \( U^B_\phi \) be the corresponding gauge transformations acting in \( \mathcal{H}_A \) and \( \mathcal{H}_B \) according to (35). A channel \( \Phi: A \rightarrow B \) is said to be \textit{gauge-covariant} if

\[
\Phi[U^A_\phi \rho(U^A_\phi)^*] = U^B_\phi \Phi[\rho](U^B_\phi)^*
\]

for all input states \( \rho \) and all \( \phi \in [0, 2\pi] \). For the Gaussian channel (32) with parameters \((K, l, \mu)\) this reduces to

\[
l(z) \equiv 0, \quad KJ_B - J_A K = 0, \quad \mu(z, z') = \mu(J_B z, J_B z').
\]

The relation (32) for gauge-covariant Gaussian channels takes the form (cf. [30])

\[
\Phi^*[D_B(z)] = D_A(Kz) \exp\{-\mu(z, z)\}, \quad z \in Z_B,
\]

where

\[
\mu \geq \pm \frac{1}{2}[J_B - K^*J_A K].
\]

The equivalence of (44) and (33) is obtained by applying Lemma 7 to the Hermitian form

\[
\beta(z, z') = \mu(z, z') - i \frac{1}{2}[\Delta_B(z, z') - \Delta_A(Kz, Kz')].
\]

A channel \( \Phi: A \rightarrow B \) is said to be \textit{gauge-contravariant} if

\[
\Phi[U^A_\phi \rho(U^A_\phi)^*] = (U^B_\phi)^* \Phi[\rho]U^B_\phi
\]

for all input states \( \rho \) and all \( \phi \in [0, 2\pi] \). For the Gaussian channel (32) with parameters \((K, l, \mu)\) this reduces to

\[
l(z) \equiv 0, \quad KJ_B + J_A K = 0, \quad \mu(z, z') = \mu(J_B z, J_B z').
\]

The relation (32) for a gauge-contravariant Gaussian channel takes the form

\[
\Phi^*[D_B(z)] = D_A(-\Lambda Kz) \exp\{-\mu(z, z)\}, \quad z \in Z_B,
\]

where \( \Lambda \) is the antilinear operator of complex conjugation: \( \Lambda^2 = I, \Lambda^* = -\Lambda \) on \( \mathbb{Z}_A \), so that \( \Lambda J_A + J_A \Lambda = 0 \), and \( K = -\Lambda K \) is a complex-linear operator from \( \mathbb{Z}_B \) to \( \mathbb{Z}_A \). Here

\[
\mu \geq \pm \frac{1}{2}[J_B + K^*J_A K].
\]

The last condition is obtained by applying Lemma 7 to the Hermitian form

\[
\beta(z, z') = \mu(z, z') - i \frac{1}{2}[\Delta_B(z, z') - \Delta_A(Kz, Kz')] = \frac{i}{2}[\Delta_B(z, z') + \Delta_A(Kz, Kz')].
\]
3.3. Attenuators and amplifiers. In what follows we restrict ourselves to Gaussian channels that are gauge-covariant or contravariant with respect to fixed complex structures. Therefore, to be specific, we regard vectors in $\mathbf{Z}$ as $s$-dimensional complex column vectors, where the operator $J$ acts as multiplication by $i$, the corresponding Hermitian inner product is $\langle z, z' \rangle = z^* z'$, and the symplectic form is $\Delta(z, z') = 2 \text{Im} z^* z'$, where $*$ denotes Hermitian conjugation. The linear operators on $\mathbf{Z}$ commuting with $J$ are represented by complex $s \times s$ matrices. The gauge group acts in $\mathbf{Z}$ as multiplication by $e^{i\phi}$. Gaussian gauge-invariant states are described by the modified characteristic functions

$$
\phi(z) = \text{Tr} \rho D(z) = \exp\{-z^* \alpha z\},
$$

where $\alpha$ is a Hermitian correlation matrix such that $\alpha \succeq I/2$, as follows from (37). For the given complex structure, the unique minimal solution of the last inequality is $I/2$, to which there correspond the vacuum state $\rho_0$ and the family $\{\rho_z; z \in \mathbf{Z}\}$ of coherent states with $\rho_z = D(z)\rho_0D(z)^*$. We have

$$
\text{Tr} \rho_w D(z) = \exp\left\{2i \text{Im} w^* z - \frac{1}{2} |z|^2 \right\},
$$

where $|z|^2 = z^* z$.

Let $\mathbf{Z}_A$ and $\mathbf{Z}_B$ be input and output spaces of dimensions $s_A$ and $s_B$. We denote by $s_A = \dim \mathbf{Z}_A$ and $s_B = \dim \mathbf{Z}_B$ the numbers of modes of the input and output of a channel. The action of a Gaussian gauge-covariant channel (43) can be described as

$$
\Phi^*[D_B(z)] = D_A(Kz) \exp\{-z^* \mu z\}, \quad z \in \mathbf{Z}_B,
$$

where $K$ is a complex $s_B \times s_A$ matrix and $\mu$ a Hermitian $s_B \times s_B$ matrix that satisfy the condition

$$
\mu \succeq \pm \frac{1}{2}(I_B - K^* K),
$$

with $I_B$ the $s_B \times s_B$ identity matrix. This follows from (44) by taking into account that the matrix of the form $j(z, \bar{z}')$ in an orthonormal basis is just the identity matrix $I$ of corresponding size. We will need the following lemma.

**Lemma 8.** The map (49) is injective if and only if $7 \ K K^* > 0$ (in which case necessarily $s_B \succeq s_A$).

**Proof.** Injectivity means that $\rho_1 = \rho_2$ if $\Phi[\rho_1] = \Phi[\rho_2]$. But $\Phi[\rho_1] = \Phi[\rho_2]$ is equivalent to $\text{Tr} \rho_1 \Phi^*[D_B(z)] = \text{Tr} \rho_2 \Phi^*[D_B(z)]$, that is, $\text{Tr} \rho_1 D_A(Kz) = \text{Tr} \rho_2 D_A(Kz)$ for all $z \in \mathbf{Z}_B$. By the irreducibility of the Weyl system, this property is equivalent to $\text{Ran} K = \mathbf{Z}_A$, that is, $\text{Ker} K^* = \{0\}$, or $KK^* > 0$. $\square$

The channel (49) is extreme if $\mu$ is a minimal solution of the inequality (50). Special cases of the maps (49) are provided by attenuator and amplifier channels, which are characterized by the matrix $K$ satisfying the inequalities $K^* K \preceq I$ and

\begin{itemize}
  \item For Hermitian matrices $M$ and $N$ the strict inequality $M > N$ means that $M - N$ is positive definite.
\end{itemize}
\(K^*K \geq I\), respectively. We are particularly interested in an extreme attenuator, corresponding to

\[K^*K \leq I_B, \quad \mu = \frac{1}{2}(I_B - K^*K), \quad (51)\]

and an extreme amplifier, corresponding to

\[K^*K \geq I_B, \quad \mu = \frac{1}{2}(K^*K - I_B). \quad (52)\]

A Gaussian gauge-contravariant channel (46) acts according to the formula

\[
\Phi^*[D_B(z)] = D_A(-\overline{K}z) \exp\{-z^*\mu z\}, \quad (53)
\]

where \(\mu\) is a Hermitian matrix satisfying the inequality

\[\mu \geq \frac{1}{2}(I_B + K^*K), \quad (54)\]

which follows from (47). Here \(z\) is the column vector consisting of complex conjugates of the components of \(z\). These maps are extreme if

\[
\mu = \frac{1}{2}(I_B + K^*K). \quad (55)
\]

The following proposition generalizes to the multimode case the useful and important decomposition of one-mode channels (see [20] and [13]).

**Proposition 9.** An arbitrary Gaussian gauge-covariant channel \(\Phi: A \to B\) is a composition \(\Phi = \Phi_2 \circ \Phi_1\) of an extreme attenuator \(\Phi_1: A \to B\) and an extreme amplifier \(\Phi_2: B \to B\).

An arbitrary Gaussian gauge-contravariant channel \(\Phi: A \to B\) is a composition of an extreme attenuator \(\Phi_1: A \to B\) and an extreme gauge-contravariant channel \(\Phi_2: B \to B\).

**Proof.** The composition \(\Phi = \Phi_2 \circ \Phi_1\) of Gaussian gauge-covariant channels \(\Phi_1\) and \(\Phi_2\) obeys the rules

\[K = K_1K_2, \quad (56)\]

\[\mu = K_2^*\mu_1K_2 + \mu_2. \quad (57)\]

Inserting the equalities

\[
\mu_1 = \frac{1}{2}(I_B - K_1^*K_1) = \frac{1}{2}(I_B - |K_1|^2),
\]

\[
\mu_2 = \frac{1}{2}(K_2^*K_2 - I_B) = \frac{1}{2}(|K_2|^2 - I_B)
\]

into (57) and using (56), we obtain

\[|K_2|^2 = K_2^*K_2 = \mu + \frac{1}{2}(K^*K + I_B) \geq \begin{cases} I_B \\ K^*K \end{cases} \quad (58)\]
from the inequality (50). Using the operator monotonicity of the square root, we have

$$|K_2| \geq I_B, \quad |K_2| \geq |K|.$$  

The first inequality in (58) implies that if we choose

$$K_2 = |K_2| = \sqrt{\mu + \frac{1}{2}(K^*K + I_B)}$$  \hspace{1cm} (59)

and the corresponding matrix $\mu_2 = (|K_2|^2 - I_B)/2$, then we obtain an extreme amplifier $\Phi_2: B \to B$.

Then with

$$K_1 = K|K_2|^{-1}$$  \hspace{1cm} (60)

we get from the second inequality in (58) and Lemma 10 (below) that

$$K_1K_1^* = K|K_2|^{-2}K^* = K\left[\mu + \frac{1}{2}(K^*K + I)\right]^{-1}K^* \leq I_A,$$  \hspace{1cm} (61)

which implies that $K_1^*K_1 \leq I_A$, so that $K_1$ together with the corresponding matrix $\mu_1 = (I_B - K_1^*K_1)/2$ give an extreme attenuator.

**Lemma 10.** If $M \geq K^*K$, then $KM^{-1}K^* \leq I_A$, where $-$ denotes the (generalized) inverse matrix.

**Proof.** By the definition of the generalized inverse matrix,

$$u^*M^{-1}u = \sup_{v: v \in \text{Ran } M, \ v \neq 0} \frac{|u^*v|^2}{v^*Mv}.$$  \hspace{1cm}

Substituting $K^*u$ in place of $u$ and using the Cauchy–Schwarz inequality in the numerator of the fraction, we get that

$$u^*KM^{-1}K^*u \leq \sup_{v: v \in \text{Ran } M, \ v \neq 0} \frac{u^*u v^*K^*Kv}{v^*Mv} \leq u^*u. \hspace{1cm} \blacksquare$$

In the case of a contravariant channel the relations (56) and (57) are replaced by

$$K = K_1K_2,$$  \hspace{1cm} (62)
$$\mu = K_2^*\mu_1K_2 + \mu_2.$$  \hspace{1cm} (63)

Substituting

$$\mu_1 = \frac{1}{2}(I - K_1^*K_1), \quad \mu_2 = \frac{1}{2}(K_2^*K_2 + I_B)$$

into (63) and using (54), we obtain

$$|K_2|^2 = K_2^*K_2 = \mu + \frac{1}{2}(K^*K - I_B) \geq K^*K.$$  \hspace{1cm} (64)

Taking $K_2 = |K_2|$ and $\mu_2 = (|K_2|^2 + I_B)/2$ gives us an extreme gauge-contravariant channel $\Phi_2: B \to B$. With

$$K_1 = K|K_2|^{-1}$$  \hspace{1cm} (65)
we get from Lemma 10 that
\[ K_1^* K_1^* = K(|K_2|^2)^2 K^* \]
\[ = K \left[ \mu + \frac{1}{2}(K^*K - I_B) \right]^{-1} K^* \leq I_A, \]  
(66)

which implies that $K_1^* K_1^* \leq I_A$, so $K_1$ with the corresponding matrix $\mu_1$ gives an extreme attenuator $\Phi_1 : A \rightarrow B$, completing the proof of the proposition. □

**Remark 11.** In the case of a gauge-covariant channel, the equality in (61) shows that $KK^* > 0$ implies that $K_1^* K_1^* > 0$, while the inequality $\mu > (K^* K - I_B)/2$ implies that $K_1 K_1^* < I_A$. In the case of a gauge-contravariant channel, the inequality $\mu > (I_B + K^* K)/2$ implies that $0 < K_1 K_1^* < I_A$ in view of (66).

**Proposition 12.** The extreme attenuator with matrix $K$ and the extreme attenuator with matrix $\tilde{K} = \sqrt{I_A - KK^*}$ are mutually complementary.

The extreme amplifier with matrix $K$ and the extreme gauge-contravariant channel with matrix $\tilde{K} = \sqrt{KK^* - I_A}$ are mutually complementary.

**Proof.** For the case of one mode see [13] or [38], §12.6.1. We sketch the multimode proof below. Define $Z_E \simeq Z_A$ and $Z_D \simeq Z_B$, so that $Z = Z_A \oplus Z_D \simeq Z_B \oplus Z_E$.

In the case of an attenuator we consider on $Z$ the block unitary matrix
\[ V = \begin{bmatrix} K & \sqrt{I_B - K^*K} \\ \sqrt{I_B - K^*K} & -K^* \end{bmatrix}, \]  
(67)

which defines a unitary dynamics $U$ on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_D \simeq \mathcal{H}_B \otimes \mathcal{H}_E$ by the relation $U^* D_{BE}(z_{BE})U = D_{AD}(Vz_{BE})$. Here $z_{BE} = [z_B \ z_E]^t$, $D_{BE}(z_{BE}) = D_B(z_B) \otimes D_E(z_E)$, and the unitarity follows from the relation
\[ K \sqrt{I_B - K^*K} = \sqrt{I_A - KK^*}. \]  
(68)

Let $\rho_D = \rho_0$ be the vacuum state and $\rho_A = \rho$ an arbitrary state. Then the formulae (7) and (8) define mutually complementary extreme attenuators as described in the first assertion. The proof is obtained by computing the characteristic function of the output states of the channels. For the state of the composite system $\rho_{BE} = U(\rho \otimes \rho_D)U^*$ we have
\[ \phi_{BE}(z_{BE}) = \operatorname{Tr} U(\rho \otimes \rho_D)U^*[D_B(z_B) \otimes D_E(z_E)] \]
\[ = \operatorname{Tr}(\rho \otimes \rho_D)U^*[D_B(z_B) \otimes D_E(z_E)]U \]
\[ = \operatorname{Tr}(\rho \otimes \rho_D)[D_A(Kz_B + \tilde{K}z_E) \otimes D_D(\sqrt{I_B - K^*K}z_B - K^*z_E)] \]
\[ = \phi_A(Kz_B + \tilde{K}z_E) \exp \left\{ -\frac{1}{2} \sqrt{I_B - K^*K}z_B - K^*z_E \right\}^2. \]  
(69)

Setting $z_E = 0$ or $z_B = 0$, we get that
\[ \phi_B(z_B) = \phi_A(Kz_B) \exp \left\{ -\frac{1}{2}z_B^*(I_B - K^*K)z_B \right\}, \]
\[ \phi_E(z_E) = \phi_A(\tilde{K}z_E) \exp \left\{ -\frac{1}{2}z_E^*KK^*z_E \right\}, \]

as required.
In the case of an amplifier, let us consider the block operator
\[ V = \begin{bmatrix} \frac{K}{\Lambda \sqrt{K^*K - I_B}} & -\sqrt{KK^* - I_A}\Lambda \\ -\Lambda \sqrt{K^*K - I_B} & \frac{\Lambda K^*\Lambda}{\Lambda K^*\Lambda} \end{bmatrix}, \]
where \( \Lambda \) is the operator of complex conjugation, which anticommutes with multiplication by \( i \). Using the property \( \Delta(\Lambda z, \Lambda z') = -\Delta(z, z') \), we get that \( V \) corresponds to a symplectic transformation on \( \mathbb{Z} \) generating the unitary dynamics \( U \) on \( \mathcal{H} \). Again let \( \rho_0 \) be the vacuum state of the environment. Then the formulae (7) and (8) define mutually complementary channels as described in the second assertion of the proposition, and the rest of the proof is similar. \( \square \)

Below we will also need the following lemma.

**Lemma 13.** Let \( \Phi_1 : A \to B \) be an extreme attenuator with \( 0 < K_1K_1^* < I_A \). Then \( \Phi_1[P\psi] = P\psi' \) (a pure state) if and only if \( P\psi \) is a coherent state.

**Proof.** According to Proposition 12, the complementary channel \( \Phi_1 \) is an extreme attenuator with the matrix \( eK = pI_A - K_1K_1^* \), so that \( 0 < eK < I_A \). Its output is also pure, \( \Phi_1[P\psi] = P\psi' \), since the output states of complementary channels have identical non-zero spectra by Lemma 3. Thus,
\[ U(\psi \otimes \psi_0) = \psi' \otimes \psi_E', \]
where \( \psi_0 \in \mathcal{H}_D \) is the vacuum vector, and \( U \) is the unitary operator on \( \mathcal{H} \) implementing the symplectic transformation corresponding to the unitary operator (67) on \( \mathbb{Z}_A \oplus \mathbb{Z}_D \cong \mathbb{Z}_B \oplus \mathbb{Z}_E \), with \( \mathbb{Z}_D \cong \mathbb{Z}_B \) and \( \mathbb{Z}_E \cong \mathbb{Z}_A \). Using the notation
\[ \phi(z) = \text{Tr} P\psi D_A(z), \quad \phi'(z_B) = \text{Tr} P\psi' D_B(z_B), \quad \phi_E(z_E) = \text{Tr} P\psi_E D_E(z_E) \]
for the quantum characteristic functions, we have in view of (69) the functional equation
\[ \phi'(z_B)\phi_E(z_E) = \phi(K_1z_B + \tilde{K}z_E)\exp\left\{-\frac{1}{2}\sqrt{I_B - K_1^*K_1}z_B - K_1^*z_E \right\}^2. \] (70)

Letting \( z_E = 0 \) and \( z_B = 0 \), respectively, we find that
\[ \phi'(z_B) = \phi(K_1z_B)\exp\left\{-\frac{1}{2}\sqrt{I_B - K_1^*K_1}z_B \right\}^2, \]
\[ \phi_E(z_E) = \phi(\tilde{K}z_E)\exp\left\{-\frac{1}{2}|K_1^*z_E |^2\right\}, \]
and thus after the change of variables \( z = K_1z_B, z' = \tilde{K}z_E \) and use of (68) we reduce the equation (70) to
\[ \phi(z)\phi(z') = \phi(z + z')\exp\{\text{Re}z^*z'\}. \]

The hypothesis of the lemma ensures that \( \text{Ran}K_1 = \text{Ran}\tilde{K} = \mathbb{Z}_A \). Substituting \( \omega(z) = \phi(z)\exp\{|z|^2/2\} \), we find that
\[ \omega(z)\omega(z') = \omega(z + z') \] (71)
for all $z, z' \in \mathbb{Z}_A$. The function $\omega(z)$ along with the characteristic function $\phi(z)$ is continuous and satisfies the condition $\omega(-z) = \bar{\omega}(z)$. The only solution of (71) satisfying these conditions is the exponential $\omega(z) = \exp\{i \operatorname{Im} w^* z\}$ for some complex number $w$. Thus,

$$\phi(z) = \exp\left\{ i \operatorname{Im} w^* z - \frac{1}{2} |z|^2 \right\}$$

is the characteristic function of the coherent state $\rho_{w/2}$. □

### 3.4. Gaussian optimizers.

The following basic result for one mode was obtained in [53]. Here we present a complete proof in the multimode case, a sketch of which was given in [23].

**Theorem 14.** (i) Let $\Phi$ be a Gaussian gauge-covariant or contravariant channel and let $f$ be a real concave function on $[0, 1]$ with $f(0) = 0$. Then

$$\operatorname{Tr} f(\Phi[\rho]) \geq \operatorname{Tr} f(\Phi[\rho_{w}]) = \operatorname{Tr} f(\Phi[\rho_0])$$

(72)

for all states $\rho$ and any coherent state $\rho_w$ (the value on the right-hand side is the same for all coherent states by the unitary covariance property of a Gaussian channel (34)).

(ii) If $f$ is strictly concave, then equality holds in (72) only if $\rho$ is a coherent state and one of the following conditions holds:

(a) $s_A = s_B$ and $\Phi$ is an extreme amplifier with

$$\mu = \frac{1}{2} (K^* K - I_B) > 0;$$

(b) $s_B \geq s_A$, $\Phi$ is gauge-covariant with $K K^* > 0$, and

$$\mu > \frac{1}{2} (K^* K - I_B);$$

(c) $s_B \geq s_A$, $\Phi$ is gauge-contravariant with $K K^* > 0$, and

$$\mu > \frac{1}{2} (I_B + K^* K).$$

**Proof.** (i) We start with the first assertion for strictly concave $f$. In this case the inequality (72) for arbitrary concave $f$ follows by monotone approximation $f(x) = \lim_{\varepsilon \downarrow 0} f_\varepsilon(x)$ by the strictly concave functions $f_\varepsilon(x) = f(x) - \varepsilon x^2$. Also, by concavity it is sufficient to prove (72) just for $\rho = P_\psi$.

By Proposition 9, $\Phi = \Phi_2 \circ \Phi_1$, where $\Phi_1$ is an extreme attenuator and $\Phi_2 : B \to B$ is either an extreme amplifier or an extreme gauge-contravariant channel. Any extreme attenuator maps the vacuum state into itself. Indeed,

$$\operatorname{Tr} \Phi_1[\rho_0] D_B(z) = \operatorname{Tr} \rho_0 \Phi_1^*[D_B(z)] = \operatorname{Tr} \rho_0 D_A(Kz) \exp \left\{ -\frac{1}{2} z^*(I_B - K^* K)z \right\}$$

$$= \exp \left\{ -\frac{1}{2} |z|^2 \right\} = \operatorname{Tr} \rho_0 D_B(z).$$
Therefore, $\text{Tr} f(\Phi[\rho_0]) = \text{Tr} f(\Phi_2[\rho_0])$. Then it is sufficient to prove (72) for all extreme amplifiers and all extreme gauge-contravariant channels $\Phi_2$. Indeed, assume that we have proved that

$$\text{Tr} f(\Phi_2[P_\psi]) \geq \text{Tr} f(\Phi_2[\rho_0])$$

(74)

for any state vector $\psi$. Consider the spectral resolution

$$\Phi_1[P_\psi] = \sum_j p_j P_{\phi_j}, \quad \text{where } p_j > 0.$$

Then

$$\text{Tr} f(\Phi[P_\psi]) = \text{Tr} f(\Phi_2[\Phi_1[P_\psi]])$$

$$\geq \sum_j p_j \text{Tr} f(\Phi_2[P_{\phi_j}])$$

$$\geq \text{Tr} f(\Phi_2[\rho_0])$$

$$= \text{Tr} f(\Phi_2[\Phi_1[\rho_0]]) = \text{Tr} f(\Phi[\rho_0]).$$

(75) (76) (77) (78)

According to the second assertion of Proposition 12 and Lemma 3,

$$\text{Tr} f(\Phi_2[P_\psi]) = \text{Tr} f(\tilde{\Phi}_2[P_\psi]),$$

where $\Phi_2$ is an extreme amplifier and $\tilde{\Phi}_2$ is an extreme gauge-contravariant channel. Thus, it is sufficient to prove (74) only for an extreme amplifier $\Phi_2: B \to B$ with Hermitian matrix $K_2 \geq I_B$.

The following result is based on a key observation by Giovannetti.

**Lemma 15.** For an extreme amplifier $\Phi_2: B \to B$ with matrix $K_2 \geq I_B$ there is an extreme attenuator $\Phi_1'$ such that for all $\psi \in \mathcal{H}_B$

$$\Phi_2(P_\psi) \sim (\Phi_2 \circ \Phi_1')(P_\psi).$$

(79)

**Proof.** By Proposition 12 and Lemma 3, $\Phi_2(P_\psi) \sim \tilde{\Phi}_2(P_\psi)$ for all $\psi \in \mathcal{H}_B$, where $\tilde{\Phi}_2$ is the extreme contravariant channel with the matrix $\tilde{K} = \sqrt{K_2^2 - I_B}$.

We define the transposition map $\mathcal{T}: B \to B$ by

$$\mathcal{T}[D(z)] = D(-z).$$

The composition $\Phi = \mathcal{T} \circ \tilde{\Phi}_2$ is a covariant Gaussian channel:

$$\Phi^*[D(z)] = \tilde{\Phi}_2^* \circ \mathcal{T}[D(z)] = D\left(\sqrt{K_2^2 - I_B} z\right) \exp\left\{-\frac{1}{2} z^*K_2^2 z\right\}.$$
Applying the decomposition in Proposition 9, we get that \( \Phi = \Phi_2 \circ \Phi'_1 \), where \( \Phi_2 \) is the original amplifier and \( \Phi'_1 : B \to B \) is another extreme attenuator with matrix \( K_1 = \sqrt{I_B - K_2^{-2}} \). This implies the relation (79). □

Lemmas 15 and 5 give us that
\[
\text{Tr} f(\Phi_2(P_\psi)) = \text{Tr} f((\Phi_2 \circ \Phi'_1)(P_\psi)).
\] (80)

Again consider the spectral resolution of the density operator
\[
\Phi'_1(P_\psi) = \sum_j p'_j P_{\psi_j}, \quad p'_j > 0.
\]

By concavity,
\[
\text{Tr} f((\Phi_2 \circ \Phi'_1)[P_\psi]) \geq \sum_j p'_j \text{Tr} f(\Phi_2[P_\psi]).
\] (81)

Since \( f \) is assumed to be strictly concave, the functional \( \rho \to \text{Tr} f(\Phi_2[\rho]) \) is strictly concave \([11]\). Assuming that \( P_{\psi} \) is a minimizer for the functional (80), we conclude that all the operators \( \Phi_2[P_{\psi}] \) must coincide, since otherwise the above inequality would be strict, contradicting the assumption. From Lemma 8 it follows that \( P_{\psi_j} = P_{\psi} \) for all \( j \) and for some \( \psi' \in \mathcal{H}_B \), and hence the output \( \Phi_1[P_{\psi}] = P_{\psi'} \) is a pure state if \( P_{\psi} \) is a minimizer.

Since \( K_1 = \sqrt{I_B - K_2^{-2}} \), the assumptions of Lemma 13 are satisfied if \( K_2 > I_B \). In this case if \( P_{\psi} \) is a minimizer, then the lemma implies that \( P_{\psi} \) is a coherent state. Thus, we obtain the inequality (74) for the amplifier \( \Phi_2 \) with \( K_2 > I_B \) and strictly concave \( f \). In this way we also obtain the case a) of the ‘only if’ statement (ii).

For an amplifier \( \Phi_2 \) with matrix \( K_2 \geq I_B \), we can take any sequence \( K_2^{(n)} \to I_B \) with \( K_2^{(n)} \to K_2 \), and the corresponding amplifiers \( \Phi_2^{(n)} \). Then \( \text{Tr} f(\Phi_2^{(n)}[\rho]) \to \text{Tr} f(\Phi_2[\rho]) \) for any concave polygonal function \( f \) on \([0,1] \) with \( f(0) = 0 \) and any \( \rho \in \mathcal{S}(\mathcal{H}_A) \). This follows from the fact that any such function is a Lipschitz function, \( |f(x) - f(y)| \leq \kappa |x - y| \), and \( ||\Phi_2^{(n)}[\rho] - \Phi_2[\rho]||_1 \to 0 \). It follows that (74) holds for all extreme amplifiers \( \Phi_2 \) in the case of concave polygonal functions \( f \). For arbitrary concave functions \( f \) on \([0,1] \) there is a monotonically non-decreasing sequence of concave polygonal functions \( f_m \) converging to \( f \) pointwise. Passing to the limit as \( m \to \infty \) gives the inequality (74) for an arbitrary extreme amplifier, and hence (72) holds for any Gaussian gauge-covariant or contravariant channel.

(ii) The ‘only if’ statements in the cases b) and c) are obtained from the decomposition \( \Phi = \Phi_2 \circ \Phi_1 \) and the relations (75)–(77) by arguments similar to the case of an extreme amplifier. We note that the conditions imposed on the channel \( \Phi \) imply that in the decomposition \( \Phi = \Phi_2 \circ \Phi_1 \) the attenuator \( \Phi_1 \) is defined by a matrix \( K_1 \) such that \( 0 < K_1 K_1^* < I_A \) (see Remark 11). Applying the argument involving the relations (80) and (81) with strictly concave \( f \) to the relations (75)–(78), we get that for any pure minimizer \( P_{\psi} \) of \( \text{Tr} f(\Phi[P_{\psi}]) \) the output of the extreme attenuator \( \Phi_1[P_{\psi}] \) is necessarily a pure state. Applying Lemma 13 to the attenuator \( \Phi_1 \), we conclude that \( P_{\psi} \) is necessarily a coherent state. □
3.5. Explicit formulae and additivity.

**Proposition 16.** For any $p > 1$ and any Gaussian gauge-covariant or contravariant channel $\Phi$

\[
\|\Phi\|_{1 \to p} = \left( \text{Tr} \Phi[\rho_0]^p \right)^{1/p}, \tag{82}
\]

\[
\tilde{R}_p(\Phi) = R_p(\Phi[\rho_0]), \tag{83}
\]

\[
\tilde{H}(\Phi) = H(\Phi[\rho_0]), \tag{84}
\]

where $\rho_0$ is the vacuum state.

The multiplicativity property (14) holds for any two Gaussian gauge-covariant (contravariant) channels $\Phi_1$ and $\Phi_2$, as well as the additivity property for the minimal Rényi entropy (15) and for the minimal von Neumann entropy (16).

**Proof.** The first statement follows from Theorem 14 by taking $f(x) = -x^p$ and $f(x) = -x \log x$.

If $\Phi_1$ and $\Phi_2$ are both gauge-covariant (contravariant), then their tensor product $\Phi_1 \otimes \Phi_2$ shares this property. The second statement then follows from the expressions (82)–(84) and the product property of the vacuum state $\rho_0 = \rho_0^{(1)} \otimes \rho_0^{(2)}$, which follows from the definition. □

From the definitions of gauge-co/contravariant channels (49) and (53) it follows that $\Phi[\rho_0]$ is a gauge-invariant Gaussian state with the correlation matrix $\mu + K^*K/2$. The spectrum of $\Phi[\rho_0]$ can be computed explicitly, leading to the expressions [41]

\[
\|\Phi\|_{1 \to p} = \left[ \det \left[ \left( \mu + \frac{1}{2} K^* K + \frac{1}{2} I_B \right)^p - \left( \mu + \frac{1}{2} K^* K - \frac{1}{2} I_B \right)^p \right] \right]^{-1/p}
\]

and

\[
\tilde{H}(\Phi) = \text{tr} g \left( \mu + \frac{1}{2} (K^* K - I_B) \right), \tag{85}
\]

where $g(x) = (x+1) \log(x+1) - x \log x$ and tr denotes the trace of operators on $Z$.

In the last case we used the formula for the entropy of a Gaussian state (48) [40]:

\[
H(\rho) = \text{tr} g \left( \alpha - \frac{1}{2} I \right).
\]

We now turn to the classical capacity of a channel $\Phi$. In infinite dimensions, there are two novel features as compared to the situation described in §2.4. First, one has to extend the notion of ensemble to embrace continual families of states. We call an arbitrary Borel probability measure $\pi$ on $\mathcal{S}(\mathcal{H}_A)$ a generalized ensemble. The **average state** of the generalized ensemble $\pi$ is defined as the barycenter of the probability measure:

\[
\bar{\rho}_\pi = \int_{\mathcal{S}(\mathcal{H}_A)} \rho \pi(d\rho).
\]

The usual ensembles correspond to finitely supported measures.

Second, one has to consider input constraints to avoid infinite values of the capacities. Let $F$ be a positive selfadjoint operator acting in $\mathcal{H}_A$ which usually
represents the energy of the system $A$. We consider input states with constrained energy: $\text{Tr } \rho F \leq E$, where $E$ is a fixed positive constant. Since the operator $F$ is usually unbounded, care should be taken in defining the trace. We put $\text{Tr } \rho F^\dagger = \int_0^\infty \lambda \, dm_\rho(\lambda)$, where $m_\rho(\lambda) = \text{Tr } \rho E(\lambda)$ and $E(\lambda)$ is the spectral function of the selfadjoint operator $F$. Then the constrained $\chi$-capacity is given by the following generalization of the expression (17):

$$C_\chi(\Phi, F, E) = \sup_{\pi : \text{Tr } \rho_\pi F \leq E} \chi(\pi),$$  \hspace{1cm} (86)

where

$$\chi(\pi) = H(\Phi[\rho]) - \int_{\mathcal{H}_A} H(\Phi[\rho]) \, \pi(d\rho).$$  \hspace{1cm} (87)

To ensure that this expression is defined correctly, we must impose on the channel $\Phi$ and the constraint operator $F$ certain additional conditions (see [38], §11.5) which, however, always hold in the Gaussian case considered below.

Let $F^{(n)} = F \otimes I \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes I \otimes F$. Then the constrained classical capacity is given by the expression

$$C(\Phi, F, E) = \lim_{n \to \infty} \frac{1}{n} C_\chi(\Phi^{\otimes n}, F^{(n)}, nE).$$  \hspace{1cm} (88)

For a Gaussian gauge-covariant channel $\Phi$ we now consider the gauge-invariant oscillator energy operator

$$F = \sum_{j,k=1}^{s_A} \epsilon_{jk} a_j^* a_k,$$

where $\epsilon = [\epsilon_{jk}]$ is a Hermitian positive definite matrix and $a_j = (q_j + ip_j)/\sqrt{2}$ is the annihilation operator for the $j$th mode. For any state $\rho$ satisfying the condition $\text{Tr } \rho F < \infty$, the first moments $\text{Tr } \rho a_j$ and the second moments $\text{Tr } \rho a_j^* a_k$ and $\text{Tr } \rho a_j a_k$ are well defined. For a gauge-invariant state $\text{Tr } \rho a_j = 0$ and $\text{Tr } \rho a_j a_k = 0$. For a Gaussian gauge-invariant state (48) we have

$$\alpha - \frac{1}{2} = [\text{Tr } \rho a_j^* a_k]_{j,k=1,\ldots,s}$$

(see, for instance, [33]).

**Proposition 17.** The constrained classical capacity of a Gaussian gauge-covariant channel $\Phi$ is

$$C(\Phi; F, E) = C_\chi(\Phi; F, E) = \max_{\nu : \text{Tr } \nu \leq E} \text{tr } g \left( K^* \nu K + \mu + \frac{1}{2}(K^* K - I_B) \right) - \text{tr } g \left( \mu + \frac{1}{2}(K^* K - I_B) \right).$$  \hspace{1cm} (89)

The optimal ensemble $\pi$ on which the supremum in (86) is attained consists of the coherent states $\rho_z = D_A(z) \rho_0 D_A(z)^*$, $z \in \mathbb{Z}_A$, distributed with a gauge-invariant Gaussian probability distribution $Q_\nu(d^2z)$ on $\mathbb{Z}_A$ having zero mean and the correlation matrix $\nu$ which solves the maximization problem in (89).
Consider a Gaussian ensemble $\pi_{\nu}$ of coherent states $\rho_z = D_A(z)\rho_0 D_A(z)^*$, $z \in \mathbb{Z}_A$, with a gauge-invariant Gaussian probability distribution $Q_{\nu}(d^2 w)$ on $\mathbb{Z}_A$ having zero mean and some correlation matrix $\nu$. The distribution is determined by the classical characteristic function

$$\int_{\mathbb{Z}_A} \exp\{2i \text{Im} \, w^* z\} Q_{\nu}(d^2 w) = \exp\{-z^* \nu z\}.$$ 

Using the covariance property (34) of a Gaussian channel, we have

$$H(\Phi[\rho_z]) = H(\Phi[D_A(z)\rho_0 D_A(z)^*]) = H(\Phi[\rho_0]) = \text{tr} \left( \mu + \frac{1}{2}(K^*K - I_B) \right),$$

which does not depend on $z$ and hence gives the value of the integral term in (87). Integration of the characteristic functions of coherent states gives

$$\text{Tr} \, \overline{\rho}_{\pi_{\nu}} D_A(z) = \exp\left\{-z^* \left( \nu + \frac{1}{2}I_A \right) z \right\}.$$ 

Then

$$\nu = \left[ \text{Tr} \, \overline{\rho}_{\pi_{\nu}} a_j^* a_k \right]_{j,k=1,...,s_A} \quad \text{and} \quad \text{Tr} \, \overline{\rho}_{\pi_{\nu}} F = \sum_{j,k=1}^s \epsilon_{jk} \text{Tr} \, \overline{\rho}_{\pi_{\nu}} a_j^* a_k = \text{tr} \nu \epsilon.$$ 

The state $\Phi[\overline{\rho}_{\pi_{\nu}}]$ is a gauge-invariant Gaussian state with the correlation matrix $K^*(\nu + I_A/2)K + \mu$, and therefore it has the entropy

$$\text{tr} \left( K^* \nu K + \mu + \frac{1}{2}(K^*K - I_B) \right).$$

Thus, for the Gaussian ensemble $\pi_{\nu}$

$$\chi(\pi_{\nu}) = \text{tr} \left( K^* \nu K + \mu + \frac{1}{2}(K^*K - I_B) \right) - \text{tr} \left( \mu + \frac{1}{2}(K^*K - I_B) \right). \quad (90)$$

Summarizing, we need to show that

$$C(\Phi; F, E) = C\chi(\Phi; F, E) = \sup_{\nu: \text{Tr} \nu \epsilon \leq E} \chi(\pi_{\nu}). \quad (91)$$

We denote by $\mathcal{G}$ the set of Gaussian gauge-invariant states on $\mathcal{H}_A$.

**Lemma 18.**

$$\max_{\rho^{(n)}: \text{Tr} \rho^{(n)} F^{(n)} \leq nE} H(\Phi^{\otimes n}[\rho^{(n)}]) \leq n \max_{\rho: \rho \in \mathcal{G}, \text{Tr} \rho F \leq E} H(\Phi[\rho]). \quad (92)$$

**Proof.** We first prove that

$$\sup_{\rho^{(n)}: \text{Tr} \rho^{(n)} F^{(n)} \leq nE} H(\Phi^{\otimes n}[\rho^{(n)}]) \leq n \sup_{\rho: \text{Tr} \rho F \leq E} H(\Phi[\rho]). \quad (93)$$
Indeed, denoting by $\rho_j$ the partial state of $\rho^{(n)}$ in the $j$th tensor factor of $\mathcal{H}^\otimes_n$ and letting $\overline{\rho} = n^{-1} \sum_{j=1}^n \rho_j$, we have

$$H(\Phi^\otimes n[\rho^{(n)}]) \leq \sum_{j=1}^n H(\Phi[\rho_j]) \leq nH(\Phi[\overline{\rho}]),$$

where in the first inequality we used subadditivity of the quantum entropy, and in the second inequality its concavity. Moreover, $\text{Tr} \overline{\rho} F = n^{-1} \text{Tr} \rho^{(n)} F^{(n)} \leq E$, which implies (93).

Using the gauge covariance of the channel $\Phi$, we can then confine the maximization on the right-hand side of (93) to the gauge-invariant states. Indeed, for a given state $\rho$ satisfying the constraint $\text{Tr} \rho F \leq E$, the averaging

$$\rho_{\text{av}} = \frac{1}{2\pi} \int_0^{2\pi} U_{\varphi} \rho U_{\varphi}^* \, d\varphi$$

of it also satisfies the constraint, while $H(\Phi[\rho]) \leq H(\Phi[\rho_{\text{av}}])$ by the concavity of the entropy.

Finally, we use the maximum entropy principle, which says that among the states with fixed second moments the Gaussian state has maximal entropy (see, for instance, [38], Lemma 12.25). This proves (92). □

We now have

$$\max_{\rho : \rho \in \mathcal{G}, \text{Tr} \rho F \leq E} H(\Phi[\rho]) = \max_{\nu : \text{tr} \nu \in \mathcal{E}} \text{tr} g\left( (K^* \nu K + \mu + \frac{1}{2} (K^* K - I_B)) \right). \quad (94)$$

Let $\nu$ be the solution of the maximization problem on the right-hand side of (94). To prove (89) observe that

$$n\chi(\pi_{\nu}) \leq nC_{\chi}(\Phi, F, E) \leq C_{\chi}(\Phi^\otimes n, F^{(n)}, nE) \leq \max_{\rho^{(n)} : \text{Tr} \rho^{(n)} F^{(n)} \leq nE} H(\Phi^\otimes n[\rho^{(n)}]) - \min_{\rho^{(n)}} H(\Phi^\otimes n[\rho^{(n)}]).$$

Using Lemma 18 and Proposition 16 we see that this is less than or equal to

$$n\left[ \max_{\rho : \rho \in \mathcal{G}, \text{Tr} \rho F \leq E} H(\Phi[\rho]) - H(\Phi[\rho_0]) \right] = n\chi(\pi_{\nu}),$$

where the equality follows from (94) and (90).

Thus, $C_{\chi}(\Phi^\otimes n, F^{(n)}, nE) = nC_{\chi}(\Phi, F, E)$, and hence the constrained classical capacity (88) of the Gaussian gauge-covariant channel is given by the expression (89). □

A similar argument applies to a Gaussian gauge-contravariant channel (53), giving the expression (89) with $\epsilon$ replaced by $\overline{\epsilon}$. Indeed, in this case $\Phi[\overline{p}_{\pi_{\nu}}]$ is a gauge-invariant Gaussian state with the characteristic function

$$\text{Tr} \Phi[\overline{p}_{\pi_{\nu}}] D(z) = \exp\left\{ - (Kz)^* \left( \nu + \frac{1}{2} I_A \right) Kz - z^* \mu z \right\}$$

$$= \exp\left\{ - (Kz)^* \left( \overline{\nu} + \frac{1}{2} I_A \right) Kz - z^* \mu z \right\}$$
and the correlation matrix $K^*(\overline{v} + I_A/2)K + \mu$. On the other hand, $\text{tr} \nu \epsilon = \text{tr} \overline{v} \epsilon$, so that redefining $\overline{v}$ as $\nu$, we get the assertion.

The maximization in (89) is a finite-dimensional optimization problem which is a quantum analogue of the ‘water-filling’ problem in classical information theory (see, for instance, [15], [40]). It can be solved analytically only in some special cases, for example, when $K, \mu, \epsilon$ commute, and it is a subject of separate study.

3.6. The case of a quantum-classical Gaussian channel. In this subsection we follow [23] in giving an alternative proof of the majorization property established in [52] for Glauber’s coherent states.

Consider the affine map which transforms quantum states $\rho \in \mathcal{S}(\mathcal{H})$ into probability densities on $\mathbb{Z}$:

$$\rho \rightarrow p_\rho(z) = \text{Tr} \rho D(z) \rho_0 D(z)^*, \quad (95)$$

where $D(z)$ are the displacement operators and $\rho_0$ is the vacuum state with the quantum characteristic function

$$\phi_0(z) \equiv \text{Tr} \rho_0 D(z) = \exp \left\{-\frac{1}{2}|z|^2 \right\}.$$ 

The function $p_\rho(z)$ is bounded by 1 and is indeed a continuous probability density, and the normalization follows from the resolution of the identity on $\mathcal{H}$:

$$\int_Z D(z) \rho_0 D(z)^* \frac{d^2s_z}{\pi^s} = I_{\mathcal{H}}.$$ 

**Proposition 19.** If $f$ is a concave function on $[0,1]$ with $f(0) = 0$, then for an arbitrary state $\rho$

$$\int_Z f(p_\rho(z)) \frac{d^2s_z}{\pi^s} \geq \int_Z f(p_{\rho_w}(z)) \frac{d^2s_z}{\pi^s}, \quad (96)$$

where $\rho_w$ is an arbitrary coherent state.

**Proof.** For any $c > 0$ consider the channel $\Phi_c$ defined by

$$\Phi_c[\rho] = \int \frac{d^2s_z}{\pi^s c^2} \text{Tr}[\rho D(c^{-1}z) \rho_0 D^*(c^{-1}z)] \rho_z. \quad (97)$$

The map (97) is a Gaussian gauge-covariant channel such that

$$\Phi_c^*[D(z)] = D(cz) \exp \left\{-\frac{c^2}{2} |z|^2 \right\}.$$ 

(cf. [22], §5). Therefore, by Theorem 14,

$$\text{Tr} f(\Phi_c[\rho]) \geq \text{Tr} f(\Phi_c[\rho_w]) \quad (98)$$

for all states $\rho$ and any coherent state $\rho_w$. We will prove Proposition 19 by taking the limit as $c \rightarrow \infty$. 

In the proof we also use the Berezin–Lieb inequalities [9]:

\[ \int_{\mathbb{Z}} f(\rho(z)) \frac{d^{2s}z}{\pi^s} \leq \text{Tr} f(\sigma) \leq \int_{\mathbb{Z}} f(\rho(z)) \frac{d^{2s}z}{\pi^s}, \]  

(99)

valid for any quantum state admitting the representation

\[ \sigma = \int_{\mathbb{Z}} \rho(z) \rho_z \frac{d^{2s}z}{\pi^s} \]

with probability density \( \rho(z) \). On the right-hand side of (99), \( \rho(z) = \text{Tr} \rho_z \). In (99) one has to assume that \( f \) is defined on \([0, \infty)\) (in fact, \( \rho(z) \) can be unbounded). We shall assume this for the moment.

Taking \( \sigma = \Phi_c[\rho] \), we get from (97) that

\[ \rho(z) = \frac{1}{c^{2s}} \text{Tr} \rho D(c^{-1}z) \rho_0 D^*(c^{-1}z) = \frac{1}{c^{2s}} \rho(c^{-1}z), \]

while

\[ \rho(z) = \text{Tr} \rho_z \Phi_c[\rho] = \int_{\mathbb{Z}} \rho(w) \text{Tr} \rho_z \rho_w \frac{d^{2s}w}{\pi^s}. \]  

(100)

We use the well-known formula (see, for instance, [48] and [33])

\[ \text{Tr} \rho_z \rho_w = \exp\{-|z - w|^2\}. \]

Introducing the Gaussian probability density

\[ q_c(z) = \frac{c^{2s}}{\pi^s} \exp\{-c^2|z|^2\}, \]

which tends to the \( \delta \)-function as \( c \to \infty \), and substituting this into (100), we have

\[ \rho(z) = \int d^{2s}w q_c(z - w) \]

\[ = \int d^{2s}w' q_c(z - cw') \]

\[ = \frac{1}{c^{2s}} \rho_q * q_c(c^{-1}z). \]  

(101)

After the change of the integration variable \( c^{-1}z \to z \) the inequalities (99) become

\[ \int_{\mathbb{Z}} f(c^{-2s} \rho(z)) \frac{d^{2s}z}{\pi^s} \leq c^{-2s} \text{Tr} \rho(f(\Phi_c[\rho])) \leq \int_{\mathbb{C}^s} f(c^{-2s} \rho_z * q_c(z)) \frac{d^{2s}z}{\pi^s}. \]

Substituting \( \rho = \rho_w \), we get that

\[ \int_{\mathbb{Z}} f(c^{-2s} \rho_w(z)) \frac{d^{2s}z}{\pi^s} \leq c^{-2s} \text{Tr} f(\Phi_c[\rho_w]) \leq \int_{\mathbb{Z}} f(c^{-2s} \rho_w * q_c(z)) \frac{d^{2s}z}{\pi^s}. \]
Combining the last two displayed formulae with (98), we obtain
\[ \int_{Z} g(p_\rho(z)) \frac{d^{2s}z}{\pi^s} - \int_{Z} g(p_\rho_\omega(z)) \frac{d^{2s}z}{\pi^s} \geq \int_{Z} g(p_\rho(z)) \frac{d^{2s}z}{\pi^s} - \int_{Z} g(p_\rho * q_c(z)) \frac{d^{2s}z}{\pi^s}, \] (102)
where \( g(x) = f(c^{-2s}x) \), which is again a concave function. Moreover, any concave polygonal function \( g \) on \([0, 1] \) with \( g(0) = 0 \) can be obtained in this way by defining
\[ f(x) = \begin{cases} g(c^{2s}x), & x \in [0, c^{-2s}], \\ g(1) + g'(1)(x - c^{-2s}), & x \in [c^{-2s}, \infty), \end{cases} \]
and hence (102) holds for any such function. Then the right-hand side of the inequality (102) tends to zero as \( c \to \infty \). Indeed, \( |g(x) - g(y)| \leq \kappa |x - y| \) for a polygonal function, and the asserted convergence follows from the convergence \( p_\rho * q_c \to p_\rho \) in \( L_1 \): if \( p(z) \) is a bounded continuous probability density, then
\[ \lim_{c \to \infty} \int_{Z} |p * q_c(z) - p(z)| \frac{d^{2s}z}{\pi^s} = 0. \]
Thus, we obtain (96) for concave polygonal functions \( f \). But for an arbitrary continuous concave function \( f \) on \([0, 1] \) there is a monotonically non-decreasing sequence of concave polygonal functions \( f_n \) converging to \( f \). Applying Beppo Levy’s theorem, we obtain the proposition. □

4. Appendix

Consider a gauge-covariant channel \( \Phi \) such that the matrices \( K^*K \) and \( \mu \) commute (in particular, this condition is satisfied by extreme amplifiers and attenuators). These channels are diagonalizable in the following sense. We have
\[ K = V_A K_d V_B, \quad \mu = V_B^* \mu_d V_B, \]
where \( V_A \) and \( V_B \) are unitary matrices and \( K_d \) and \( \mu_d \) are diagonal (rectangular) matrices with non-negative values on the diagonal. Then \( K^*K = V_B^* K_d^2 V_B \) and
\[ \Phi[\rho] = U_B \Phi_d [U_A \rho U_A^*] U_B^*, \] (103)
where \( U_A \) and \( U_B \) are canonical unitary (‘metaplectic’ [2]) transformations acting on \( \mathcal{H}_A \) and \( \mathcal{H}_B \) such that
\[ U_B^* D_B(z) U_B = D_B(V_B z), \quad U_A^* D_A(z) U_A = D_A(V_A z). \]
To describe the action of the ‘diagonal’ channel \( \Phi_d \) in more detail, we have to consider separately the cases \( s_A = s_B, s_A < s_B, \) and \( s_A > s_B \).
In the case \( s_A = s_B \) we have
\[ K_d = \text{diag}[k_j]_{j=1,...,s_B}, \quad \mu_d = \text{diag}[\mu_j]_{j=1,...,s_B}. \]
Then $\Phi_d = \bigotimes_{j=1}^{s_B} \Phi_j$, where, in self-explanatory notation,

$$\Phi_j[D_j(z_j)] = D_j(k_jz_j) \exp\{-\mu_j|z_j|^2\}. \quad (104)$$

In the case $s_A < s_B$

$$K_d = \begin{bmatrix} \text{diag}[k_j]_{j=1,\ldots,s_A} \\ 0 \end{bmatrix},$$

where 0 denotes a block of zeros of size $(s_B - s_A) \times s_A$. Then

$$\Phi_d[\rho] = \bigotimes_{j=1}^{s_A} \Phi_j[\rho] \otimes \rho_0^{[s_A+1,\ldots,s_B]},$$

where for $j = 1, \ldots, s_A$ the one-mode channels $\Phi_j$ are given by (104) and $\rho_0^{[s_A+1,\ldots,s_B]}$ is the vacuum state of the modes $s_A + 1, \ldots, s_B$.

In the case $s_A > s_B$

$$K_d = \begin{bmatrix} \text{diag}[k_j]_{j=1,\ldots,s_B} \\ 0 \end{bmatrix},$$

where 0 denotes a block of zeros of size $s_B \times (s_A - s_B)$. Then

$$\Phi_d[\rho] = \left( \bigotimes_{j=1}^{s_A} \Phi_j \right) \left[ \text{Tr}_{s_B+1,\ldots,s_A} \rho \right],$$

where $\text{Tr}_{s_B+1,\ldots,s_A}$ denotes the partial trace over the last $s_A - s_B$ modes of the operator $\rho$.

There is a similar reduction to diagonal form for extreme gauge-contravariant channels.

The author is grateful to M. E. Shirokov (Steklov Mathematical Institute of RAS) for comments and discussions. Thanks are due to David Ding (Stanford University) for pointing out some typos.

Bibliography

[1] Г.Г. Амосов, А.С. Холево, Р.Ф. Вернер, “О гипотезе аддитивности в квантовой теории информации”, Пробл. передачи информ. 36:4 (2000), 25–34; English transl., G. G. Amosov, A. S. Holevo, and R. F. Werner, “On the additivity conjecture in quantum information theory”, Problems Inform. Transmission 36:4 (2000), 305–313.

[2] Arvind, B. Dutta, N. Mukunda, and R. Simon, “The real symplectic groups in quantum mechanics and optics”, Pramana 45:6 (1995), 471–497.

[3] G. Aubrun, S. Szarek, and E. Werner, “Hastings’s additivity counterexample via Dvoretzky’s theorem”, Comm. Math. Phys. 305:1 (2011), 85–97.

[4] K. M. R. Audenaert, “A note on the $p \rightarrow q$ norms of completely positive maps”, Linear Algebra Appl. 430:4 (2009), 1436–1440.

[5] К.И. Бабенко, “Об одном неравенстве в теории интегралов Фурье”, Изв. АН СССР. Сер. матем. 25:4 (1961), 531–542; English transl., K. I. Babenko, “An inequality in the theory of Fourier integrals”, Amer. Math. Soc. Transl. Ser. 2, vol. 44, Amer. Math. Soc., Providence, RI 1965, pp. 115–128.
[6] W. Beckner, “Inequalities in Fourier analysis”, *Ann. of Math. (2)* **102**:1 (1975), 159–182.

[7] S. T. Belinschi, B. Collins, and I. Nechita, *Almost one bit violation for the additivity of the minimum output entropy*, 2013 (v3 – 2014), 24 pp., arXiv:1305.1567.

[8] C. H. Bennett, C. A. Fuchs, and J. A. Smolin, “Entanglement-enhanced classical communication on a noisy quantum channel”, *Quantum communication, computing, and measurement* (O. Hirota, A. S. Holevo, and C. M. Caves, eds.), Plenum, New York 1997, pp. 79–88.

[9] Ф. А. Березин, “Ковариантные и контравариантные символы операторов”, *Изв. АН СССР. Сер. матем.* **36**:5 (1972), 1134–1167; English transl., F. A. Berezin, “Covariant and contravariant symbols of operators”, *Math. USSR-Izv.* **6**:5 (1972), 1117–1151.

[10] F. G. S. L. Brandão and M. Horodecki, “On Hastings’ counterexamples to the minimum output entropy additivity conjecture”, *Open Syst. Inf. Dyn.* **17**:1 (2010), 31–52.

[11] E. Carlen, “Trace inequalities and quantum entropy: an introductory course”, *Entropy and the quantum*, Contemp. Math., vol. 529, Amer. Math. Soc., Providence, RI 2010, pp. 73–140.

[12] E. A. Carlen and E. H. Lieb, “A Minkowski type trace inequality and strong subadditivity of quantum entropy”, *Differential operators and spectral theory*, Amer. Math. Soc. Transl. Ser. 2, vol. 189, Amer. Math. Soc., Providence, RI 1999, pp. 59–68.

[13] F. Caruso, V. Giovannetti, and A. S. Holevo, “One-mode bosonic Gaussian channels: a full weak-degradability classification”, *New J. Phys.* **8** (2006), 310, 18 pp.

[14] F. Caruso, J. Eisert, V. Giovannetti, and A. S. Holevo, “Multi-mode bosonic Gaussian channels”, *New J. Phys.* **10** (2008), 083030, 33 pp.

[15] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley Ser. Telecom., John Wiley & Sons, Inc., New York 1991, xxiv+542 pp.

[16] E. B. Davies, *Quantum theory of open systems*, Academic Press, London–New York 1976, x+171 pp.

[17] B. Demoen, P. Vanheuverzwijn, and A. Verbeure, “Completely positive quasi-free maps on the CCR-algebra”, *Rep. Math. Phys.* **15**:1 (1979), 27–39.

[18] M. Fukuda, C. King, and D. K. Moser, “Comments on Hastings’ additivity counterexamples”, *Comm. Math. Phys.* **296**:1 (2010), 111–143.

[19] M. Fukuda and M. M. Wolf, “Simplifying additivity problems using direct sum constructions”, *J. Math. Phys.* **48**:7 (2007), 072101, 7 pp.

[20] R. García-Patrón, C. Navarrete-Benlloch, S. Lloyd, J. H. Shapiro, and N. J. Cerf, “Majorization theory approach to the Gaussian channel minimum entropy conjecture”, *Phys. Rev. Lett.* **108**:11 (2012), 110505, 5 pp.

[21] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, “Classical capacity of the lossy bosonic channel: the exact solution”, *Phys. Rev. Lett.* **92**:2 (2004), 027902, 4 pp.

[22] V. Giovannetti, A. S. Holevo, and R. García-Patrón, “A solution of Gaussian optimizer conjecture for quantum channels”, *Comm. Math. Phys.* **334**:3 (2015), 1553–1571.

[23] В. Джованнетти, А. С. Холево, А. Мари, “Мажоризация и аддитивность для многомодовых бозонных гауссовских каналов”, *ТМФ* **182**:2 (2015), 338–349; English transl., V. Giovannetti, A. S. Holevo, and A. Mari, “Majorization and additivity for multimode bosonic Gaussian channels”, *Theoret. and Math. Phys.* **182**:2 (2015), 284–293.
[24] V. Giovannetti and S. Lloyd, “Additivity properties of a Gaussian channel”, *Phys. Rev. A* **69**:6 (2004), 062307, 9 pp.

[25] A. Grudka, M. Horodecki, and L. Pankowski, *Constructive counterexamples to additivity of minimum output Rényi entropy of quantum channels for all* $p > 2$, 2009 (v2 – 2010), 4 pp., arXiv:0911.2515.

[26] M. B. Hastings, “Superadditivity of communication capacity using entangled inputs”, *Nature Physics* **5**:4 (2009), 255–257.

[27] M. Hayashi, H. Imai, K. Matsumoto, M.-B. Ruskai, and T. Shimono, “Qubit channels which require four inputs to achieve capacity: implications for additivity conjectures”, *Quantum Inf. Comput.* **5**:1 (2005), 13–31.

[28] P. Hayden, *The maximal p-norm multiplicativity conjecture is false*, 2007, 12 pp., arXiv:0707.3291.

[29] P. Hayden and A. Winter, “Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$”, *Comm. Math. Phys.* **284**:1 (2008), 263–280.

[30] T. Heinosaari, A. S. Holevo, and M. M. Wolf, “The semigroup structure of Gaussian channels”, *Quantum Inf. Comput.* **10**:7-8 (2010), 619–635.

[31] T. Hiroshima, “Additivity and multiplicativity properties of some Gaussian channels for Gaussian inputs”, *Phys. Rev. A* **73**:1 (2006), 012330, 9 pp.

[32] A. S. Holevo, “The capacity of the quantum channel with general signal states”, *IEEE Trans. Inform. Theory* **44**:1 (1998), 269–273.

[33] A. S. Холево, *Вероятностные и статистические аспекты квантовой теории*, 2-е изд., доп. и испр., Ин-т компьют. исслед., М.–Ижевск 2003, 404 с.; English transl., A.S. Holevo, *Probabilistic and statistical aspects of quantum theory*, 2nd ed., Quaderni. Monographs, vol. 1, Edizioni della Normale, Pisa 2011, xvi+323 pp.

[34] A. S. Холево, *Семиаргическая структура квантовой теории*, Ин-т компьют. исслед., М.–Ижевск 2003, 191 с.; English transl., A. S. Holevo, *Statistical structure of quantum theory*, Lect. Notes Phys. Monogr., vol. 67, Springer-Verlag, Berlin 2001, x+159 pp.

[35] A. S. Холево, “Классические пропускные способности квантового канала с ограничением на входе”, *Теория вероятн. и ее примен.* **48**:2 (2003), 359–374; English transl., A. S. Holevo, “Entanglement-assisted capacities of constrained quantum channels”, *Theory Probab. Appl.* **48**:2 (2004), 243–255.

[36] A. S. Холево, “Комплементарные каналы и проблема аддитивности”, *Теория вероятн. и ее примен.* **51**:1 (2006), 133–143; English transl., A. S. Holevo, “Complementary channels and the additivity problem”, *Theory Probab. Appl.* **51**:1 (2007), 92–100.

[37] A. S. Холево, “Мультипликативность $p$-норм вполне положительных отображений и проблема аддитивности в квантовой теории информации”, *VMH* **61**:2(368) (2006), 113–152; English transl., A. S. Holevo, “Multiplicativity of $p$-norms of completely positive maps and the additivity problem in quantum information theory”, *Russian Math. Surveys* **61**:2 (2006), 301–339.

[38] A. S. Холево, *Квантовые системы, каналы, информация*, МЦНМО, М. 2010, 327 с.; English transl., A. S. Holevo, *Quantum systems, channels, information. A mathematical introduction*, de Gruyter Stud. Math. Phys., vol. 16, de Gruyter, Berlin 2012, xiv+349 pp.

[39] A. S. Холево, “Об экстремальных бозонных линейных каналах”, *ТМФ* **174**:2 (2013), 331–341; English transl., A. S. Holevo, “Extreme bosonic linear channels”, *Theoret. and Math. Phys.* **174**:2 (2013), 288–297.

[40] A. S. Holevo, M. Sohma, and O. Hirota, “Capacity of quantum Gaussian channels”, *Phys. Rev. A* **59**:3 (1999), 1820–1828.
A. S. Holevo, M. Sohma, and O. Hirota, “Error exponents for quantum channels with constrained inputs”, Rep. Math. Phys. 46:3 (2000), 343–358.

A. S. Holevo and R. F. Werner, “Evaluating capacities of bosonic Gaussian channels”, Phys. Rev. A 63:3 (2001), 032312, 14 pp.

M. Horodecki, P. W. Shor, and M. B. Ruskai, “Entanglement breaking channels”, Rev. Math. Phys. 15:6 (2003), 629–641.

C. King, “Maximal $p$-norms of entanglement breaking channels”, Quantum Inf. Comput. 3:2 (2003), 186–190; 2002, 7 pp., arXiv:quant-ph/0212057.

C. King, “Additivity for unital qubit channels”, J. Math. Phys. 43:10 (2002), 4641–4653.

C. King, “The capacity of the quantum depolarizing channel”, IEEE Trans. Inform. Theory 49:1 (2003), 221–229.

C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai, “Properties of conjugate channels with applications to additivity and multiplicativity”, Markov Process. Related Fields 13:2 (2007), 391–423.

J. R. Klauder and E. C. G. Sudarshan, Fundamentals of quantum optics, W. A. Benjamin, Inc., New York–Amsterdam 1968, xi+279 pp.

A. I. Kostrikin, Ю. И. Манин, Линейная алгебра и геометрия, 2-е изд., Наука, М. 1986, 304 с.; English transl., A. I. Kostrikin and Yu. I. Manin, Linear algebra and geometry, Algebra, vol. 1, Gordon and Breach Science Publishers, New York 1989, x+309 pp.

E. H. Lieb, “Proof of an entropy conjecture of Wehrl”, Comm. Math. Phys. 62:1 (1978), 35–41.

E. H. Lieb, “Gaussian kernels have only Gaussian maximizers”, Invent. Math. 102:1 (1990), 179–208.

E. H. Lieb and J. P. Solovej, “Proof of an entropy conjecture for Bloch coherent spin states and its generalizations”, Acta Math. 212:2 (2014), 379–398.

A. Mari, V. Giovannetti, and A. S. Holevo, “Quantum state majorization at the output of bosonic Gaussian channels”, Nature Communications 5 (2014), 3826.

V. Paulsen, Completely bounded maps and operator algebras, Cambridge Stud. Adv. Math., vol. 78, Cambridge Univ. Press, Cambridge 2002, xii+300 pp.

M. Reed and B. Simon, Methods of modern mathematical physics, vol. I: Functional analysis, Academic Press, New York–London 1972, xvii+325 pp.

A. Serafini, J. Eisert, and M. M. Wolf, “Multiplicativity of maximal output purities of Gaussian channels under Gaussian inputs”, Phys. Rev. A 71:1 (2005), 012320, 9 pp.

M. E. Širokov, “О супераддитивности выпуклого замыкания выходной энтропии квантового канала”, УМН 61:6(372) (2006), 191–192; English transl., M. E. Shirokov, “Superadditivity of the convex closure of the output entropy of a quantum channel”, Russian Math. Surveys 61:6 (2006), 1186–1188.

M. E. Širokov, “О свойствах квантовых каналов, связанных с классической пропускной способностью”, Теория вероятн. и ее примен. 52:2 (2007), 301–335; English transl., M. E. Shirokov, “On properties of quantum channels related to their classical capacity”, Theory Probab. Appl. 52:2 (2008), 250–276.

P. W. Shor, “Additivity of the classical capacity of entanglement-breaking quantum channels”, J. Math. Phys. 43:9 (2002), 4334–4340.

П. В. Шор, “О равенстве добавимости в теории квантовой информации”, Proc. Amer. Math. Soc. 6:2 (1955), 211–216.
[62] K. Temme, F. Pastawski, and M. J. Kastoryano, “Hypercontractivity of quasi-free quantum semigroups”, *J. Phys. A* **47**:40 (2014), 405303, 27 pp.; 2014, 26 pp., arXiv:1403.5224.

[63] A. Wehrl, “General properties of entropy”, *Rev. Modern Phys.* **50**:2 (1978), 221–260.

[64] R. F. Werner and A. S. Holevo, “Counterexample to an additivity conjecture for output purity of quantum channels”, *J. Math. Phys.* **43**:9 (2002), 4353–4357.

[65] A. Winter, “The maximum output $p$-norm of quantum channels is not multiplicative for any $p > 2$”, 2007 (v3 – 2008), 4 pp., arXiv:0707.0402.

Alexander S. Holevo
Steklov Mathematical Institute of Russian Academy of Sciences
E-mail: holevo@mi.ras.ru

Received 11/JAN/15
Translated by THE AUTHOR