Communications in
Mathematical
Physics

Convergence Rates for the Quantum Central Limit Theorem

Simon Becker[1] 📖, **Nilanjana Datta**[1]**, Ludovico Lami**[2,3]**, Cambyse Rouzé**[4]

[1] Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences University of Cambridge, Cambridge CB3 0WA, UK. E-mail: simon.becker@damtp.cam.ac.uk; n.datta@damtp.cam.ac.uk
[2] School of Mathematical Sciences and Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, University Park, Nottingham NG7 2RD, UK.
[3] Institute of Theoretical Physics and IQST, Universität Ulm, Albert-Einstein-Allee, 89069 Ulm, Germany. E-mail: ludovico.lami@gmail.com
[4] Zentrum Mathematik, Technische Universität München, 85748 Garching, Germany. E-mail: rouzecambyse@gmail.com

Contents

Abstract: Various quantum analogues of the central limit theorem, which is one of
the cornerstones of probability theory, are known in the literature. One such analogue,
due to Cushen and Hudson, is of particular relevance for quantum optics. It implies
that the state in any single output arm of an n-splitter, which is fed with n copies of a
centred state ρ with finite second moments, converges to the Gaussian state with the
same first and second moments as ρ. Here we exploit the phase space formalism to
carry out a refined analysis of the rate of convergence in this quantum central limit
theorem. For instance, we prove that the convergence takes place at a rate $\mathcal{O}\left(n^{-1/2}\right)$ in
the Hilbert–Schmidt norm whenever the third moments of ρ are finite. Trace norm or
relative entropy bounds can be obtained by leveraging the energy boundedness of the
state. Via analytical and numerical examples we show that our results are tight in many
respects. An extension of our proof techniques to the non-i.i.d. setting is used to analyse
a new model of a lossy optical fibre, where a given m-mode state enters a cascade of n
beam splitters of equal transmissivities $\lambda^{1/n}$ fed with an arbitrary (but fixed) environment
state. Assuming that the latter has finite third moments, and ignoring unitaries, we show
that the effective channel converges in diamond norm to a simple thermal attenuator,
with a rate $\mathcal{O}\left(n^{-\frac{1}{2(m+1)}}\right)$. This allows us to establish bounds on the classical and quantum
capacities of the cascade channel. Along the way, we derive several results that may be
of independent interest. For example, we prove that any quantum characteristic function
χ_ρ is uniformly bounded by some $\eta_\rho < 1$ outside of any neighbourhood of the origin;
also, η_ρ can be made to depend only on the energy of the state ρ.

1. Introduction

The Central Limit Theorem (CLT) is one of the cornerstones of probability theory. This
theorem and its various extensions have found numerous applications in diverse fields
including mathematics, physics, information theory, economics and psychology. Any
limit theorem becomes more valuable if it is accompanied by estimates for rates of con-
vergence. The Berry–Esseen theorem (see e.g. [1]), which gives the rate of convergence
of the distribution of the scaled sum of independent and identically distributed (i.i.d.)
random variables to a normal distribution, thus provides an important refinement of the
CLT.

The first results on quantum analogues of the CLT were obtained in the early 1970s by Cushen and Hudson [2], and Hepp and Lieb [3,4]. The approach of [3] was generalised by Giri and von Waldenfels [5] a few years later. These papers were followed by numerous other quantum versions of the CLT in the context of quantum statistical mechanics [6–14], quantum field theory [15–17], von Neumann algebras [18,19], free probability [20], noncommutative stochastic processes [21] and quantum information theory [22–24]. For a more detailed list of papers on noncommutative or quantum central limit theorems (QCLT), see for example [19,25] and references therein. A partially quantitative central limit theorem for unsharp measurements has been obtained in [26].

An important pair of non-commuting observables is the pair (x, p) of canonically conjugate operators, which obey Heisenberg's canonical commutation relations (CCR) $[x, p] = iI$, where I denotes the identity operator.[1] These observables could be, for example, the position and momentum operators of a quantum particle, or the so-called position and momentum quadratures of a single-mode bosonic field, described in the quantum mechanical picture by the Hilbert space $\mathcal{H}_1 := L^2(\mathbb{R})$ – the space of square integrable functions on $\mathbb{R}$. The corresponding annihilation and creation operators are constructed as $a := (x + ip)/\sqrt{2}$ and $a^\dagger := (x - ip)/\sqrt{2}$. When expressed in terms of $a, a^\dagger$, the CCR take the form $[a, a^\dagger] = I$.

Quantum states are represented by density operators, i.e. positive semi-definite trace class operators with unit trace. A state ρ of a continuous variable quantum system is uniquely identified by its characteristic function, defined for all $z \in \mathbb{C}$ by $\chi_\rho(z) :=$ $\mathrm{Tr}\left[\rho\, e^{za^\dagger - z^* a}\right]$. The special class of Gaussian states comprises all quantum states whose characteristic function is the (classical) characteristic function of a normal random variable on $\mathbb{C}$.[2] Exactly as in the classical case, a quantum Gaussian state is uniquely defined by its mean and covariance matrix.

Cushen and Hudson [2] proved a quantum CLT for a sequence of pairs of such canonically conjugate operators $\{(x_n, p_n) : n = 1, 2, \ldots\}$, with each pair acting on a distinct copy of the Hilbert space $\mathcal{H}_1$. More precisely, they showed that sequences that are stochastically independent and identically distributed, and have finite covariance matrix and zero mean with respect to a quantum state ρ (given by a density operator on $\mathcal{H}_1$), are such that their scaled sums converge in distribution to a normal limit distribution [2, Theorem 1].

Their result admits a physical interpretation in terms of a passive quantum optical element known as the n-splitter. This can be thought of as the unitary operator $U_{n\text{-split}}$ that acts on n annihilation operators of n independent optical modes as $U_{n\text{-split}}\, a_j\, U_{n\text{-split}}^\dagger = \sum_k F_{jk} a_k$, where $F_{jk} = e^{\frac{2\pi jk}{n} i}$ is the discrete Fourier transform matrix. Passivity here means that $U_{n\text{-split}}$ commutes with the canonical Hamiltonian of the field, i.e. $\left[U_{n\text{-split}}, \sum_j a_j^\dagger a_j\right] = 0$. When n identical copies of a state ρ are combined by means of an n-splitter, and all but the first output modes are traced away, the resulting output state is called the *n-fold quantum convolution* of ρ, and denoted by $\rho^{\boxplus n}$. This nomenclature is justified by the fact that the characteristic function $\chi_{\rho \boxplus \sigma}$ of two states ρ and σ is equal to the product of the characteristic functions of ρ and σ, a relation analogous to that satisfied by characteristic functions of convolutions of classical random variables. Observe state $\rho^{\boxplus n}$ can also be obtained as the output of a cascade of $n - 1$

---

[1] Throughout this paper we set $\hbar = 1$.

[2] The characteristic function of a complex-valued random variable X is defined by $\chi_X(z) := \mathbb{E}\left[e^{zX^* - z^* X}\right]$.

beam splitters with suitably tuned transmissivities $\lambda_j = j/(j+1)$ for $j = 1, 2, \ldots n-1$ (see Fig. 1a).

Cushen and Hudson's result is that if ρ is a centred state (i.e. with zero mean) and has finite second moments, its convolutions $\rho^{\boxplus n}$ converge to the Gaussian state ρ_G with the same first and second moments as ρ in the limit $n \to \infty$ (Theorem 3). In [2, Theorem 1], the convergence is with respect to the weak topology of the Banach space of trace class operators, which translates to pointwise convergence of the corresponding characteristic functions, by a quantum analogue of Levy's lemma that is also proven in [2]. This in turn implies that the convergence actually is with respect to the strong topology, i.e. in trace norm (see [27], or [28, Lemma 4]).

In this paper, we focus on the framework proposed by Cushen and Hudson, and provide a refinement of their result by deriving estimates for the associated rates of convergence. We consider a quantum system composed of m modes of the electromagnetic field, each modelled by an independent quantum harmonic oscillator, so that the corresponding Hilbert space becomes $\mathcal{H}_1^{\otimes m} = L^2(\mathbb{R}^m)$. The main contribution of this paper consists of estimates on rate of convergence of $\rho^{\boxplus n}$ to the 'Gaussification' ρ_G of ρ, obtained under suitable assumptions on ρ – typically, the finiteness of higher-order moments. In analogy with the classical case, we refer to our Theorems 6 and 7 as *quantum Berry–Esseen theorems*. Our estimates are given in the form of bounds on the Schatten p-norms (for $p = 1$ and 2) of the difference $(\rho^{\boxplus n} - \rho_G)$ in the limit of large n, as well as bounds on the relative entropy of $\rho^{\boxplus n}$ with respect to ρ_G in the same limit.

We also show that the assumption of finiteness of the second moments cannot be removed from the Cushen–Hudson theorem. Namely, we construct a simple example of a single-mode quantum state σ such that $\text{Tr}\left[\sigma\, (aa^\dagger)^{1-\delta}\right]$ is finite for all $\delta > 0$ (and infinite for $\delta = 0$), yet $\sigma^{\boxplus n}$ does not converge to any quantum state as $n \to \infty$.

As an application, we propose and study a new model of optical fibre, represented as a cascade of n beam splitters, each with transmissivity $\lambda^{1/n}$ and fed with a fixed environment state ρ, which is assumed to have bounded energy and thermal Gaussification. Such a model may be relevant to the mathematical modelisation of a channel running across an integrated optical circuit [29,30]. We are able to show that for $n \to \infty$ the cascade channel converges in diamond norm, up to irrelevant symplectic unitaries, to a thermal attenuator channel with transmissivity λ and the same photon number as that of the environment state ρ. Furthermore, an extension of our results to the non-i.i.d. setting allows us to bound the rate of convergence in terms of the diamond norm distance. Finally, combining existing continuity bounds on entropies and energy-constrained channel capacities [31,32], obtained by Winter [33,34] and Shirokov [35,36], with the known formulae expressing or estimating energy-constrained classical [37,38] and quantum [39–45] capacities of thermal attenuator channels, we derive bounds on the same capacities for the cascade channel.

Finally, along the way we derive several novel results concerning quantum characteristic functions, which we believe to be of independent interest. First, we prove the simple yet remarkable fact that *convolving any two quantum states (i.e. mixing them in a* 50 : 50 *beam splitter) always results in a state with non-negative Wigner function (Lemma 16).* This allows us to interpret the quantum central limit theorem as a result on classical random variables, in turn enabling us to transfer techniques from classical probability theory to the quantum setting. Secondly, we derive new decay bounds on the behaviour of the quantum characteristic function both at the origin and at infinity. For instance, we prove that for any m-mode quantum state ρ and for any $\varepsilon > 0$ there exists a constant $\eta = \eta(\rho, \varepsilon) < 1$ such that $|\chi_\rho(z)| \leq \eta(\rho, \varepsilon)$ for all $z \in \mathbb{C}^m$ with $|z| \geq \varepsilon$
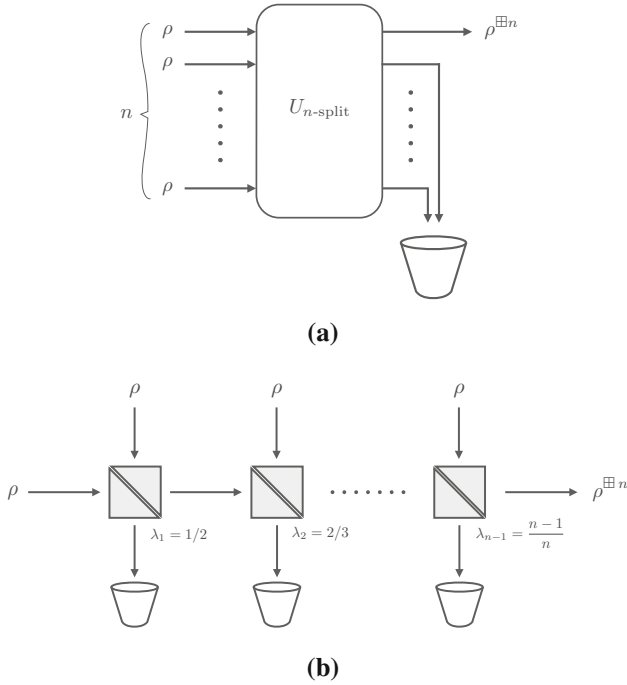
(a)

(b)

Fig. 1. The n-fold convolution $\rho^{\boxplus n}$ of a state ρ can be realised by mixing n copies of it either: **a** in an n-splitter; or **b** in a cascade of beam splitters with suitably tuned transmissivities

(Proposition 14). Moreover, we show that such a constant can be made to depend only on the second moments of the state, assuming they are finite (Proposition 15). As an explicit example, consider a single-mode state ρ with mean energy E. We then prove that $|\chi_\rho(z)| \leq 1 - \frac{c}{E^2}$ for all z with $|z| \geq \frac{C}{\sqrt{E}}$, where c, C are universal constants. Note that any such bound must depend on the energy, as one can construct a sequence of highly squeezed Gaussian states for which the modulus of the characteristic function approaches one at any designated point in phase space (Example 2).

Layout of the paper: In Sect. 2 we introduce the notation and definitions used in the paper. In Sect. 3 we recall the Cushen and Hudson quantum central limit theorem. Our main results are presented in Sect. 4. The rest of the paper is devoted to the proofs of these results. We start with the novel properties of quantum characteristic functions (Sect. 5), which lie at the heart of our approach. Then, in Sect. 6 we prove our quantum Berry–Esseen theorems. Section 7 is devoted to the discussion of the optimality and sharpness of our results. In Sect. 8 we apply our quantitative non-i.i.d. extension of the Cushen–Hudson theorem to an optical fibre subject to non-Gaussian environment noise. The paper contains a technical appendix (Appendix A) that makes the connection between moments and the regularity of the quantum characteristic function and shows that our definition of moments induces a canonical family of interpolation spaces.

2. Notation and Definitions

In this section, we fix the basic notations used in the paper, and introduce the necessary definitions.

2.1. Mathematical notation. Let $\mathcal{H}$ denote a separable Hilbert space, and let $\mathcal{B}(\mathcal{H})$ denote the set of bounded linear operators acting on $\mathcal{H}$. Let $\mathcal{D}(\mathcal{H})$ denote the set of *quantum states* of a system with Hilbert space $\mathcal{H}$, that is the set of density operators ρ (positive semi-definite, i.e. $\rho \geq 0$, trace class operators[3] with unit trace) acting on $\mathcal{H}$. We denote by $\| \cdot \|_{\mathcal{T}_p(\mathcal{H})} \equiv \| \cdot \|_p$ the Schatten p-norm, defined as $\|X\|_p = (\mathrm{Tr} \, |X|^p)^{1/p}$. The *Schatten p-class* $\mathcal{T}_p(\mathcal{H})$ is the Banach subspace of $\mathcal{B}(\mathcal{H})$ formed by all bounded linear operators whose Schatten p-norm is finite. We shall hereafter refer to $\mathcal{T}_1(\mathcal{H})$ as the set of *trace class* operators, to the corresponding norm $\| \cdot \|_1$ as the trace norm, and to the induced distance (e.g. between quantum states) as the trace distance. The case $p = 2$ is also special, as the norm $\| \cdot \|_2$ coincides with the *Hilbert–Schmidt norm*.

Let A, B be positive semi-definite operators defined on some domains $\mathrm{Dom}(A), \mathrm{Dom}(B) \subseteq \mathcal{H}$. According to [46, Definition 10.15], we write that $A \geq B$ if and only if $\mathrm{Dom}\left(A^{1/2}\right) \subseteq \mathrm{Dom}\left(B^{1/2}\right)$ and $\left\|A^{1/2} \, |\psi\rangle\right\|^2 \geq \left\|B^{1/2} \, |\psi\rangle\right\|^2$ for all $|\psi\rangle \in \mathrm{Dom}\left(A^{1/2}\right)$. Now, let A be a positive semi-definite operator, and let ρ be a quantum state with spectral decomposition $\rho = \sum_i p_i \, |e_i\rangle\langle e_i|$. We define the *expected value* of A on ρ as

$$\mathrm{Tr}[\rho A] := \sum_{i \, : \, p_i > 0} p_i \left\|A^{1/2} \, |e_i\rangle\right\|^2 \in \mathbb{R}_+ \cup \{+\infty\}, \tag{1}$$

with the convention that $\mathrm{Tr}[\rho A] = +\infty$ if the above series diverges or if there exists an index i such that $p_i > 0$ and $|e_i\rangle \notin \mathrm{Dom}\left(A^{1/2}\right)$. To extend this definition to a generic densely defined self-adjoint operator X on $\mathcal{H}$, it is useful to consider its decomposition $X = X_+ - X_-$ into positive and negative part [46, Example 7.1]. We will say that X has *finite expected value on ρ* if $|e_i\rangle \in \mathrm{Dom}\left(X_+^{1/2}\right) \cap \mathrm{Dom}\left(X_-^{1/2}\right)$ for all i such that $p_i > 0$, and moreover the two series $\sum_i p_i \left\|X_{\pm}^{1/2} \, |e_i\rangle\right\|^2$ both converge. In this case, we call

$$\mathrm{Tr}[\rho X] := \sum_{i \, : \, p_i > 0} p_i \left\|X_+^{1/2} \, |e_i\rangle\right\|^2 - \sum_{i \, : \, p_i > 0} p_i \left\|X_-^{1/2} \, |e_i\rangle\right\|^2 \tag{2}$$

the *expected value* of X on ρ. Clearly, given two operators $A \geq B \geq 0$, we have that $\mathrm{Tr}[\rho A] \geq \mathrm{Tr}[\rho B]$.

For two real sequences $(a_n(\lambda))_n$, $(b_n(\lambda))_n$ that depend on some parameter λ, we write $a_n(\lambda) = \mathcal{O}_\lambda (b_n(\lambda))$ if there exists a constant $c_\lambda > 0$ that only depends on λ such that $|a_n(\lambda)| \leq c_\lambda |b_n(\lambda)|$ holds in the limit $n \to \infty$. We also write $a_n(\lambda) = \mathcal{O}_\lambda \left(b_n(\lambda)^\infty\right)$ if for every $N \in \mathbb{N}$ we have that $a_n(\lambda) = \mathcal{O}_\lambda \left(b_n(\lambda)^N\right)$.

For an n-linear tensor $A : \times_{i=1}^n \mathbb{C}^m \to \mathbb{C}^k$, we write $A(x^{\times n}) := A \underbrace{(x, \ldots, x)}_{n \text{ times}}$ if the vector we apply the tensor to is the same in every component. For functions f, we sometimes abuse the notation by denoting the norm of this function as $\|f(z)\|$ instead of $\|f\|$. We denote with $*$ the entry-wise complex conjugation, with $\top$ the standard transposition of vectors, and with $\dagger$ the combination of the two.

---

[3] That is, operators $A \in \mathcal{B}(\mathcal{H})$ for which $\|A\|_1 := \mathrm{Tr} \, |A| < \infty$.

For partial derivatives with respect to complex variables z, z^* we write ∂_z and ∂_{z^*}. Consider an m-dimensional multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ with $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_m$. Then $\partial_z^\alpha := \partial_{z_1}^{\alpha_1} \partial_{z_2}^{\alpha_2} \ldots \partial_{z_m}^{\alpha_m}$, and analogously for z^*. The total derivatives of order k of a function $f : \mathbb{C}^m \to \mathbb{C}$ we denote by $D^k f := \left(\partial_z^\alpha \partial_{z^*}^\beta f \right)_{|\alpha|+|\beta|=k}$. We then recall the definition of the Fréchet derivative for functions $f : \mathbb{C}^m \to \mathbb{C}$ such that $D^k f : \mathbb{C}^m \to \mathcal{B}(\underbrace{\mathbb{C}^m \times \cdots \times \mathbb{C}^m}_{k \text{ times}}, \mathbb{C})$ and therefore

$$D^k f(z) \left(v^{(1)}, \ldots, v^{(k)} \right) = \sum_{|\alpha|+|\beta|=k} \partial_z^\alpha \partial_{z^*}^\beta f(z) \left(\prod_{\ell=1}^{|\alpha|} v_{j_\alpha(\ell)}^{(\ell)} \right) \left(\prod_{\ell=|\alpha|+1}^{k} v_{j_\beta(\ell-|\alpha|)}^{(\ell)} \right)^*, \tag{3}$$

with $j_\alpha(\ell) := \min \left\{ j \in \{1, \ldots, m\} : \ell \le \sum_{j'=1}^{j} \alpha_{j'} \right\}$. Let $C_0(\mathbb{C}^m)$ denote the space of continuous functions $f : \mathbb{C}^m \to \mathbb{C}$ that tend to zero as $|z| \to \infty$, where for $z \in \mathbb{C}^m$ we set

$$|z| := \sqrt{\sum_{j=1}^{m} |z_j|^2} . \tag{4}$$

We write $C_c^\infty(\mathbb{C}^m)$ to denote the space of smooth and compactly supported functions on $\mathbb{C}^m$. For some open set $\Omega \subseteq \mathbb{C}^m$ with closure $\overline{\Omega}$, a function $f : \overline{\Omega} \to \mathbb{C}$, and a non-negative integer $k \in \mathbb{N}_0$, we denote by $C^k(\overline{\Omega})$ the space of functions for which the norm

$$\|f\|_{C^k(\overline{\Omega})} := \max_{|\alpha|+|\beta| \le k} \sup_{z \in \Omega} \left| \partial_z^\alpha \partial_{z^*}^\beta f(z) \right| \tag{5}$$

is finite. Here, $\alpha, \beta \in \mathbb{N}_0^m$ are multi-indices. When $k \ge 0$ is *not* an integer, we define instead

$$\|f\|_{C^k(\overline{\Omega})} := \|f\|_{C^{\lfloor k \rfloor}(\overline{\Omega})} + \max_{|\alpha|+|\beta|=\lfloor k \rfloor} \sup_{\substack{z,w \in \Omega, \\ z \ne w}} \frac{\left| \partial_z^\alpha \partial_{z^*}^\beta f(z) - \partial_z^\alpha \partial_{z^*}^\beta f(w) \right|}{\|z-w\|^{k-\lfloor k \rfloor}} . \tag{6}$$

This extension allows us to consider the normed spaces $C^k(\overline{\Omega})$ for all $k \ge 0$. Typically, we will deal with the case where Ω is bounded, so that $C^k(\overline{\Omega})$ is in fact a Banach space. Finally, $L^2(\Omega)$ will denote the space of equivalence classes of measurable functions $f : \Omega \to \mathbb{C}$ whose L^2 norm $\|f\|_{L^2(\Omega)}^2 := \int_\Omega |f(z)|^2 \, d^{2m}z$ is finite.

2.2. Definitions.

2.2.1. Quantum information with continuous variables
In this paper, we focus on continuous variable quantum systems. The Hilbert space of a set of m harmonic oscillators, in this context called 'modes', is the space $\mathcal{H}_m := L^2(\mathbb{R}^m)$ of square-integrable functions on $\mathbb{R}^m$. Let x_j, p_j be the canonical position and momentum operators on the jth mode. The m annihilation and creation operators, denoted by $a_j := (x_j + i p_j)/\sqrt{2}$ and $a_j^\dagger := (x_j - i p_j)/\sqrt{2}$ $(j = 1, \ldots, m)$, satisfy the commutation relations

$$[a_j, a_k] = 0, \qquad [a_j, a_k^\dagger] = \delta_{jk} I , \tag{7}$$

where I is the identity on $\mathcal{H}_m$. An m-mode quantum state ρ is said to be *centred* if

$$\text{Tr}[\rho \, a_j] := \frac{1}{\sqrt{2}} \left(\text{Tr}[\rho \, x_j] + i \, \text{Tr}[\rho \, p_j] \right) \equiv 0 \qquad \forall \, j = 1, \ldots, m \, , \tag{8}$$

i.e. if all expected values of the canonical operators on ρ, defined according to (2), vanish. For an m-tuple of non-negative integers $n = (n_1, \ldots, n_m) \in \mathbb{N}_0^m$, the corresponding *Fock state* is defined by $|n\rangle := (n_1! \ldots n_m!)^{-1/2} \bigotimes_{j=1}^m \left(a_j^\dagger \right)^{n_j} |0\rangle$, where $|0\rangle \in \mathcal{H}_m$ denotes the (multi-mode) vacuum state. In what follows, we often consider $m = 1$.

The (von Neumann) *entropy* of a quantum state ρ is defined as

$$S(\rho) := - \text{Tr} \left[\rho \log \rho \right], \tag{9}$$

which is well defined although possibly infinite.[4] The *relative entropy* between two states ρ and σ is usually written as follows [47]

$$D(\rho \| \sigma) := \text{Tr} \left[\rho \left(\log \rho - \log \sigma \right) \right]. \tag{10}$$

Again, the above expression is well defined and possibly infinite [48].[5]

For two Hilbert spaces $\mathcal{H}, \mathcal{H}'$, a quantum channel $\mathcal{N} : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H}')$ is a completely positive, trace-preserving linear map. For a linear map $\mathcal{L} : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H}')$, we define its *diamond norm* as

$$\| \mathcal{L} \|_\diamond := \sup_{k \in \mathbb{N}} \sup_{X \in \mathcal{T}_1(\mathcal{H} \otimes \mathbb{C}^k)} \frac{\left\| (\mathcal{L} \otimes \text{id}_{\mathbb{C}^k})(X) \right\|_1}{\|X\|_1}, \tag{11}$$

where the supremum is over all non-zero trace class operators X on $\mathcal{H} \otimes \mathbb{C}^k$.

Consider a quantum system with Hilbert space $\mathcal{H}$, governed by a *Hamiltonian H*, which is taken to be a positive (possibly unbounded) operator on $\mathcal{H}$. The energy of a state $\rho \in \mathcal{D}(\mathcal{H})$ is the quantity $\text{Tr}[\rho H] \in \mathbb{R}_+ \cup \{+\infty\}$ defined as in (1).

Given two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, a Hamiltonian H on $\mathcal{H}$, and some energy bound $E > \inf_{|\psi\rangle \in \mathcal{H}} \langle \psi | H | \psi \rangle \geq 0$, the corresponding *energy-constrained classical capacity* of a channel $\mathcal{N} : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H}')$ is given by [31,49–52]

$$\begin{aligned}
\mathscr{C}_H \left(\mathcal{N}, E \right) &= \lim_{n \to \infty} \frac{1}{n} \chi_{H^{(n)}} \left(\mathcal{N}^{\otimes n}, nE \right), \\
\chi_H \left(\mathcal{N}, E \right) &:= \sup_{\substack{\{p_i, \rho_i\}_i \\ \sum_i p_i \, \text{Tr}[\rho_i H] \leq E}} \left\{ S \left(\sum_i p_i \rho_i \right) - \sum_i p_i \, S \left(\rho_i \right) \right\}, 11
\end{aligned} \tag{12}$$

where it is understood that the Hamiltonian $H^{(n)}$ on $\mathcal{H}^{\otimes n}$ is given by $H^{(n)} := \sum_{j=1}^n H_j$, where H_j acts on the jth tensor factor, and tensor products with the identity operator

---

[4] One way to define it is via the infinite sum $S(\rho) = \sum_i (-p_i \log p_i)$, where $\rho = \sum_i p_i |e_i\rangle\langle e_i|$ is the spectral decomposition of ρ. Since all terms of this sum are non-negative, the sum itself can be assigned a well-defined value, possibly $+\infty$.

[5] To define it one considers the infinite sum $D(\rho \| \sigma) := \sum_{i,j} |\langle e_i | f_j \rangle|^2 \left(p_i \log p_i - p_i \log q_j + q_j - p_i \right)$, where $\rho = \sum_i p_i |e_i\rangle\langle e_i|$ and $\sigma = \sum_j q_j |f_j\rangle\langle f_j|$ are the spectral decompositions of ρ and σ, respectively. As detailed in [48], the convexity of $x \mapsto x \log x$ implies that all terms of this sum are non-negative, which makes the expression well defined.

are omitted for notational simplicity. With the same notation, one can also define the *energy-constrained quantum capacity* of $\mathcal{N}$, given by [32,34,53–55]

$$Q_H(\mathcal{N}, E) := \lim_{n \to \infty} \frac{1}{n} Q_{H^{(n)}}^{(1)}(\mathcal{N}^{\otimes n}, nE),$$

$$Q_H^{(1)}(\mathcal{N}, E) := \sup_{\substack{|\Psi\rangle \in \mathcal{H} \otimes \widetilde{\mathcal{H}} \\ \langle \Psi | H \otimes I | \Psi \rangle \leq E}} \left\{ S\left(\mathcal{N}\left(\operatorname{Tr}_{\widetilde{\mathcal{H}}} |\Psi\rangle\langle\Psi|\right)\right) - S\left((\mathcal{N} \otimes I)\left(|\Psi\rangle\langle\Psi|\right)\right) \right\}, \quad (13)$$

where $\operatorname{Tr}_{\widetilde{\mathcal{H}}}$ is the partial trace over the entirely arbitrary ancillary Hilbert space $\widetilde{\mathcal{H}}$. In this paper we are interested in the simple case $\mathcal{H} = \mathcal{H}_m = L^2(\mathbb{R}^m)$ and $\mathcal{H}' = \mathcal{H}_{m'} = L^2(\mathbb{R}^{m'})$, so that there is a natural choice for H, namely, the *canonical Hamiltonian*

$$H_m := \sum_{j=1}^{m} a_j^\dagger a_j \quad (14)$$

of m modes. In this case, we will omit the subscripts and simply write the energy-constrained capacities as $\mathscr{C}(\mathcal{N}, E)$ and $\mathcal{Q}(\mathcal{N}, E)$.

2.2.2. Phase space formalism
We define the *displacement operator* $\mathscr{D}(z)$ associated with a complex vector $z \in \mathbb{C}^m$ as

$$\mathscr{D}(z) = \exp\left[\sum_j (z_j a_j^\dagger - z_j^* a_j)\right]. \quad (15)$$

Thus, $\mathscr{D}(z)$ is a unitary operator and satisfies $\mathscr{D}(z)^\dagger = \mathscr{D}(-z)$ and

$$\mathscr{D}(z)\mathscr{D}(w) = \mathscr{D}(z + w)\, e^{\frac{1}{2}(z^\mathsf{T} w^* - z^\dagger w)}, \quad (16)$$

valid for all $z, w \in \mathbb{C}^m$.

Let $H_{\text{quad}} = \sum_{j,k} \left(X_{jk} a_j^\dagger a_k + Y_{jk} a_j a_k + Y_{jk}^* a_j^\dagger a_k^\dagger \right)$, where $X = X^\dagger$ is an $m \times m$ Hermitian matrix, and $Y = Y^\mathsf{T}$ is an $m \times m$ complex symmetric matrix. The unitaries $e^{-iH_{\text{quad}}}$ generated by such Hamiltonians, and products thereof,[6] are called *symplectic unitaries*, because they induce a symplectic linear transformation at the phase space level $(z_R, z_I) \in \mathbb{R}^{2m}$, where $z_R := \Re z$ and $z_I := \Im z$ [56,57]. A symplectic unitary is called *passive* if it commutes with the number operator $\sum_j a_j^\dagger a_j$, which happens whenever the generating Hamiltonian H_{quad} satisfies $Y = 0$. A passive symplectic unitary V acts on annihilation operators as $V a_j V^\dagger = \sum_k U_{jk} a_k$, where U is an $m \times m$ unitary matrix.

For trace class operators $T \in \mathcal{T}_1(\mathcal{H}_m)$, the *quantum characteristic function* $\chi_T : \mathbb{C}^m \to \mathbb{C}$ is given by

$$\chi_T(z) := \operatorname{Tr}[T \mathscr{D}(z)]. \quad (17)$$

Conversely, the operator T can be reconstructed from χ_T via the weakly defined identity

$$T = \int \frac{d^{2m}z}{\pi^m} \chi_T(z)\mathscr{D}(-z). \quad (18)$$

---

[6] While not all products of unitaries of the form $e^{-iH_{\text{quad}}}$ can be written as a single exponential, two such factors always suffice. See [56, p.37], combined with [56, Propositions 2.12, 2.18, and 2.19] and with the observation that the exponential Lie map of the unitary group is surjective.

Observe that the adjoint $T^\dagger$ of T satisfies $\chi_{T^\dagger}(z) = \chi_T(-z)^*$ for all $z \in \mathbb{C}^m$, so that T is self-adjoint if and only if $\chi_T(-z) \equiv \chi_T(z)^*$. The characteristic function χ_T of a trace class operator T is bounded and uniformly continuous [58, § 5.4]. If T is positive semi-definite (e.g. if T is a density operator), then $\max_\alpha |\chi_T(\alpha)| = \chi_T(0) = \text{Tr}[T]$.

We write $|\psi_f\rangle$ to denote the pure state corresponding to the wave function $f \in L^2(\mathbb{R}^m)$, so that the corresponding rank-one state $\psi_f := |\psi_f\rangle\langle\psi_f|$ has the following characteristic function:

$$\chi_{\psi_f}(z) = e^{-iz_I^\mathsf{T} z_R} \int d^m x \, f^*(x) \, f\left(x - \sqrt{2}z_R\right) e^{\sqrt{2}\,i\,z_I^\mathsf{T} x}, \tag{19}$$

where as usual $z = z_R + iz_I$.

The Fourier transform of the characteristic function is known as the *Wigner function*. For a trace class operator T, the Wigner function is given by [59, Eq. (4.5.12) and (4.5.19)]

$$W_T(z) := \int \frac{d^{2m} w}{\pi^{2m}} \chi_T(w) \, e^{z^\mathsf{T} w^* - z^\dagger w} \tag{20}$$

$$= \frac{2^m}{\pi^m} \text{Tr}\left[\mathscr{D}(-z)T\mathscr{D}(z)\,(-1)^{\sum_j a_j^\dagger a_j}\right]. \tag{21}$$

Observe that $W_{T^\dagger}(z) = W_T(z)^*$, so that T is self-adjoint if and only if $W_T(z) \in \mathbb{R}$ for all $z \in \mathbb{C}^m$. From (21) it is not difficult to see that $|W_T(z)| \le \frac{2^m}{\pi^m}\|T\|_1$, where $\|T\|_1 = \text{Tr}\,|T|$ reduces to 1 when T is a density operator. By taking the Fourier transform of (19), one can show that

$$W_{\psi_f}(z) = \frac{2}{\pi} \int d^m x \, f^*\left(x + \sqrt{2}z_R\right) f\left(-x + \sqrt{2}z_R\right) e^{2\sqrt{2}\,i\,z_I^\mathsf{T} x}. \tag{22}$$

Moreover, the energy of any density matrix, ρ, can be obtained as a phase space integral

$$\int d^{2m} z \,\|z\|^2 \, W_\rho(z) = \text{Tr}\left[\rho\left(H_m + \frac{m}{2}\,I\right)\right] \tag{23}$$

The displacement operator $\mathscr{D}(z)$ induces a translation or *displacement* of the Wigner function as follows, hence the nomenclature:

$$\chi_{\mathscr{D}(z)\,\rho\,\mathscr{D}(z)^\dagger}(u) = e^{z^\dagger u - z^\mathsf{T} u^*}\chi_\rho(u), \qquad W_{\mathscr{D}(z)\,\rho\,\mathscr{D}(z)^\dagger}(u) = W_\rho(u - z). \tag{24}$$

The map $T \mapsto \chi_T$, defined for trace class operators T in (17), extends uniquely to an isomorphism between the space of Hilbert–Schmidt operators and that of square-integrable functions $L^2(\mathbb{C}^m)$. In fact, the *quantum Plancherel theorem* guarantees that this is also an isometry, namely

$$\text{Tr}[S^\dagger T] = \int \frac{d^{2m} z}{\pi^m} \chi_S(z)^* \,\chi_T(z) = \pi^m \int d^{2m} z \, W_S(z)^* W_T(z) \tag{25}$$

and therefore

$$\|\rho - \sigma\|_2^2 = \int \frac{d^{2m} z}{\pi^m} \left|\chi_\rho(z) - \chi_\sigma(z)\right|^2 = \pi^m \int d^{2m} z \,\left(W_\rho(z) - W_\sigma(z)\right)^2. \tag{26}$$

Henceforth, we refer to (26) as the *quantum Plancherel identity*.

Gaussian states on $\mathcal{H}_m$ are the density operators $\rho \in \mathcal{D}(\mathcal{H}_m)$ such that $W_\rho(z)$ is a Gaussian probability distribution on the real space $(z_R, z_I) \in \mathbb{R}^{2m}$ and are uniquely defined by their first and second moments. A particularly simple example of a single-mode Gaussian state is a *thermal state* with mean photon number $N \in [0, \infty)$, given by

$$\tau_N := \frac{1}{N+1} \sum_{n=0}^{\infty} \left(\frac{N}{N+1} \right)^n |n\rangle\langle n| \,. \tag{27}$$

The thermal state is the maximiser of the entropy among all states with a fixed maximum average energy:

$$\max \left\{ S(\rho) : \rho \in \mathcal{D}(\mathcal{H}_1), \ \mathrm{Tr}\left[\rho \, a^\dagger a \right] \leq N \right\} = S(\tau_N) = g(N) \tag{28}$$

for all $N \geq 0$, where the function g is defined by

$$g(x) := (x+1)\log(x+1) - x \log x \,. \tag{29}$$

The characteristic function and Wigner function of the thermal state evaluate to [59, Eq. (4.4.21) and (4.5.31)]

$$\chi_{\tau_N}(z) = e^{-(2N+1)|z|^2/2} \,, \qquad W_{\tau_N}(z) = \frac{2}{\pi(2N+1)} e^{-2|z|^2/(2N+1)} \,, \tag{30}$$

respectively, so that τ_N is easily seen to be a centred Gaussian state.

2.3. Moments.

Definition 1 *(Standard Moments).* An *m*-mode quantum state ρ is said to have *finite standard moments* of order up to k, for some $k \in [0, \infty)$, if

$$M_k(\rho) := \mathrm{Tr}\left[\rho H_m^{k/2} \right] < \infty \,, \tag{31}$$

where H_m is the canonical Hamiltonian (14), and the above trace is defined as in (1).

Remark. The above condition is fairly easy to check once the matrix representation of ρ in the Fock basis is given. Namely, resorting to (1) and exchanging the order of summation for infinite series with non-negative terms, we see that (31) is equivalent to

$$M_k(\rho) = \sum_{n \in \mathbb{N}_0^m} (m + |n|)^{k/2} \langle n|\rho|n\rangle < \infty \,, \tag{32}$$

where as usual $|n| = \sum_j n_j$.

Given $k > 0$ and $m \in \mathbb{N}$, we can also define, by analogy with classical harmonic analysis, the *m-mode bosonic Sobolev space* of order k as follows

$$\mathscr{W}^{k,1}(\mathcal{H}_m) := \left\{ X \in \mathscr{T}_1(\mathcal{H}_m); \ \|X\|_{\mathscr{W}^{k,1}(\mathcal{H}_m)} < \infty \right\},$$

where as usual $\mathcal{H}_m = L^2(\mathbb{R}^m)$. Here, we set

$$\|X\|_{\mathscr{W}^{k,1}(\mathcal{H}_m)} := \left\| (H_m + mI)^{k/4} \, X \, (H_m + mI)^{k/4} \right\|_1 \,,$$

with the canonical Hamiltonian on m modes being defined by (14). For density operators ρ it holds, using monotone convergence and cyclicity of the trace, that

$$
\begin{aligned}
\|\rho\|_{\mathscr{W}^{k,1}(\mathcal{H}_m)} &= \sup_E \mathrm{Tr}\left[\mathbb{1}_{[0,E]}(H_m)\,(H_m + mI)^{k/4}\,\rho\;(H_m + mI)^{k/4}\,\mathbb{1}_{[0,E]}(H_m)\right]\\
&= \sup_E \mathrm{Tr}\left[\rho\;(H_m + mI)^{k/2}\,\mathbb{1}_{[0,E]}(H_m)\right]\\
&= \mathrm{Tr}\left[\rho\;(H_m + mI)^{k/2}\right]
\end{aligned}
$$

where $\mathbb{1}_{[0,E]}$ is the indicator function of the interval $[0, E]$.

It is well known that the characteristic function of any classical random variable with finite moments of order up to k (with k being a positive integer) is continuously differentiable k times everywhere. We can draw inspiration from this fact to devise an alternative way to introduce moments, relying on the regularity of the quantum characteristic function, in the quantum setting as well. We refer to moments defined in this manner as *phase space moments*.

Definition 2 *(Phase space moments).* An m-mode quantum state ρ is said to have *finite phase space moments* of order up to k, for some $k \in [0, \infty)$, if

$$
M'_k(\rho, \varepsilon) := \big\|\chi_\rho\big\|_{C^k(B(0,\varepsilon))} < \infty \tag{33}
$$

for some $\varepsilon > 0$, where $B(0, \varepsilon) := \{z \in \mathbb{C}^m : |z| \le \varepsilon\}$ is the Euclidean ball of radius ε centred in 0, and the norm on the space $C^k\,(B(0, \varepsilon))$ is defined by (5) and (6).

In complete analogy with the classical case, finiteness of standard moments implies local differentiability of the characteristic function, and hence finiteness of phase space moments. See Theorem 9 of Sect. 4.

However, the converse is not true in general. This is not surprising, as the same phenomenon is observed for classical random variables. In fact, a famous example by Zygmund [60] shows the existence of classical random variables with continuously differentiable characteristic function whose first absolute moments do not exist. We can swiftly carry over his example to the quantum realm, e.g. by considering a particular displaced vacuum state $\rho := c \sum_{n=2}^\infty \frac{1}{n^2 \log n} \big(\mathscr{D}(n)\,|0\rangle\langle 0|\,\mathscr{D}(n)^\dagger + \mathscr{D}(n)^\dagger\,|0\rangle\langle 0|\,\mathscr{D}(n)\big)$. One can show that its characteristic function is $\chi_\rho(z) = e^{-|z|^2} \sum_{n=2}^\infty \frac{\cos(2nz_I)}{n^2 \log n}$, which turns out to be continuously differentiable everywhere [60]. However,

$$
\mathrm{Tr}[\rho|x|] \ge c \sum_{n=2}^\infty \frac{1}{n^2 \log n}\,2\sqrt{2}n = +\infty,
$$

which implies that ρ has no finite first-order moments (see Lemma 24).

In spite of the above counterexample, we show in Theorem 28 that at least if k is an even integer, then the existence of kth order phase space moment implies the existence of the kth order standard moment. Again, this is in total analogy with the classical case [61, Theorem 1.8.16].

Remark. Due to the above, for even k, we simply use the word *moment* in the statements of our theorems, instead of differentiating between standard moments and phase space moments.

2.4. Quantum convolution. A *beam splitter* with transmissivity $\lambda \in [0, 1]$ acting on two sets of m modes is a particular type of a passive symplectic unitary, which we express as[7]

$$U_\lambda := \exp\left[\left(\arccos\sqrt{\lambda}\right)\sum_j (a_j^\dagger b_j - a_j b_j^\dagger)\right], \tag{34}$$

where a_j and b_j $(j = 1, \ldots, m)$ are the creation operators of the first and second sets of modes, respectively. Its action on annihilation operators can be represented as follows

$$U_\lambda a_j U_\lambda^\dagger = \sqrt{\lambda}\, a_j - \sqrt{1-\lambda}\, b_j, \qquad U_\lambda b_j U_\lambda^\dagger = \sqrt{1-\lambda}\, a_j + \sqrt{\lambda}\, b_j \quad \forall j \in \{1, .., m\}. \tag{35}$$

Accordingly, displacement operators are transformed by

$$U_\lambda \left(\mathscr{D}(z) \otimes \mathscr{D}(w)\right) U_\lambda^\dagger = \mathscr{D}\left(\sqrt{\lambda}z + \sqrt{1-\lambda}\, w\right) \otimes \mathscr{D}\left(-\sqrt{1-\lambda}\, z + \sqrt{\lambda}\, w\right). \tag{36}$$

The beam splitter unitary can be used to define the following (λ-dependent) *quantum convolution*: for two m-mode quantum states ρ, σ and $\lambda \in [0, 1]$, their (λ-dependent) *quantum convolution* is given by the state $\rho \boxplus_\lambda \sigma$ which is defined according to [62] as

$$\rho \boxplus_\lambda \sigma := \mathrm{Tr}_2\left[U_\lambda(\rho \otimes \sigma)U_\lambda^\dagger\right]. \tag{37}$$

In terms of characteristic functions, this definition corresponds to

$$\chi_{\rho \boxplus_\lambda \sigma}(z) = \chi_\rho\left(\sqrt{\lambda}\, z\right) \chi_\sigma\left(\sqrt{1-\lambda}\, z\right). \tag{38}$$

It is not difficult to verify that for all symplectic unitaries V and all $\lambda \in [0, 1]$, the beam splitter unitary U_λ of (34) satisfies $[V \otimes V,\ U_\lambda] = 0$. In particular,

$$V(\rho \boxplus_\lambda \sigma)V^\dagger = (V\rho V^\dagger) \boxplus_\lambda (V\sigma V^\dagger) \tag{39}$$

for any state σ. Also, using (35) it can be shown that the mean photon number of a quantum convolution is just the convex combination of those of the input states, i.e.

$$\mathrm{Tr}\left[(\rho \boxplus_\lambda \sigma)H_m\right] = \lambda\, \mathrm{Tr}\left[\rho H_m\right] + (1 - \lambda)\, \mathrm{Tr}\left[\sigma H_m\right], \tag{40}$$

where the canonical Hamiltonian is defined by (14).

For all m-mode quantum states σ and all $\lambda \in [0, 1]$, we can use the corresponding convolution to define a quantum channel $\mathscr{N}_{\sigma,\lambda} : \mathcal{T}_1(\mathcal{H}_1) \to \mathcal{T}_1(\mathcal{H}_1)$, whose action is given by

$$\mathscr{N}_{\sigma,\lambda}(\rho) := \rho \boxplus_\lambda \sigma. \tag{41}$$

When $\sigma = \tau_N$ is a thermal state (with mean photon number N), the channel $\mathscr{N}_{\tau_N,\lambda} =: \mathscr{E}_{N,\lambda}$ is called a *thermal attenuator channel*. Its action, obtained by combining (38) and (30), is given by

$$\mathscr{E}_{N,\lambda} : \chi_\rho(z) \longmapsto \chi_{\mathscr{E}_{N,\lambda}(\rho)}(z) := \chi_\rho\left(\sqrt{\lambda}\, z\right) e^{-\frac{1-\lambda}{2}(2N+1)|z|^2}. \tag{42}$$

For the thermal attenuator channel, the energy-constrained classical capacity (defined in (12)) can be shown to reduce to can be shown to be given by [37,38]

$$\mathscr{C}\left(\mathscr{E}_{N,\lambda}, E\right) = g\left(\lambda E + (1 - \lambda)N\right) - g\left((1 - \lambda)N\right), \tag{43}$$

[7] Tensor products are omitted here.

where g is given by (29).

In what follows, we will be interested in the symmetric quantum convolutions $\rho_1 \boxplus \cdots \boxplus \rho_n$, iteratively defined for a positive integer n and states $\rho_1, \ldots, \rho_n$, by the relations $\rho \boxplus \sigma := \rho \boxplus_{1/2} \sigma$ and

$$\rho_1 \boxplus \cdots \boxplus \rho_n := (\rho_1 \boxplus \cdots \boxplus \rho_{n-1}) \boxplus_{1-1/n} \rho_n . \tag{44}$$

We will also use the shorthand

$$\rho^{\boxplus n} := \underbrace{\rho \boxplus \cdots \boxplus \rho}_{n \text{ times}} . \tag{45}$$

In terms of characteristic and Wigner functions, we can also write

$$\chi_{\rho_1 \boxplus \cdots \boxplus \rho_n}(z) = \chi_{\rho_1}\left(z/\sqrt{n}\right) \ldots \chi_{\rho_n}\left(z/\sqrt{n}\right), \tag{46}$$

$$W_{\rho_1 \boxplus \cdots \boxplus \rho_n}(z) = n^m \left(W_{\rho_1} \star \cdots \star W_{\rho_n}\right)\left(\sqrt{n}z\right). \tag{47}$$

Here, $\star$ denotes convolution, which is defined for n functions $f_1, \ldots, f_n : \mathbb{C}^m \to \mathbb{R}$ by

$$(f_1 \star \cdots \star f_n)(x) := \int d^m y_1 \ldots d^m y_{n-1} \, f_1(y_1) \ldots f_{n-1}(y_{n-1}) f_n\left(x - y_1 - \ldots - y_{n-1}\right) \tag{48}$$

Equation (46) shows that the quantum characteristic function of the symmetric quantum convolution satisfies the same scaling property as a sum of classical i.i.d. (independent and identically distributed) random variables. The important special case $\rho_i \equiv \rho$ of (46) for all $i \in \{i, 2, \ldots, n\}$, on which we will focus most of our efforts, reads

$$\chi_{\rho^{\boxplus n}}(z) = \chi_\rho\left(z/\sqrt{n}\right)^n . \tag{49}$$

Iterating (39), using (44), shows that

$$V\left(\rho^{\boxplus n}\right) V^\dagger = (V\rho V^\dagger)^{\boxplus n} \tag{50}$$

holds for all symplectic unitaries V.

3. Cushen and Hudson's Quantum Central Limit Theorem

In [2], Cushen and Hudson proved the following quantum mechanical analogue of the central limit theorem, which is the starting point of our study.

Theorem 3 [2, Theorem 1] . *Let $\rho \in \mathcal{D}(\mathcal{H}_m)$ be a centred m-mode quantum state with finite second moments. Then the sequence $(\rho^{\boxplus n})_{n \in \mathbb{N}}$ converges weakly to the Gaussian state ρ_G of same first and second moments as ρ:*

$$\mathrm{Tr}\left[\rho^{\boxplus n} X\right] \xrightarrow[n \to \infty]{} \mathrm{Tr}\left[\rho_G X\right], \qquad \forall X \in \mathcal{B}(\mathcal{H}), \tag{51}$$

where $\mathcal{B}(\mathcal{H}_m)$ is the set of bounded operators on $\mathcal{H}_m$.

Remark. The state ρ_G is commonly called the *Gaussification* of ρ.

In fact, the proof of Theorem 3 relies on the equivalence between weak convergence of states and pointwise convergence of their characteristic functions. More precisely, the following holds:

Lemma 4 ([27, Lemma 4.3] and [28, Lemma 4]). *Let $(\rho_n)_{n\in\mathbb{N}}$ be a sequence of density operators on $\mathcal{H}_m$. The following are equivalent:*

- *$(\rho_n)_{n\in\mathbb{N}}$ converges to a density operator in the weak operator topology, namely, it holds that $\lim_{n\to\infty}\langle x|\rho_n|y\rangle = \langle x|\rho|y\rangle$ for all $|x\rangle, |y\rangle \in \mathcal{H}_m$;*
- *$(\rho_n)_{n\in\mathbb{N}}$ converges in trace distance to a trace class operator;*
- *the sequence $(\chi_{\rho_n})_{n\in\mathbb{N}}$ of characteristic functions converges pointwise to a function that is continuous at 0.*

Together, the above lemma and Theorem 3 allow us to conclude the following seemingly stronger convergence:

Theorem 5. *Under the assumptions of Theorem 3, we have that*

$$\lim_{n\to\infty} \left\| \rho^{\boxplus n} - \rho_G \right\|_1 = 0. \tag{52}$$

4. Main Results

The main objective of this paper is to refine Theorem 5 of the previous section in the following directions:

- First, in the case in which the state ρ satisfies the conditions of the Cushen–Husdon theorem, we provide quantitative bounds on the rate at which the sequence of states $(\rho^{\boxplus n})_{n\in\mathbb{N}}$ converges to ρ_G, under the assumption of finiteness of certain phase space moments of ρ. We also show how finiteness of phase space moments is implied by finiteness of the corresponding standard moments, the latter having the advantage of being a more easily verifiable condition. Moreover, we show that finiteness of even integer phase space moments implies finiteness of even integer standard moments (Sect. 4.1).
- Secondly, we provide an example to show that the assumption that the second moments be finite in the Cushen–Hudson theorem cannot be weakened (Sect. 4.2).
- Thirdly, we extend our results to the non-i.i.d. setting, i.e. we consider a scaling in the quantum convolution different from (44). This allows us to analyse the propagation of states through cascades of beam splitters with varying transmissivities (Sect. 4.3).
- Finally, we provide a precise asymptotic analysis of the behaviour of quantum characteristic functions at zero and at infinity (Sect. 4.4).

4.1. Quantitative bounds in the QCLT. In this section, we state our results on rates of convergence in the Cushen–Hudson quantum central limit theorem. We call them quantum Berry–Esseen theorems, as is customary in the literature. Our first theorem provides convergence rates $\mathscr{O}\left(n^{-1/2}\right)$ in the quantum central limit theorem under a fourth-order moment condition. The rate of convergence is boosted to $\mathscr{O}\left(n^{-1}\right)$ if the third derivative of the characteristic function at zero vanishes:

Theorem 6 (Quantum Berry–Esseen theorem; High regularity). *Let ρ be a centred m-mode quantum state with finite fourth-order phase space moments. Then, the convergence in the quantum central limit theorem in Hilbert–Schmidt norm satisfies*

$$\left\| \rho^{\boxplus n} - \rho_G \right\|_2 = \mathscr{O}_{M_4'}\left(n^{-1/2}\right).$$

Here, $M_4' = M_4'(\rho, \varepsilon)$ is the moment defined in (33), and $\varepsilon > 0$ is sufficiently small. Moreover, if $D^3\chi_\rho(0) = 0$ then the convergence is at least with rate $\mathscr{O}_{M_4'}\left(n^{-1}\right)$.

The proof of Theorem 6 is provided in Sect. 6. In the next Theorem, we weaken the assumption on the moments of the state ρ, which leads to a slower rate of convergence.

Theorem 7 (Quantum Berry–Esseen theorem; Low regularity). *Let ρ be a centred m-mode quantum state with finite $(2 + \alpha)$-order phase space moments, where $\alpha \in (0, 1]$. The convergence in the quantum central limit theorem in Hilbert–Schmidt norm is given by*

$$\|\rho^{\boxplus n} - \rho_G\|_2 = \mathscr{O}_{M'_{2+\alpha}}\left(n^{-\alpha/2}\right).$$

Here, $M'_{2+\alpha} = M'_{2+\alpha}(\rho, \varepsilon)$ is the phase space moment defined in (33), and $\varepsilon > 0$ is sufficiently small.

The proof of Theorem 7 is provided in Sect. 6. The variable α allows us to obtain a convergence rate under the assumption of finiteness of phase space moments of order all the way down to 2 (excluded), which is the assumption required in the Cushen–Hudson QCLT. The above results can further be used to find convergence rates in other, statistically more relevant, distance measures:

Corollary 8 (Convergence in trace distance and relative entropy). *Assume that an m-mode quantum state ρ has finite third-order phase space moments. Then,*

$$\left\|\rho^{\boxplus n} - \rho_G\right\|_1 = \mathscr{O}_{M'_3}\left(n^{-\frac{1}{2(m+1)}}\right), \qquad D\left(\rho^{\boxplus n} \big\| \rho_G\right) = \mathscr{O}_{M'_3}\left(n^{-\frac{1}{2(m+1)}}\right),$$

where $M'_3 = M'_3(\rho, \varepsilon)$ is defined in (33), and $\varepsilon > 0$ is sufficiently small. The above rates are replaced by $\mathscr{O}_{M'_{2+\alpha}}\left(n^{-\alpha/(2m+2)}\right)$ when ρ only satisfies the conditions of Theorem 7.

The proof of this Corollary is given in Sect. 6.

Remark. (Condition on the existence of moments). The error bounds in Theorems 6 and 7 are stated in terms of assumptions on the phase space moments M'_k given by (33), of the state. It is possible to bound the phase space moments M'_k directly in terms of the standard moments M_k defined in (31). This is stated in the following Theorem, whose proof is given in Appendices A–C

Theorem 9. *Let $k \in [0, \infty)$, m a positive integer, and $\varepsilon > 0$ be given. Then every m-mode quantum state with finite standard moments of order up to k also has finite phase space moments of the same order. More precisely, there is a constant $c_{k,m}(\varepsilon) < \infty$ such that*

$$M'_k(\rho, \varepsilon) = \left\|\chi_\rho\right\|_{C^k(B(0,\varepsilon))} \leq c_{k,m}(\varepsilon) M_k(\rho).$$

Conversely, if the characteristic function is 2k times totally differentiable at $z = 0$ for some integer k, then the $2k^{th}$ standard moment is finite as well.

The importance of Theorem 9 for us comes from the fact that most of our proofs rest upon local differentiability properties of the characteristic function. While mathematically useful, such properties have no direct physical meaning and may be hard to verify in practice. Instead, the condition of finiteness of higher-order standard moments, as given in Definition 1, bears a straightforward physical meaning, related to the properties of the photon number distribution of the state, and is often easier to verify.

The key to proving Theorem 9 for fractional k lies in an interpolation argument. To state it precisely, we briefly recall some basic facts about real interpolation theory (see

[63] for more details): given two Banach spaces $(\mathcal{X}_0, \|.\|_{\mathcal{X}_0})$ and $(\mathcal{X}_1, \|.\|_{\mathcal{X}_1})$, and a parameter $0 \leq \theta \leq 1$, define the *K-function* as follows:

$$K(t, X) := \inf_{X = X_0 + X_1} \left(\|X_0\|_{\mathcal{X}_0} + t\|X_1\|_{\mathcal{X}_1} \right) \quad \forall t > 0, \tag{53}$$

and derive from this the function $\Phi_\theta(K(X)) = \sup_{t>0} t^{-\theta} K(t, X)$. The real interpolation spaces, parametrised by $\theta \in (0, 1)$, are then defined as

$$(\mathcal{X}_0, \mathcal{X}_1)_\theta := \{ X \in \mathcal{X}_0;\ \Phi_\theta(K(X)) < \infty \} \text{ with norm } \|X\|_{(\mathcal{X}_0, \mathcal{X}_1)_\theta} = \Phi_\theta(K(X)).$$

Now, given two couples of Banach spaces $\mathcal{X}_0, \mathcal{X}_1$ and $\mathcal{Y}_0, \mathcal{Y}_1$, and a map $\Psi : \mathcal{X}_0 + \mathcal{X}_1 \to \mathcal{Y}_0 + \mathcal{Y}_1$ such that $\Psi : \mathcal{X}_0 \to \mathcal{Y}_0$ and $\Psi : \mathcal{X}_1 \to \mathcal{Y}_1$ are bounded, the map $\Psi : (\mathcal{X}_0, \mathcal{X}_1)_\theta \to (\mathcal{Y}_0, \mathcal{Y}_1)_\theta$ is bounded and:

$$\|\Psi : (\mathcal{X}_0, \mathcal{X}_1)_\theta \to (\mathcal{Y}_0, \mathcal{Y}_1)_\theta\| \leq \|\Psi : \mathcal{X}_0 \to \mathcal{Y}_0\|^{1-\theta} \|\Psi : \mathcal{X}_1 \to \mathcal{Y}_1\|^\theta.$$

We want to apply this to the map $\rho \mapsto \chi_\rho$.

The following interpolation result for density operators then holds:

Proposition 10. *Let* $k_1 \geq k_0 \geq 0$ *be real numbers. The* m-*mode bosonic Sobolev spaces* $\mathscr{W}^{k_0, 1}(\mathcal{H}_m)$ *and* $\mathscr{W}^{k_1, 1}(\mathcal{H}_m)$ *form a compatible couple such that for any* m-*mode quantum state* ρ *and* $\theta \in (0, 1)$ *the real interpolation norm satisfies*

$$\|\rho\|_{(\mathscr{W}^{k_0, 1}(\mathcal{H}_m), \mathscr{W}^{k_1, 1}(\mathcal{H}_m))_{\theta, \infty}} \leq \|\rho\|_{\mathscr{W}^{(1-\theta)k_0 + \theta k_1, 1}(\mathcal{H}_m)}.$$

The proof of Proposition 10 is stated in Appendix B.

4.2. Optimality of convergence rates and necessity of finite second moments in the QCLT.

The results stated in the previous section lead naturally to the following questions:

(i) Can the assumption of finiteness of second moments in the Cushen–Hudson theorem be weakened?

(ii) Are the convergence rates of Theorems 6 and 7 and Corollary 8 optimal?

We start by answering the first question in the negative: there exists a state with finite moments of all orders $2(1 - \delta)$ (for $\delta > 0$) for which neither Theorem 3 nor Theorem 5 holds.

Proposition 11. *Consider the one-mode state* $\psi_f := |\psi_f\rangle\langle\psi_f|$ *with wave function*

$$f(x) := \frac{1}{\sqrt{2}} \frac{1}{(1 + x^2)^{3/4}}. \tag{54}$$

Then: (a) ψ_f *is centred; (b)* $M_{2(1-\delta)}(\psi_f) = \langle\psi_f|(aa^\dagger)^{1-\delta}|\psi_f\rangle < \infty$ *for all* $\delta > 0$; *yet (c) the sequence* $\left(|\psi_f\rangle\langle\psi_f|^{\oplus n} \right)_n$ *does not converge to any quantum state. Hence, the assumption of finiteness of second moments in the Cushen–Hudson QCLT (Theorems 3 and 5) cannot be weakened.*

The proof of the above proposition is given in Sect. 7.

We now come to the second question (ii) regarding tightness of the estimates in Theorems 6 and 7 and Corollary 8. In Sect. 7 below, we study several explicit examples and provide convincing numerical evidence that our estimates are indeed tight, at least as far as the Hilbert–Schmidt convergence rates are concerned. Our findings are summarised as follows.

- We start by looking at the pure state $|\psi\rangle = (|0\rangle + |3\rangle)/\sqrt{2}$, with density matrix $\psi = |\psi\rangle\langle\psi|$ and thermal Gaussification $\psi_G = \tau_{3/2}$. Our findings indicate that $\left\| \psi^{\boxplus n} - \psi_G \right\|_2 \sim c\, n^{-1/2}$, in the sense that the ratio between the two sides tends to 1 as $n \to \infty$, for some absolute constant c (Example 5 and Fig. 4). Hence, the $\mathcal{O}(n^{-1/2})$ convergence rate of Theorem 7 is attained.

- Next, we focus on the second estimate of Theorem 6, and show that it is also tight. Namely, we compute the differences $\left\| \psi^{\boxplus n} - \psi_G \right\|_\zeta$ for the simple case of a single-photon state $\psi = |1\rangle\langle 1|$ and for $\zeta = 1, 2$, and find numerical evidence that again $\left\| \psi^{\boxplus n} - \psi_G \right\|_\zeta \sim c\, n^{-1}$ for some absolute constant c (Example 4 and Fig. 4). This shows that the $\mathcal{O}(n^{-1})$ convergence rate stated in Theorem 6, under the assumption that $D^3 \chi_\rho(0) = 0$, is also attained.

4.3. Applications to capacity of cascades of beam splitters with non-Gaussian environment. We now discuss applications of our results to the study of channels that arise naturally in the analysis of lossy optical fibres. We model a physical fibre of overall transmissivity λ as a cascade of n beam splitters, in each of which the signal state ω is mixed via an elementary beam splitter of transmissivity $\lambda^{1/n}$ with a fixed state ρ, modelling the environmental noise (Fig. 2). Each step corresponds to the action of the channel $\mathcal{N}_{\rho,\,\lambda^{1/n}} : \omega \mapsto \omega \boxplus_{\lambda^{1/n}} \rho$ (cf. the definition (41)), so that the whole cascade can be represented by the n-fold composition $\mathcal{N}^n_{\rho,\,\lambda^{1/n}}$. Note that this is in general a non-Gaussian channel, albeit it is Gaussian *dilatable* [28,64]. We are interested in the asymptotic expression of the output state $\mathcal{N}^n_{\rho,\,\lambda^{1/n}}(\omega)$ as the number n tends to infinity, as a function of the input state ω. In other words, we want to study the asymptotic channel $\lim_{n\to\infty} \mathcal{N}^n_{\rho,\,\lambda^{1/n}}$.

At this point, it should not come as a surprise that such a channel exists and coincides with $\mathcal{N}_{\rho_G,\,\lambda}$.

Before we see why, let us justify why the above model may be relevant to applications. The recently flourishing field of integrated quantum photonics sets as its goal that of implementing universal quantum computation on miniaturised optical chips [29,30,65, 66]. A quantum channel that runs across such a circuit is susceptible to noise generated by other active elements of the same circuit, e.g. single-photon sources. While we expect such noise to be far from thermal, it may become so in the limit $n \to \infty$ of many interactions. In a regime where n is finite, albeit large, our setting will thus be the appropriate one. The forthcoming Corollary 13 allows us to study the classical and quantum capacity of the effective channel in such a regime.

Let us note in passing that the cascade architecture we are investigating now, in spite of some apparent resemblance, is different from that depicted in Fig. 1b. While we regard the former as more operationally motivated, the latter is mathematically convenient, as the transmissivities are tuned in such a way as to yield the symmetric convolution $\rho^{\boxplus n}$ at the output.

Theorem 12 (Approximation of thermal attenuators channels by cascades of beam splitters). *Let ρ be a centred m-mode quantum state with finite third-order phase space moments M'_3, cf. (33), and denote by ρ_G its Gaussification. Then,*

$$\left\| \mathcal{N}^n_{\rho,\,\lambda^{1/n}} - \mathcal{N}_{\rho_G,\,\lambda} \right\|_\diamond = \mathcal{O}_{M'_3}\left(n^{-\frac{1}{2(m+1)}} \right) ,$$
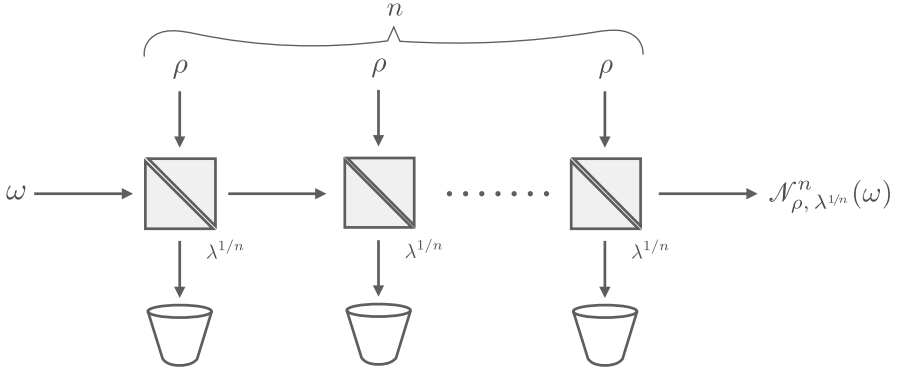
Fig. 2. An input state ω enters an optical fibre modelled by a cascade of n beam splitters with equal transmissivities $\lambda^{1/n}$ and environment states ρ

where $\| \cdot \|_{\diamond}$ stands for the diamond norm (11).

One can further make use of the recently derived continuity bounds under input energy constraints [33–36] in order to find bounds on capacities of the cascade channel $\mathcal{N}^n_{\rho, \lambda^{1/n}}$ in the physically relevant case where the Gaussification ρ_G of ρ is a thermal state.[8]

Corollary 13. *Consider a single-mode quantum state ρ with finite third-order phase space moments M'_3 (cf. (33)) and thermal Gaussification $\rho_G = \tau_N$ as in (27). Then, for $\lambda \in [0, 1]$, mean photon number $N := \mathrm{Tr}\left[\rho\, a^\dagger a\right] < \infty$, and some input energy $E > 0$, the energy-constrained classical and quantum capacity of the cascade channel $\mathcal{N}^n_{\rho, \lambda^{1/n}}$ relative to the canonical Hamiltonian $a^\dagger a$ satisfy*

$$\left| \mathscr{C}\left(\mathcal{N}^n_{\rho, \lambda^{1/n}}, E\right) - g(\lambda E + (1-\lambda)N) + g((1-\lambda)N) \right| \leq \Delta_c(n; N, M'_3, \lambda, E) \quad (55)$$

and

$$\left| \mathcal{Q}\left(\mathcal{N}^n_{\rho, \lambda^{1/n}}, E\right) - \mathcal{Q}\left(\mathscr{E}_{N, \lambda}, E\right) \right| \leq \Delta_q(n; N, M'_3, \lambda, E), \quad (56)$$

where $g(x) := (x+1)\log(x+1) - x\log x$ (as in (29)), and $\mathcal{Q}\left(\mathscr{E}_{N, \lambda}, E\right)$ is the quantum capacity of the thermal attenuator.[9]
The remainder terms are such that

$$\begin{aligned}
\Delta_c(n; N, M'_3, \lambda, E) &\leq C(M'_3)\, n^{-1/4} \log n, \\
\Delta_q(n; N, M'_3, \lambda, E) &\leq C(M'_3)\, n^{-1/8} \log n.
\end{aligned} \quad (57)$$

for some constant $C = C(M'_3)$ and all sufficiently large $n \geq n_0\left(\lambda E + (1-\lambda)N, M'_3\right)$.

The proofs of Theorem 12 and Corollary 13 are postponed to Sect. 8.

---

[8] This amounts to assuming that ρ can be brought to its so-called Williamson form (see (63) of Sect. 6) by a passive symplectic unitary only.
[9] An analytical formula for this quantity is currently not known. We report the best lower [39,45] and upper [40–44] bounds known to date in (105)–(106) and (107)–(109), respectively. These results can be used together with (56) to find bounds on $\mathcal{Q}\left(\mathcal{N}^n_{\rho, \lambda^{1/n}}, E\right)$.

4.4. New results on quantum characteristic functions. In this subsection we state our refined asymptotic analysis of the decay of quantum characteristic functions that we employ in the proofs of our main theorems. For arbitrary quantum states, we have the following asymptotic result on the quantum characteristic function at infinity. It states that the quantum characteristic function can, in absolute value, only attain the value one at zero and decays to zero at infinity. Both these properties do not hold for general classical random variables, see Sect. 5.2.

Proposition 14. *The quantum characteristic function of an m-mode quantum state ρ is a continuous function that is arbitrarily small in absolute value outside of a sufficiently large compact set, i.e. χ_ρ belongs to the Banach space $C_0(\mathbb{C}^m)$ of asymptotically vanishing functions. Moreover, for any $\varepsilon > 0$ we have*

$$\max_{z \in \mathbb{C}^m \setminus B(0,\varepsilon)} |\chi_\rho(z)| < 1, \tag{58}$$

where $B(0, \varepsilon) := \{z \in \mathbb{C}^m : |z| \leq \varepsilon\}$ denotes a Euclidean ball of radius ε centred at the origin.

The proof of Proposition 14 is given in Sect. 5.2. Interestingly, we can obtain a much more refined asymptotic on the decay of quantum characteristic functions if we assume that the state has finite second order moments.

Proposition 15. *Let ρ be an m-mode state with finite average energy $E := \mathrm{Tr}\left[\rho\left(H_m + \frac{m}{2} I\right)\right]$, where we have explicitly accounted for the non-zero energy of the vacuum state. Then, for all $z \in \mathbb{C}^m$ and all $\delta \in [0, 1]$ it holds that*

$$|\chi_\rho(z)| \leq 1 - \frac{(1-\delta)^3 \delta^{2m-1} ((2m+1)!!)^2}{6 \cdot 2^{4m} \cdot E^{2m-1}} \min\left\{|z|^2, \frac{\pi^2 \delta}{4E}\right\}.$$

The proof of Proposition 15 is given in Sect. 5.2.

5. New Results on Quantum Characteristic Functions: Proofs

Quantum characteristic functions constitute a central tool in our approach. Therefore, the first step in our path towards the quantum Berry–Esseen theorems is to prove the results stated in Sect. 4.4. The structure of this section is as follows:

- *Quantum–classical correspondence:* We derive a quantum–classical correspondence of the central limit theorems by showing that the quantum convolution of two arbitrary density operators naturally induces a classical random variable (Sect. 5.1).
- *Decay bounds:* We derive new decay estimates and asymptotic properties of the quantum characteristic function at infinity (Sect. 5.2).

5.1. Quantum–Classical Correspondence. In this section we show that the quantum convolution $\rho \boxplus \sigma$ of any two states ρ and σ has a non-negative Wigner function. While the mathematics behind this is known (see e.g. [67, Proposition (1.99)], [2, Proposition 5], and [68, Eq. (8)]), we believe that its physical implications have not been appreciated to the extent they deserve.

Lemma 16. *Let ρ and σ be arbitrary m-mode quantum states. Then the Wigner function of their convolution $\rho \boxplus \sigma$ defined by (37), with $\lambda = 1/2$, is given by*

$$W_{\rho \boxplus \sigma}(z) = \frac{2^m}{\pi^m} \, \mathrm{Tr}\left[\rho \, \mathscr{D}(\sqrt{2}z) J\sigma J \mathscr{D}(\sqrt{2}z)^\dagger\right], \qquad (59)$$

where $J := (-1)^{\sum_j a_j^\dagger a_j}$ is the unitary and self-adjoint operator that implements a phase space inversion (in the sense of Eq. (61) below). In particular,

$$W_{\rho \boxplus \sigma}(z) \geq 0 \qquad \forall\, z \in \mathbb{C}^m. \qquad (60)$$

Proof. We start by verifying that J actually corresponds to a phase space inversion, in the sense that

$$W_{J\rho J}(z) = W_\rho(-z) \qquad (61)$$

for all m-mode quantum states ρ and all $z \in \mathbb{C}^m$. This follows from the easily verified fact that $J a_j J = -a_j$ for all j, which also implies that $J\mathscr{D}(z)J = \mathscr{D}(-z)$. In fact, using (21) we find that

$$W_{J\rho J}(z) = \frac{2^m}{\pi^m} \, \mathrm{Tr}\left[\mathscr{D}(-z)J\rho J\mathscr{D}(z)\, J\right] = \frac{2^m}{\pi^m} \, \mathrm{Tr}\left[J\mathscr{D}(z)\rho\mathscr{D}(-z)\right] = W_\rho(-z).$$

We now compute

$$
\begin{aligned}
W_{\rho \boxplus \sigma}(z) &\overset{1}{=} 2^m \left(W_\rho \star W_\sigma\right)\left(\sqrt{2}z\right) \\
&\overset{2}{=} 2^m \int d^{2m}u \, W_\rho(u) \, W_\sigma\left(\sqrt{2}z - u\right) \\
&\overset{3}{=} 2^m \int d^{2m}u \, W_\rho(u) \, W_{J\sigma J}\left(u - \sqrt{2}z\right) \\
&\overset{4}{=} 2^m \int d^{2m}u \, W_\rho(u) \, W_{\mathscr{D}(\sqrt{2}z)J\sigma J\mathscr{D}(\sqrt{2}z)^\dagger}(u) \\
&\overset{5}{=} \frac{2^m}{\pi^m} \, \mathrm{Tr}\left[\rho \, \mathscr{D}(\sqrt{2}z)J\sigma J\mathscr{D}(\sqrt{2}z)^\dagger\right].
\end{aligned}
$$

In 1, we use the convolution property for the Wigner function in (47), where in 2 we just write out the convolution of several functions as in (48). In 3 we then first flip phase space variables according to (61) and use the displacement operator in 4 to translate them by $\sqrt{2}z$, cf (24). Finally, in 5 we use the quantum Plancherel identity (25) to transform the integral over Wigner functions in a trace over density operators.

(24)

The above equalities are labelled by the equation numbers corresponding to the identities that justify them. □

Remark. It is not difficult to see that $\lambda = 1/2$ is the only special value for which Lemma 16 can hold, i.e. such that $W_{\rho \boxplus_\lambda \sigma}(z) \geq 0$ for all m-mode states ρ, σ and for all $z \in \mathbb{C}^m$. To see why, consider the case where $m = 1$ and ρ, σ are the first two Fock states. The action of the beam splitter unitary on the annihilation operators, as expressed by (35), leads to the identity $|0\rangle\langle 0| \boxplus_\lambda |1\rangle\langle 1| = \lambda |0\rangle\langle 0| + (1 - \lambda)|1\rangle\langle 1|$. Using the expression for the Wigner function of Fock states [59, Eq. (4.5.31)], we see that

$$W_{|0\rangle\langle 0| \boxplus_\lambda |1\rangle\langle 1|}(z) = W_{\lambda|0\rangle\langle 0|+(1-\lambda)|1\rangle\langle 1|}(z) = \frac{2}{\pi} e^{-2|z|^2}\left(\lambda - (1 - \lambda)\left(1 - 4|z|^2\right)\right).$$

Hence, $W_{|0\rangle\langle0|\boxplus_\lambda|1\rangle\langle1|}(0) < 0$ as soon as $0 \le \lambda < 1/2$. For $1/2 < \lambda \le 1$, we arrive at the same conclusion by looking at the state $|1\rangle\langle1| \boxplus_\lambda |0\rangle\langle0| = |0\rangle\langle0| \boxplus_{1-\lambda} |1\rangle\langle1|$, obtained by sending $\lambda \mapsto 1 - \lambda$.

We proceed by showing how the above result bridges the gap between classical and quantum central limit theorems. We now fix an m-mode quantum state ρ, and notice that $\rho^{\boxplus 2n} = (\rho \boxplus \rho)^{\boxplus n}$. Consider the probability density function $f_X := W_{\rho \boxplus \rho} \ge 0$, where positivity holds by (60). Let X be a random variable with density f_X. The mean and covariance matrix of X coincide with those of $\rho \boxplus \rho$, which are in turn the same as those of ρ. Hence, at the level of Gaussifications, $f_G = W_{\rho_G}$. We write for an i.i.d. family of random variables X_i with law f_X

$$
\begin{aligned}
W_{\rho^{\boxplus 2n}}(u) &= W_{(\rho \boxplus \rho)^{\boxplus n}}(u) \\
&\overset{1}{=} n^m W_{\rho \boxplus \rho}^{\star n}\left(\sqrt{n}u\right) \\
&= n^m f_X^{\star n}\left(\sqrt{n}u\right) \\
&\overset{2}{=} f_{(X_1+\cdots+X_n)/\sqrt{n}}(u)
\end{aligned}
$$

where 1 follows from (47) and 2 follows from the change of variables $u \mapsto \sqrt{n}u$. This implies by applying the classical and quantum Plancherel identities (26) that

$$
\begin{aligned}
\left\| \rho^{\boxplus 2n} - \rho_G \right\|_2^2 &= \pi^{-m} \left\| \chi_{\rho^{\boxplus 2n}} - \chi_{\rho_G} \right\|_{L^2(\mathbb{R}^{2m})}^2 \\
&= \pi^m \left\| W_{\rho^{\boxplus 2n}} - W_{\rho_G} \right\|_{L^2(\mathbb{R}^{2m})}^2 \\
&= \pi^m \left\| f_{(X_1+\cdots+X_n)/\sqrt{n}} - f_G \right\|_{L^2(\mathbb{R}^{2m})}^2
\end{aligned}
\tag{62}
$$

which shows that the QCLT is equivalent to a certain CLT for classical i.i.d. random variables. The problem with this approach is that the right classical tool to use here would be an estimate on the rate of convergence of $(X_1 + \cdots + X_n)/\sqrt{n}$ to the normal variable X_G with respect to the L^2 norm. However, it is known that convergence fails to hold in general, and even under some finiteness of moments assumption there does not seem to be a readily available result in the literature, that is powerful enough to be successfully employed here. Therefore, we do not pursue this route further here.

5.2. Decay estimates on the quantum characteristic function. Before studying the rate of convergence in the quantum central limit theorem, we show that quantum characteristic functions have the so-called *strict non-lattice property*. To motivate this property, we start by recalling some basic properties of characteristic functions from classical probability theory.

The characteristic function χ_X^{cl} of a classical random variable X always attains the value one at zero. However, it can also attain the value one, in absolute value, at any other point. The random variables that exhibit this latter behaviour are precisely those that are *lattice-distributed*;[10] see also [69, Section 3.5]. Examples include the Dirac, Bernoulli, geometric and Poisson distributions.

---

[10] These are discrete random variables with probability distributions supported on a lattice.

Knowing that $\left|\chi_X^{\mathrm{cl}}(t)\right| < 1$ for all values $t \neq 0$ however does not imply that $\limsup_{t\to\infty} \left|\chi_X^{\mathrm{cl}}(t)\right| < 1$. This latter condition is known as the *strict non-lattice* property of a random variable. An example of a non-lattice distributed random variable which does not satisfy the strict non-lattice property is as follows.

Example 1 ([69, Section 3.5]). Consider an enumeration of the positive rationals $q_1, q_2, \ldots \in \mathbb{Q}_+$ with $q_i \leq i$ and a non-lattice random variable X defined by

$$\mathbb{P}(X = q_n) = \mathbb{P}(X = -q_n) = 2^{-(n+1)}.$$

The random variable X is then given by

$$\chi_X^{\mathrm{cl}}(t) = \mathbb{E}(e^{itX}) = \sum_{n=1}^{\infty} (e^{itq_n} + e^{-itq_n})\mathbb{P}(X = q_n)$$

which simplifies to

$$\chi_X^{\mathrm{cl}}(t) = \sum_{i=1}^{\infty} \frac{\cos(tq_i)}{2^i}.$$

Let $q_i = \frac{p_i}{r_i}$ where $p_i \in \mathbb{Z}$ and $r_i \in \mathbb{N}_0$, by considering times $t_n = 2\pi \prod_{i=1}^{n} r_i$ for arbitrarily large n, one has $\limsup_{t\to\infty} \left|\chi_X^{\mathrm{cl}}(t)\right| = 1$.

We now show the surprising fact that quantum characteristic functions do not exhibit this somewhat pathological behaviour. Instead, for any quantum state ρ it holds that $\limsup_{|z|\to\infty} \left|\chi_\rho(z)\right| = 0$, as the proof of Proposition 14 below shows.

Proof of Proposition 14. Thanks to the spectral theorem and by the dominated convergence theorem, it suffices to prove that $\lim_{|z|\to\infty} \chi_{\psi_f}(z) = 0$ for all wave function $f \in L^2(\mathbb{R}^m)$, where $\psi_f := |\psi_f\rangle\langle\psi_f|$, and $|\psi_f\rangle$ is the pure state with wave function f. We rephrase this as the requirement that χ_{ψ_f} belongs to the Banach space $C_0(\mathbb{C}^m)$, where the norm on $C_0(\mathbb{C}^m)$ is the supremum norm.

We consider smooth compactly supported functions f first. For such functions, the claim follows by combining (i) Eq. (19); (ii) the fact that f is normalised, i.e. $\int d^m x |f(x)|^2 = 1$; and (iii) the Riemann–Lebesgue lemma. For general $f \in L^2(\mathbb{R}^m)$, the result then follows by a density argument: for an arbitrary $f \in L^2(\mathbb{R}^m)$ there is a sequence of smooth and compactly supported functions $f_n \in C_c^\infty(\mathbb{R}^m)$ converging to $f \in L^2(\mathbb{R}^m)$, so that

$$\sup_{z\in\mathbb{C}^m} \left|\left(\chi_{\psi_{f_n}} - \chi_{\psi_f}\right)(z)\right|$$

$$= \sup_{z\in\mathbb{C}^m} \left|\langle\psi_{f_n}|\mathscr{D}(z)|\psi_{f_n}\rangle - \langle\psi_f|\mathscr{D}(z)|\psi_f\rangle\right|$$

$$\leq \sup_{z\in\mathbb{C}^m} \left\{ \left|\langle\psi_{f_n}|\mathscr{D}(z)|\psi_{f_n}\rangle - \langle\psi_{f_n}|\mathscr{D}(z)|\psi_f\rangle\right| + \left|\langle\psi_{f_n}|\mathscr{D}(z)|\psi_f\rangle - \langle\psi_f|\mathscr{D}(z)|\psi_f\rangle\right| \right\}$$

$$\leq 2 \left\||\psi_{f_n}\rangle - |\psi_f\rangle\right\| = 2 \|f_n - f\|_{L^2(\mathbb{R}^m)} \xrightarrow[n\to\infty]{} 0.$$

Since $C_0(\mathbb{C}^m)$ is a Banach space and $\chi_{\psi_{f_n}} \in C_0(\mathbb{C}^m)$, this implies that also the limit $\chi_{\psi_f} \in C_0(\mathbb{C}^m)$. Thus, to complete the proof of (58) it suffices to show that for every $\varepsilon > 0$ and any $z \in \mathbb{C}^m \setminus B(0, \varepsilon)$ one has that $\left|\chi_{\psi_f}(z)\right| < 1$. If this were not the case, then $|\psi_f\rangle$ would be an eigenvector of the displacement operator $\mathscr{D}(z)$. This is well known to be impossible, see e.g. [28, Lemma 10]. □

For a given state ρ and some fixed $\varepsilon > 0$, Proposition 14 tells us that there exists a constant $\eta(\rho, \varepsilon) < 1$ such that $\max_{z \in \mathbb{C}^m \setminus B(0,\varepsilon)} |\chi_\rho(z)| \leq \eta(\rho, \varepsilon)$ (cf. (58)). However, the problem of characterising the quantity $\eta(\rho, \varepsilon)$ in terms of some physically meaningful property of the state ρ remains. To this end, a natural candidate turns out to be the energy of the state. To see why this is the case, consider the following simple example.

Example 2 (Squeezed states). For every $z \in \mathbb{C}^m$ and every $\delta \in (0, 1)$ there is a (Gaussian) state ρ_G of mean photon number $\mathrm{Tr}\,[\rho_G H_m] \leq \frac{t^2}{8 \ln \frac{1}{1-\delta}} - \frac{1}{4}$ such that $|\chi_{\rho_G}(z)| \geq 1 - \delta$.

To see that this is the case, up to the application of passive symplectic unitaries, it suffices to consider the case $z = (t, 0, \ldots, 0)$, where $t > 0$. Consider the 'squeezed' Gaussian state [70–72] defined by the characteristic function

$$\chi_{\rho_G}(z) := e^{-\frac{\eta}{2}(\Re z_1)^2 - \frac{1}{2\eta}(\Im z_1)^2 - \frac{1}{2}\sum_{j>1}|z_j|^2},$$

where we set $\eta := \min\left\{\frac{2}{t^2} \ln \frac{1}{1-\delta}, 1\right\} > 0$. The mean photon number of ρ_G is well known to be given by $\mathrm{Tr}\,[\rho_G H_m] = \frac{1}{4}\left(\eta + \frac{1}{\eta}\right) - \frac{1}{2} \leq \frac{1}{4\eta} - \frac{1}{4}$, where we used the fact that $\eta \leq 1$.

The above example shows that any estimate on $\eta(\rho, \varepsilon)$ can be reasonably expected to depend on the energy. We now show that our preliminary work on the quantum–classical correspondence allows us to derive a general upper estimate for $|\chi_\rho(z)|$ at any designated point $z \in \mathbb{C}^m$ in terms of the energy of the state ρ. For this purpose, we draw upon some important mathematical results from the well-developed theory of *classical* characteristic functions. Proposition 15, whose proof we present now, implies e.g. that for a one-mode state ρ, we can take $\eta(\rho, \varepsilon) = 1 - \frac{c}{E} \min\left\{\varepsilon^2, \frac{C}{E}\right\}$, where E is the energy of ρ, and c, C are universal constants.

Proof of Proposition 15. Denoting as usual with $|z|$ the Euclidean norm (4) of $z \in \mathbb{C}^m$, we write the following chain of inequalities.

$$
\begin{aligned}
\left|\chi_\rho(z)\right| &\overset{1}{=} \left|\chi_{\rho \boxplus \rho}\left(\sqrt{2}z\right)\right|^{1/2} \\
&\overset{2}{=} \left|\chi^{\mathrm{cl}}_{X(\rho \boxplus \rho)}\left(\sqrt{2}z\right)\right|^{1/2} \\
&\overset{3}{\leq} \left(1 - \frac{2(1-\delta)^3 \delta^{2m-1} ((2m+1)!!)^2}{3 \cdot 2^{4m} \cdot E^{2m-1}} \min\left\{|z|^2, \frac{\pi^2 \delta}{4E}\right\}\right)^{1/2} \\
&\overset{4}{\leq} 1 - \frac{2(1-\delta)^3 \delta^{2m-1} ((2m+1)!!)^2}{6 \cdot 2^{4m} \cdot E^{2m-1}} \min\left\{|z|^2, \frac{\pi^2 \delta}{4E}\right\}.
\end{aligned}
$$

Here, 1 is an application of the quantum convolution rule (cf. the $n = 2$ case of (49)). In 2 we introduced the classical random vector $X(\rho \boxplus \rho)$ taking values in $\mathbb{C}^m$, with probability distribution given by the Wigner function $W_{\rho \boxplus \rho}$, which is everywhere non-negative by Lemma 16. The inequality in 3, which is the non-trivial one, follows from [61, Corollary 2.7.2]: we set $a := \sup_{z \in \mathbb{C}^m} W_{\rho \boxplus \rho}(z) \leq \frac{2^m}{\pi^m}$, with the latter estimate coming from (21), and $\alpha = 2$, so that

$$\gamma_\alpha = \gamma_2 = \int d^m z\, |z|^2\, W_{\rho \boxplus \rho}(z) \overset{(23)}{=} \mathrm{Tr}\left[(\rho \boxplus \rho)\left(H_m + \frac{m}{2}I\right)\right] \overset{(40)}{=} \mathrm{Tr}\left[\rho\left(H_m + \frac{m}{2}I\right)\right] =: E;$$

also, we substituted $m \mapsto 2m$, because our phase space $\mathbb{C}^m$ has real dimension $2m$; finally, we used the well-known formula $\Gamma(m + 1/2) = \sqrt{\pi}\, 2^{-m}(2m - 1)!!$, where $(\cdot)!!$ is the bi-factorial. Lastly, the inequality in 4 is just an application of the elementary estimate $\sqrt{1 - x} \leq 1 - \frac{x}{2}$ for $0 \leq x < 1$. $\quad \square$

Remark. In [61, Section 2.7], several other estimates for $\left|\chi_X^{\text{cl}}(t)\right|$ are derived. While we decided to stick to the simplest one, as it is already very instructive, it is possible to substantially improve over it, e.g. by resorting to non-isotropic estimates (cf. for instance [61, Theorem 2.7.14]). Notably, our quantum–classical correspondence allows us to translate *all* of these inequalities to the quantum setting, up to an irrelevant factor of $1/2$ in the associated constants (see step 4 in the above proof). We do not pursue this approach further, though we want to stress that it immediately leads to a plethora of further results.

6. Quantitative Bounds in the QCLT: Proofs

In this section, we provide proofs of the convergence rates in our quantum Berry–Esseen theorems. We also provide proofs of some of the statements in Sect. 4.3 on the convergence rate for cascades of beam splitters converging to thermal attenuator channels.

Outline of this section:. To fix ideas, we give a high-level outline of our proofs:

- *Williamson form*: We apply a suitable symplectic unitary to the state, so as to make the Hessian of its characteristic function diagonal and larger than the identity. Subsequently, we use the quantum Plancherel identity to express the difference of the convolved state and its Gaussification in Hilbert–Schmidt norm as a difference of quantum characteristic functions in L^2 norm (Sect. 6.1).
- *Local-tail decomposition*: We then split the integral of the L^2 norm of the difference of the quantum characteristic functions of the convolved state and the Gaussification of the original state into a regime around zero (Lemma 17), in which we can control the behaviour of the quantum characteristic function by its Taylor expansion, and a tail-regime in which we estimate the difference using Proposition 14. The error in the Taylor expansion is controlled by the phase space moments of the state, cf. Lemma 18.
- *Hilbert–Schmidt convergence*: We implement the above ideas to prove Theorems 6 and 7, and Proposition 22 (Sect. 6.2).
- *Trace norm and entropic convergence:* We then use the preservation of the boundedness of the second moment under quantum convolutions to obtain a quantitative estimate of convergence in trace distance, employing Markov's inequality and the Gentle Measurement Lemma [73], and in relative entropy, using entropic continuity bounds [33] (Sect. 6.3).
- *Convergence rates for cascades of beam splitters:* In the final subsection, we prove the results claimed in Sect. 4.3, namely convergence rates for cascades of beam splitters converging to thermal attenuator channels (Sect. 8).

6.1. Preliminary steps.

6.1.1. Williamson form Let ρ be a centred m-mode quantum state with finite second moments, as in the Cushen–Hudson theorem. It is known that one can find a symplectic unitary V and numbers $\nu_1, \ldots, \nu_m \geq 1$ such that

$$\rho' := V^\dagger \rho V \tag{63}$$

satisfies

$$\chi_{\rho'}(z) = 1 - \frac{1}{2} \sum_j \nu_j |z_j|^2 + o\left(|z|^2\right) \qquad (z \to 0). \tag{64}$$

With a slight abuse of terminology, we will call ρ' the *Williamson form* of ρ [74]. Bringing a state to its Williamson form allows us to assume that (i) the smallest eigenvalue of its covariance matrix is at least one. Also, (ii) the transformation in (63) does not change the first moments of the state, so that if ρ is centred then ρ' remains centred. Finally, (iii) the same unitary V brings not only ρ but also its Gaussification ρ_G to their Williamson forms simultaneously, so that

$$\chi_{\rho'_G}(z) = \exp\left[-\frac{1}{2}\sum_j \nu_j |z_j|^2\right], \qquad W_{\rho'_G}(z) = \left(\frac{2}{\pi}\right)^m \exp\left[-2\sum_j \nu_j |z_j|^2\right]. \tag{65}$$

holds as well. Thanks to the covariance of the quantum convolution with respect to symplectic unitaries (50), we see that

$$\left\|\rho^{\boxplus n} - \rho_G\right\|_2 = \left\|V^\dagger\left(\rho^{\boxplus n} - \rho_G\right)V\right\|_2$$
$$= \left\|(\rho')^{\boxplus n} - \rho'_G\right\|_2.$$

Combining this with the quantum Plancherel identity (26) yields

$$\left\|\rho^{\boxplus n} - \rho_G\right\|_2^2 = \int \frac{d^{2m}z}{\pi^m} \left|\chi_{(\rho')^{\boxplus n}}(z) - \chi_{\rho'_G}(z)\right|^2 \tag{66}$$
$$= \pi^m \int d^{2m}z \left(W_{(\rho')^{\boxplus n}}(z) - W_{\rho'_G}(z)\right)^2. \tag{67}$$

In short, when estimating any unitarily invariant distance of $\rho^{\boxplus n}$ from its limit ρ_G, we can assume without loss of generality that all states are in their Williamson forms. When the Hilbert–Schmidt norm is employed, we can compute the distance as an L^2 norm at the level of characteristic functions, or equivalently at that of Wigner functions.

6.1.2. Local-tail decomposition We continue with an important technical lemma that reduces the convergence in the quantum central limit theorem to the behaviour of the quantum characteristic function around zero.

Lemma 17. *Let ρ be an m-mode quantum state with finite second-order phase space moment. Without loss of generality, we assume that ρ is centred and in Williamson form, and that its Gaussification ρ_G has characteristic function as in (65). Then for every $\varepsilon > 0$ we have that*

$$\left\|\rho^{\boxplus n} - \rho_G\right\|_2^2 = \frac{1}{\pi^m} \int_{|z| \leq \sqrt{n}\,\varepsilon} d^{2m}z \left|\chi_\rho\left(\frac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2}\right|^2 + \mathcal{O}\left(n^{-\infty}\right) \tag{68}$$

as $n \to \infty$. If ρ has also finite third-order phase space moments, then

$$\|\rho^{\boxplus n} - \rho_{\mathrm{G}}\|_2 \leq \frac{\sqrt{m(m+1)(m+2)}}{6} \frac{\|D^3\chi_\rho(0)\|}{\sqrt{n}} + \mathcal{O}(n^{-\infty})$$

$$+ \frac{1}{\pi^{m/2}} \left(\int_{|z| \leq \sqrt{n}\,\varepsilon} d^{2m}z \left| \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \left(1 + \tfrac{1}{6\sqrt{n}} D^3\chi_\rho(0)\left(z^{\times 3}\right) \right) \right|^2 \right)^{1/2},$$

(69)

where the Fréchet derivative of χ_ρ is defined by (3).

Proof. The first identity (68) follows along the lines of the second one (69) and so we focus on verifying the latter. Using the quantum Plancherel identity (26) and the relation (46), we apply the triangle inequality and split the integration domain into two disjoint sets such that

$$\pi^{\frac{m}{2}} \|\rho^{\boxplus n} - \rho_{\mathrm{G}}\|_2 = \left(\int d^{2m}z \left| \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \right|^2 \right)^{1/2}$$

$$\leq \left(\int_{|z| \leq \sqrt{n}\,\varepsilon} d^{2m}z \left| \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \left(1 + \tfrac{1}{6\sqrt{n}} D^3\chi_\rho(0)\left(z^{\times 3}\right) \right) \right|^2 \right)^{1/2}$$

$$+ \left(\int_{|z| > \sqrt{n}\,\varepsilon} d^{2m}z \left| \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \left(1 + \tfrac{1}{6\sqrt{n}} D^3\chi_\rho(0)\left(z^{\times 3}\right) \right) \right|^2 \right)^{1/2}$$

$$+ \frac{1}{\sqrt{n}} \left(\int d^{2m}z \left| e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \left(\tfrac{1}{6} D^3\chi_\rho(0)\left(z^{\times 3}\right) \right) \right|^2 \right)^{1/2}.$$

(70)

The last term on the rightmost side of (70) can be estimated explicitly using spherical coordinates. Namely, combining the fact that the coefficients appearing in the Williamson form satisfy $\nu_j \geq 1$ with the bound $\left| D^3\chi_\rho(0)\left(z^{\times 3}\right) \right| \leq \left\| D^3\chi_\rho(0) \right\| |z|^3$, we obtain that

$$\int d^{2m}z \left| e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \left(\tfrac{1}{6} D^3\chi_\rho(0)(z^{\times 3}) \right) \right|^2 \leq \frac{\mathrm{vol}(\mathbb{S}^{2m-1}) \|D^3\chi_\rho(0)\|^2}{36} \int_0^\infty dr\, e^{-r^2} r^{2m+5}$$

$$= \frac{\Gamma(m+3)\,\mathrm{vol}(\mathbb{S}^{2m-1})}{72} \|D^3\chi_\rho(0)\|^2$$

$$= \frac{\pi^m}{36} m(m+1)(m+2) \|D^3\chi_\rho(0)\|^2,$$

where we used that $\int_0^\infty dr\, e^{-r^2} r^{2m+5} = \frac{\Gamma(m+3)}{2}$, and recalled the expression $\mathrm{vol}\left(\mathbb{S}^{N-1}\right) = \frac{2\pi^{N/2}}{\Gamma(N/2)}$ for the volume of the $(N-1)$-sphere. Furthermore, the second-to-last term in (70) can be shown to be exponentially small. In fact,

$$\int_{|z| > \sqrt{n}\,\varepsilon} d^{2m}z \left| \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \left(1 + \tfrac{1}{6\sqrt{n}} D^3\chi_\rho(0)\left(z^{\times 3}\right) \right) \right|^2$$

$$\leq 2 \int_{|z| > \sqrt{n}\,\varepsilon} d^{2m}z \left(\left| \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n \right|^2 + \left| e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \left(1 + \tfrac{|z|^3}{6\sqrt{n}} \left\| D^3\chi_\rho(0) \right\| \right) \right|^2 \right)$$

$$\leq 2 \left(\sup_{|z| > \varepsilon} \left| \chi_\rho(z) \right| \right)^{2(n-1)} \int_{|z| > \sqrt{n}\,\varepsilon} d^{2m}z \left| \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right) \right|^2$$

$$+ 2 \int_{|z|>\sqrt{n}\,\varepsilon} d^{2m}z \left| e^{-\frac{1}{2}\sum_j v_j|z_j|^2} \left(1 + \frac{|z|^3}{6\sqrt{n}} \left\| D^3 \chi_\rho(0) \right\| \right) \right|^2$$

$$\overset{3}{=} 2n^m \left(\sup_{|z|>\varepsilon} |\chi_\rho(z)| \right)^{2(n-1)} \int_{|u|>\varepsilon} d^{2m}u \, |\chi_\rho(u)|^2$$

$$+ 2 \int_{|z|>\sqrt{n}\,\varepsilon} d^{2m}z \left| e^{-\frac{1}{2}\sum_j v_j|z_j|^2} \left(1 + \frac{|z|^3}{6\sqrt{n}} \left\| D^3 \chi_\rho(0) \right\| \right) \right|^2$$

$$\overset{4}{\leq} 2n^m \left(\sup_{|z|>\varepsilon} |\chi_\rho(z)| \right)^{2(n-1)} + 2 \operatorname{vol}\left(\mathbb{S}^{2m-1}\right) \int_{\sqrt{n}\,\varepsilon}^\infty dr \, e^{-r^2} r^{2m-1} \left(1 + \frac{r^3 \| D^3 \chi_\rho(0) \|}{6\sqrt{n}} \right)^2$$

$$\overset{5}{\leq} 2n^m \left(\sup_{|z|>\varepsilon} |\chi_\rho(z)| \right)^{2(n-1)} + \frac{4\pi^m}{(m-1)!} e^{-\frac{\varepsilon^2}{2}n} \int_{\sqrt{n}\,\varepsilon}^\infty dr \, e^{-\frac{r^2}{2}} r^{2m-1} \left(1 + \frac{r^3 \| D^3 \chi_\rho(0) \|}{6\sqrt{n}} \right)^2,$$

where in 1 we use that $(a+b)^2 \leq 2(a^2 + b^2)$, in 2 we use that

$$\int |(fg)(x)|^2 \, dx \leq \sup_x |f(x)|^2 \int |(g)(x)|^2 \, dx,$$

and in (3) we changed variables in the first integral to $u := \frac{z}{\sqrt{n}}$. Finally, in 4, we used that the L^2 norm of the characteristic function is at most one and switched to spherical coordinates to compute the second integral. In 5, instead, we estimated $e^{-r^2} < e^{-\frac{\varepsilon^2}{2}n}e^{-\frac{r^2}{2}}$ for $r > \sqrt{n}\,\varepsilon$. Note that the first addend goes to zero faster than any inverse power of n for $n \to \infty$ by Proposition 14. The second decays exponentially, essentially because the integral is bounded in n (in fact, it tends to 0 as $n \to \infty$). This concludes the proof. □

The first term on the right-hand side of (69) features an explicit dependence on n, while the second decays faster than any inverse power of n. Therefore, all that is left to do is to estimate the third term, which can be done by looking at the behaviour of the characteristic function in a neighbourhood of the origin. The first step in this direction, rather unsurprisingly, involves a Taylor expansion of χ_ρ around 0. In the subsequent lemma we record various important estimates of this sort, which will play a key role in the proofs of our quantum Berry–Esseen theorems.

Lemma 18. *For $\varepsilon > 0$ and $k \in [0, \infty)$, let ρ be an m-mode state with finite phase space moments of order up to k (namely, with the notation of Definition 2, assume that $M_k'(\rho, \varepsilon) < \infty$). Then for all $z \in \mathbb{C}^m$ with $|z| \leq \sqrt{n}\,\varepsilon$ it holds that*

$$\left| \chi_\rho\left(\frac{z}{\sqrt{n}}\right) - \sum_{h=0}^{\lceil k \rceil - 1} \frac{1}{h! \, n^{h/2}} D^h \chi_\rho(0) \left(z^{\times h}\right) \right| \leq \frac{1}{\lfloor k \rfloor!} \binom{2m + \lfloor k \rfloor - 1}{\lfloor k \rfloor} M_k'(\rho, \varepsilon) \frac{|z|^k}{n^{k/2}}. \tag{71}$$

In particular, if ρ is centred and in Williamson form,

$$\left| \chi_\rho\left(\frac{z}{\sqrt{n}}\right) - 1 \right| \leq \frac{m(2m+1)}{2} M_2'(\rho, \varepsilon) \frac{|z|^2}{n}, \tag{72}$$

$$\left| \chi_\rho\left(\frac{z}{\sqrt{n}}\right) - 1 + \frac{1}{2n} \sum_j v_j |z_j|^2 \right| \leq \frac{m(2m+1)}{2} M_{2+\alpha}'(\rho, \varepsilon) \frac{|z|^{2+\alpha}}{n^{1+\frac{\alpha}{2}}}, \tag{73}$$

$$\left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) - 1 + \frac{1}{2n} \sum_j \nu_j |z_j|^2 \right| \leq \frac{m(m+1)(2m+1)}{9} \, M_3'(\rho, \varepsilon) \, \frac{|z|^3}{n^{3/2}} , \tag{74}$$

$$\left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) - 1 + \frac{1}{2n} \sum_j \nu_j |z_j|^2 - \frac{1}{6n^{3/2}} D^3 \chi_\rho(0) \left(z^{\times 3} \right) \right| \leq \frac{m(m+1)(2m+1)(2m+3)}{144}$$
$$M_4'(\rho, \varepsilon) \, \frac{|z|^4}{n^2} , \tag{75}$$

depending on what phase space moments are finite. In (73)*, we assumed that* $\alpha \in (0, 1)$.

The estimate in (71) follows immediately from using Hölder continuity of the derivative.

6.2. Proofs of convergence rates in Hilbert–Schmidt distance.
We start with the proof of Theorem 6 assuming fourth-order moments.

Proof of Theorem 6. By the discussion in Sect. 6.1.1, we can assume that ρ is in Williamson form, namely, that its characteristic function satisfies (64), with $\nu_1, \ldots, \nu_m \geq 1$. Since $M_2'(\rho, \varepsilon)$ is monotonically non-decreasing in ε, for any fixed $\mu \in (0, 2)$ we can chose $\varepsilon > 0$ small enough so that for any $z \in B := B\left(0, \sqrt{n}\, \varepsilon \right)$ it holds that

$$\frac{m(2m + 1)}{2} \, M_2'(\rho, \varepsilon) \, \frac{|z|^2}{n} \leq \frac{m(2m + 1)}{2} \, \varepsilon^2 \, M_2'(\rho, \varepsilon) \leq \frac{\mu}{2} . \tag{76}$$

Looking at (72), this implies that $2 \left| 1 - \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right| \leq \mu$. Now, for $x \in \mathbb{C}$ with $|x| < 2$ define the function

$$a(x) := -\frac{4}{x^2} \left(\log \left(1 - \frac{x}{2} \right) + \frac{x}{2} \right) = \sum_{\ell=0}^{\infty} \frac{x^\ell}{2^\ell \, \ell} . \tag{77}$$

Substituting $x = 2 \left(1 - \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right)$, we then have that

$$\left| \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right) + \left(1 - \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right) \right| = \left| - \left(1 - \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right)^2 a \left(2 \left(1 - \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right) \right) \right|$$
$$\leq \frac{m^2(2m + 1)^2}{4} \, M_2'(\rho, \varepsilon)^2 \, a(\mu) \, \frac{|z|^4}{n^2} , \tag{78}$$

where to deduce the last inequality we observed that $|x| \leq \mu$ implies that $|a(x)| \leq a(\mu)$. Then, thanks to (78) and (74), an application of the triangle inequality yields

$$\left| \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n \right) + \frac{1}{2} \sum_j \nu_j |z_j|^2 \right|$$
$$\leq n \left| \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right) + 1 - \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right| + n \left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) - 1 + \frac{1}{2} \sum_j \nu_j |z_j|^2 \right|$$
$$\leq \frac{m^2(2m + 1)^2}{4} \, M_2'(\rho, \varepsilon)^2 \, a(\mu) \, \frac{|z|^4}{n} + \frac{m(m + 1)(2m + 1)}{9} \, M_3'(\rho, \varepsilon) \, \frac{|z|^3}{\sqrt{n}} \tag{79}$$
$$\leq \frac{C_1 |z|^3}{\sqrt{n}} ,$$

where for fixed m the constant C_1 depends only on M'_3 (remember that $M'_2 \leq M'_3$ by construction). Using again (78) but now in conjunction with (75), by a swift application of the triangle inequality we see that

$$
\left| \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n \right) + \frac{1}{2} \sum_j \nu_j |z_j|^2 - \frac{1}{6\sqrt{n}} D^3 \chi_\rho(0) \left(z^{\times 3} \right) \right|
$$

$$
\leq n \left| \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right) + 1 - \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) \right| + n \left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right) - 1 + \frac{1}{2} \sum_j \nu_j |z_j|^2 - \frac{1}{6\sqrt{n}} D^3 \chi_\rho(0) \left(z^{\times 3} \right) \right|
$$

$$
\overset{2}{\leq} \frac{m^2 (2m+1)^2}{4} M'_2(\rho, \varepsilon)^2 a(\mu) \frac{|z|^4}{n} + \frac{m(m+1)(2m+1)(2m+3)}{144} M'_4(\rho, \varepsilon) \frac{|z|^4}{n}
$$

$$
\leq \frac{C_2 |z|^4}{n},
$$

(80)

where for fixed m the constant C_2 depends only on M'_4 (remember that $M'_2 \leq M'_4$ by construction). We now estimate

$$
e^{\frac{1}{2} \sum_j \nu_j |z_j|^2} \left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \left(1 + \frac{1}{6\sqrt{n}} D^3 \chi_\rho(0) \left(z^{\times 3} \right) \right) \right|
$$

$$
= \left| \exp \left(\log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n \right) + \frac{1}{2} \sum_j \nu_j |z_j|^2 \right) - \left(1 + \frac{1}{6\sqrt{n}} D^3 \chi_\rho(0) \left(z^{\times 3} \right) \right) \right|
$$

$$
\overset{1}{\leq} \left| \exp \left(\log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n \right) + \frac{1}{2} \sum_j \nu_j |z_j|^2 \right) - \left(1 + \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n \right) + \frac{1}{2} \sum_j \nu_j |z_j|^2 \right) \right|
$$

$$
+ \left| \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n \right) + \frac{1}{2} \sum_j \nu_j |z_j|^2 - \frac{1}{6\sqrt{n}} D^3 \chi_\rho(0) \left(z^{\times 3} \right) \right|
$$

(81)

$$
\overset{2}{\leq} \frac{C_1^2 |z|^6}{n} e^{C_1 |z|^3 / \sqrt{n}} + \frac{C_2 |z|^4}{n}
$$

$$
\overset{3}{\leq} \frac{1}{n} e^{\frac{1}{4} |z|^2} \left(C_1^2 |z|^6 + C_2 |z|^4 \right).
$$

Here, 1 follows simply by the triangle inequality. In 2, we (i) observed that $|e^u - (1+u)| \leq |u|^2 e^{|u|}$; (ii) operated the substitution $u = \log \left(\chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n \right) + \frac{1}{2} \sum_j \nu_j |z_j|^2$; (iii) noted that $\mathbb{R} \ni x \mapsto x^2 e^x$ is a monotonically increasing function; and (iv) used the fact – proved in (79) – that $|u| \leq \frac{C_1 |z|^3}{\sqrt{n}}$. Finally, in 3 we remembered that $|z| \leq \sqrt{n}\, \varepsilon$ and assumed that $\varepsilon > 0$ is small enough so that $\varepsilon C_1 \leq \frac{1}{4}$. Now, since $\nu_1, \ldots, \nu_m \geq 1$, we can rephrase the above estimate as

$$
\left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \left(1 + \frac{1}{6\sqrt{n}} D^3 \chi_\rho(0) \left(z^{\times 3} \right) \right) \right| \leq \frac{1}{n} e^{-\frac{1}{4} |z|^2} \left(C_1^2 |z|^6 + C_2 |z|^4 \right). \quad (82)
$$

Upon integration, (82) naturally yields an upper bound for the second term on the right-hand side of (69). We obtain that

$$\int_{|z|\le\sqrt{n}\,\varepsilon} d^{2m}z\,\left|\chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j v_j|z_j|^2}\left(1+\frac{1}{6\sqrt{n}}D^3\chi_\rho(0)\left(z^{\times 3}\right)\right)\right|^2$$

$$\le \frac{1}{n^2}\int_{|z|\le\sqrt{n}\,\varepsilon} d^{2m}z\, e^{-\frac{1}{2}|z|^2}\left(C_1^2|z|^6 + C_2|z|^4\right)^2$$

$$\le \frac{1}{n^2}\int d^{2m}z\, e^{-\frac{1}{2}|z|^2}\left(C_1^2|z|^6 + C_2|z|^4\right)^2$$

$$\overset{4}{=} \frac{1}{n^2}\operatorname{vol}\left(\mathbb{S}^{2m-1}\right)\int_0^\infty dr\, r^{2m-1}e^{-\frac{1}{2}r^2}\left(C_1^2 r^6 + C_2 r^4\right)^2 \tag{83}$$

$$\overset{5}{=} \frac{1}{n^2}2^{m+3}\operatorname{vol}\left(\mathbb{S}^{2m-1}\right)\int_0^\infty ds\, e^{-s}\left(4C_1^4 s^{m+5} + 4C_1^2 C_2 s^{m+4} + C_2^2 s^{m+3}\right)$$

$$\overset{6}{=} \frac{2^{m+4}\pi^m m(m+1)(m+2)(m+3)}{n^2}\left(4C_1^4(m+4)(m+5) + 4C_1^2 C_2(m+4) + C_2^2\right)$$

$$\overset{7}{\le} \frac{C_3^2}{n^2}.$$

The justification of the above steps goes as follows: in 4 we switched to spherical coordinates; in 5 we performed the change of variables $s := \frac{1}{2}r^2$; in 6 we computed the gamma integrals, also remembering that $\operatorname{vol}\left(\mathbb{S}^{2m-1}\right) = \frac{2\pi^m}{(m-1)!}$; finally, the constant C_3 introduced in 7 depends – for fixed m – only on M_4' (note that $M_3' \le M_4'$ by construction). The proof of the first claim is completed once one inserts (83) into (69). In particular, if $D^3\chi_\rho(0) = 0$ we see that the convergence rate is $\mathcal{O}_{M_4'}\left(n^{-1}\right)$. This proves also the second claim. $\square$

We continue with the proof of the low-regularity QCLT that assumes finiteness of phase space moments of order up to $2+\alpha$, for some $\alpha \in (0,1]$.

Proof of Theorem 7. We just deal with the case where $\alpha \in (0,1)$. As above, we start by fixing $\mu \in (0,2)$ and choosing a sufficiently small $\varepsilon > 0$ so that for any $z \in B := B\left(0,\sqrt{n}\,\varepsilon\right)$ the inequality (76) holds. By a similar estimate as in (79), but now leveraging (73) instead of (74), we have that for any $z \in B\left(0,\sqrt{n}\,\varepsilon\right)$

$$\left|\log\left(\chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n\right) + \frac{1}{2}\sum_j v_j|z_j|^2\right|$$

$$\le n\left|\log\left(\chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)\right) + 1 - \chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)\right| + n\left|\chi_\rho\left(\tfrac{z}{\sqrt{n}}\right) - 1 + \frac{1}{2}\sum_j v_j|z_j|^2\right| \tag{84}$$

$$\le \frac{m^2(2m+1)^2}{4}M_2'(\rho,\varepsilon)^2 a(\mu)\frac{|z|^4}{n} + \frac{m(2m+1)}{2}M_{2+\alpha}'(\rho,\varepsilon)\frac{|z|^{2+\alpha}}{n^{\alpha/2}}$$

$$\le \frac{C_4|z|^{2+\alpha}}{n^{\alpha/2}},$$

where the constant C_4 introduced in the last line depends only on $M_{2+\alpha}'$ (note that $M_2' \le M_{2+\alpha}'$).

$$e^{\frac{1}{2}\sum_j v_j|z_j|^2}\left|\chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n - e^{-\frac{1}{2}\sum_j v_j|z_j|^2}\right| = \left|\exp\left(\log\left(\chi_\rho\left(\tfrac{z}{\sqrt{n}}\right)^n\right) + \frac{1}{2}\sum_j v_j|z_j|^2\right) - 1\right|$$

$$\overset{1}{\le} \frac{C_4|z|^{2+\alpha}}{n^{\alpha/2}}e^{C_4|z|^{2+\alpha}/n^{\alpha/2}}$$

$$\overset{2}{\le} \frac{C_4|z|^{2+\alpha}}{n^{\alpha/2}}e^{\frac{1}{4}|z|^2}.$$

Here, in 1 we used the elementary estimate $|e^u - 1| \leq |u|e^{|u|}$, together with the observation that the function $\mathbb{R} \ni x \mapsto xe^x$ is monotonically increasing. In 2 we used the fact that $|z| \leq \sqrt{n}\,\varepsilon$, and chose $\varepsilon > 0$ sufficiently small so that $\varepsilon^\alpha C_4 \leq \frac{1}{4}$. Combining the above estimate with the fact that $\nu_1, \ldots, \nu_m \geq 1$ yields

$$\left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right| \leq \frac{C_4 |z|^{2+\alpha}}{n^{\alpha/2}} \, e^{-\frac{1}{4}|z|^2} , \tag{85}$$

which upon integration in turn leads to

$$
\begin{aligned}
\int_{|z| \leq \sqrt{n}\,\varepsilon} d^{2m}z \left| \chi_\rho \left(\tfrac{z}{\sqrt{n}} \right)^n - e^{-\frac{1}{2}\sum_j \nu_j |z_j|^2} \right|^2
&\leq \frac{C_4^2}{n^\alpha} \int_{|z| \leq \sqrt{n}\,\varepsilon} d^{2m}z \, e^{-\frac{1}{2}|z|^2} |z|^{4+2\alpha} \\
&\leq \frac{C_4^2}{n^\alpha} \int d^{2m}z \, e^{-\frac{1}{2}|z|^2} |z|^{4+2\alpha} \\
&\overset{3}{=} \frac{C_4^2}{n^\alpha} \, \mathrm{vol}\left(\mathbb{S}^{2m-1} \right) \int_0^\infty dr \, e^{-\frac{1}{2}r^2} r^{2m+3+2\alpha} \\
&\overset{4}{=} \frac{2^{m+\alpha+2} \, C_4^2 \, \pi^m \, \Gamma(m+\alpha+1)}{n^\alpha \, (m-1)!} \\
&\overset{5}{\leq} \frac{C_5^2}{n^\alpha} .
\end{aligned}
\tag{86}
$$

Here, in 3 we switched to spherical coordinates; in 4 we operated the change of variables $s := \frac{1}{2}r^2$ and computed the gamma integrals; the constant introduced in 5 depends, for fixed α, only on $M'_{2+\alpha}$. Inserting (86) into the right-hand side of (69) completes the proof. $\square$

6.3. Convergence in trace distance and relative entropy. In this section, we further use the assumption of finiteness of the second moments of the state in order to find convergence rates in trace distance and in relative entropy.

Proof of Corollary 8. The hypothesis implies in particular that ρ has finite phase space moments of the second order. By Theorem 28, this amounts to saying that ρ has also finite standard moments of the second order, that is, that $\mathrm{Tr}\,[\rho H_m] \leq E < \infty$. Iterating (40) and passing to the limit, we see that in fact

$$\mathrm{Tr}\left[\rho^{\boxplus n} H_m \right] = \mathrm{Tr}\,[\rho_{\mathrm{G}} H_m] = \mathrm{Tr}\,[\rho H_m] \leq E .$$

Now, for any $E' > 0$, denote by $P_{E'}$ the projection onto the finite dimensional subspace generated by the eigenvectors of the canonical Hamiltonian H_m of eigenvalue less than E'. Then, by Markov's inequality, for any $\varepsilon > 0$,

$$\mathrm{Tr}\left[\rho^{\boxplus n} P_{E/\varepsilon} \right] , \; \mathrm{Tr}\left[\rho_{\mathrm{G}} P_{E/\varepsilon} \right] \geq 1 - \varepsilon .$$

From the so-called 'gentle measurement lemma' [73, Lemma 9], we have that

$$\left\| \rho^{\boxplus n} - P_{E/\varepsilon}\, \rho^{\boxplus n} P_{E/\varepsilon} \right\|_1 , \; \left\| \rho_{\mathrm{G}} - P_{E/\varepsilon}\, \rho_{\mathrm{G}}\, P_{E/\varepsilon} \right\|_1 \leq 2\sqrt{\varepsilon} .$$

Then,

$$\left\| \rho^{\boxplus n} - \rho_G \right\|_1 \leq \left\| \rho^{\boxplus n} - P_{E/\varepsilon}\, \rho^{\boxplus n}\, P_{E/\varepsilon} \right\|_1 + \left\| P_{E/\varepsilon} \left(\rho^{\boxplus n} - \rho_G \right) P_{E/\varepsilon} \right\|_1 + \left\| P_{E/\varepsilon}\, \rho_G\, P_{E/\varepsilon} - \rho_G \right\|_1$$

$$\leq 4\sqrt{\varepsilon} + \left\| P_{E/\varepsilon} \left(\rho^{\boxplus n} - \rho_G \right) P_{E/\varepsilon} \right\|_1$$

$$\leq 4\sqrt{\varepsilon} + \left\| P_{E/\varepsilon} \right\|_2 \left\| P_{E/\varepsilon} \left(\rho^{\boxplus n} - \rho_G \right) P_{E/\varepsilon} \right\|_2$$

$$\leq 4\sqrt{\varepsilon} + (E/\varepsilon)^{m/2} \left\| \rho^{\boxplus n} - \rho_G \right\|_2$$

The result follows after optimising over $\varepsilon > 0$. In particular, if $\left\| \rho^{\boxplus n} - \rho_G \right\|_2 = \mathcal{O}\left(n^{-\alpha} \right)$, we find that $\left\| \rho^{\boxplus n} - \rho_G \right\|_1 = \mathcal{O}\left(n^{-\frac{\alpha}{m+1}} \right)$.

We now turn to the proof of the convergence in relative entropy. Observe that, since $\rho^{\boxplus n}$ and ρ_G share the same first and second moments, $\mathrm{Tr}\left[\rho^{\boxplus n} \log \rho_G \right] = \mathrm{Tr}\left[\rho_G \log \rho_G \right]$ and thus $D\left(\rho^{\boxplus n} \big\| \rho_G \right) = S\left(\rho_G \right) - S\left(\rho^{\boxplus n} \right)$. The result follows directly from [33, Lemma 18]. □

7. Optimality of Convergence Rates and Necessity of Finite Second Moments in the QCLT: Proofs

In this section we discuss the optimality of our results in two different directions:

- First, we provide examples of states ρ that do not have finite second moments and for which $\rho^{\boxplus n}$ does not converge to any quantum state. This shows the necessity of the assumptions on finite second moments in the Cushen–Hudson Theorem (Sect. 7.1).
- Secondly, we provide examples of explicit states which saturate our convergence rates in Theorems 6 and 7 (Sect. 7.2).

7.1. Failure of convergence for states with unbounded energy. We now show that the assumption of finiteness of second moments in Theorems 3 and 5 cannot be weakened, e.g. by replacing it with finiteness of some lower-order moments. Some examples of states with undefined moments that do not satisfy Theorems 3 and 5 can be obtained by drawing inspiration from probability theory. For instance, remembering that a classical Cauchy-distributed random variable does not satisfy the central limit theorem, we construct the following example.

Example 3 (Cauchy-based wave function). Consider the pure state $|\psi_f\rangle$ with wave function $f(x) := \frac{1}{\sqrt{\pi}} \frac{1}{x+i}$. The characteristic function of this state can be computed thanks to (19), which in this case evaluates to

$$\chi_{|\psi_f\rangle\langle\psi_f|}(z) = \frac{\sqrt{2}\, e^{-|z_I|\left(\sqrt{2} + i z_R \right)}}{\sqrt{2} + i z_R}.$$

The absolute value of this characteristic function is illustrated in Fig. 3.

We then find the pointwise limit $\lim_{n\to\infty} \chi_{|\psi_f\rangle\langle\psi_f|}\left(z/\sqrt{n} \right)^n = \delta_{z,0}$ which again is not continuous at 0 and hence is not the characteristic function of any quantum state.

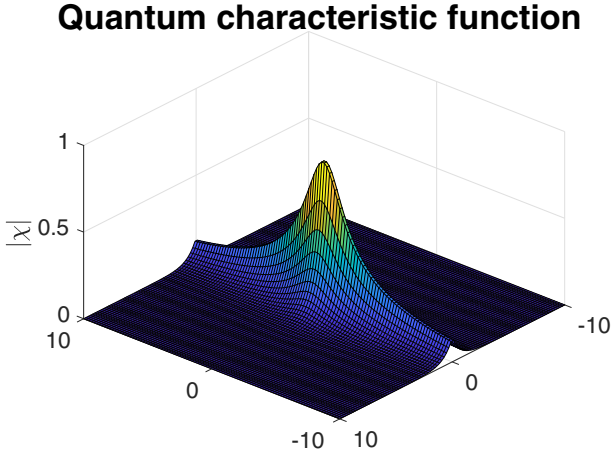## Quantum characteristic function

Fig. 3. Example of the modulus of a quantum characteristic function, taken from Example 3, with heavy tails in a single direction

The main drawback of the above state is that it does not have even first order moments. We can fix this by considering a slightly more sophisticated example. To proceed further, we first need to recall a well-known integral representation of fractional matrix powers.

Lemma 19 ([46, Proposition 5.16]). *For all $r \in (0, 1)$, all positive (possibly unbounded) operators A, and all $|\psi\rangle \in \mathrm{Dom}\left(A^{1/2}\right)$, we have that*

$$\left\| A^{r/2} |\psi\rangle \right\|^2 = \frac{\sin(\pi r)}{\pi} \int_0^\infty t^{r-1} \langle\psi| \frac{A}{tI + A} |\psi\rangle \, dt \,, \tag{87}$$

where all functions of A are defined by means of its spectral decomposition.

Proof of Proposition 11. The state is clearly centred, for instance because the wave function is symmetric under inversion $x \mapsto -x$. We proceed to prove claim (b). Note that, since $x^2 + p^2 = I + 2a^\dagger a \geq I$, $2aa^\dagger = x^2 + p^2 + I \leq 2(x^2 + p^2)$, where $p := -i\frac{d}{dx}$ is the momentum operator. We now apply the operator inequality $(A + B)^r \leq A^r + B^r$, which can be shown to hold for all $r \in [0, 1]$ and all positive (possibly unbounded) self-adjoint operators A, B. To prove this explicitly in the non-trivial case where $r \in (0, 1)$, we apply (87) to $A + B$. For a generic $|\psi\rangle \in \mathrm{Dom}\left(A^{r/2}\right) \cap \mathrm{Dom}\left(B^{r/2}\right)$, we obtain that

$$
\begin{aligned}
\left\| (A+B)^{r/2} |\psi\rangle \right\|^2 &= \frac{\sin(\pi r)}{\pi} \int_0^\infty t^{r-1} \langle\psi| \frac{A+B}{tI + A + B} |\psi\rangle \, dt \\
&= \frac{\sin(\pi r)}{\pi} \int_0^\infty t^{r-1} \langle\psi| \left(\frac{A}{tI + A + B} + \frac{B}{tI + A + B} \right) |\psi\rangle \, dt \\
&\leq \frac{\sin(\pi r)}{\pi} \int_0^\infty t^{r-1} \left(\langle\psi| \frac{A}{tI + A} |\psi\rangle + \langle\psi| \frac{B}{tI + B} |\psi\rangle \right) dt \\
&= \left\| A^{r/2} |\psi\rangle \right\|^2 + \left\| B^{r/2} |\psi\rangle \right\|^2,
\end{aligned}
$$

where the inequality in the above derivation follows e.g. from [46, Corollary 10.13]. Now, setting $A = x^2$, $B = p^2$ and $r = 1 - \delta$, we obtain that

$$(aa^\dagger)^{1-\delta} \leq |x|^{2(1-\delta)} + |p|^{2(1-\delta)} \leq |x|^{2(1-\delta)} + 1 + p^2 \,.$$

Computing the expectation value on $|\psi_f\rangle$ yields

$$\langle\psi_f|(aa^\dagger)^{1-\delta}|\psi_f\rangle \le \langle\psi_f|\left(|x|^{2(1-\delta)}+1+p^2\right)|\psi_f\rangle = \frac{1}{\sqrt{\pi}}\,\Gamma\left(\frac{3}{2}-\delta\right)\Gamma(\delta)+1+\frac{7}{10}\,,$$

where the last step is by explicit computation. This proves (b). We now move on to (c). For this we evaluate the characteristic function of the convolution $|\psi_f\rangle\langle\psi_f|^{\boxplus n}$ on the purely imaginary line. For $t \in \mathbb{R}$, using (19) we obtain that

$$\chi_{|\psi_f\rangle\langle\psi_f|}(it) = \int_{-\infty}^{+\infty} dx\,|f(x)|^2 e^{\sqrt{2}itx} = \sqrt{2}\,|t|\,K_1\left(\sqrt{2}\,|t|\right),$$

were K_1 is a modified Bessel function of the second kind, and the last equality follows from (54) and [75, Eq. (9.6.25)]. Therefore, for any fixed $t > 0$ it holds that

$$\lim_{n\to\infty}\chi_{|\psi_f\rangle\langle\psi_f|^{\boxplus n}}(it) = \lim_{n\to\infty}\chi_{|\psi_f\rangle\langle\psi_f|}\left(\frac{it}{\sqrt{n}}\right)^n$$

$$= \lim_{n\to\infty}\left(1+\left(c+\log t-\frac{1}{2}\log n\right)\frac{t^2}{n}+O\left(n^{-3/2}\right)\right)^n = 0\,,$$

where we have used the expansion in [75, Eq. (9.6.53)] (see also [75, Eq. (6.3.2) and (9.6.7)]). Since $\chi_{|\psi_f\rangle\langle\psi_f|^{\boxplus n}}(0) = 1$ for all n because $|\psi_f\rangle\langle\psi_f|^{\boxplus n}$ is a valid quantum state, the sequence of functions $\chi_{|\psi_f\rangle\langle\psi_f|^{\boxplus n}}$ does not possess a continuous limit. Hence, it cannot converge to the characteristic function of any quantum state. This proves (c).

□

7.2. Optimality of the convergence rates. The following two examples show that the bounds stated in Theorems 6 and 7 are indeed saturated. Both examples consist of states constructed using the Fock basis. The construction of examples saturating the bounds in Theorems 6 and 7 is motivated by the following Proposition.

Proposition 20. *Let ρ be a one-mode density operator satisfying the assumptions of Theorem 6 and also $\langle i|\rho|j\rangle = 0$ for $|i-j| \in \{1,3\}$. Then the state $\rho^{\boxplus n}$ converges at least with rate $\mathcal{O}\left(n^{-1}\right)$ to its Gaussification*

$$\left\|\rho^{\boxplus n}-\rho_{\mathrm{G}}\right\|_2 = \mathcal{O}\left(n^{-1}\right).$$

In particular, every density operator satisfying the assumptions of Theorem 6 that is diagonal in the Fock basis achieves a $\mathcal{O}(n^{-1})$ rate.

Proof of Proposition 20. By Theorem 6 it suffices to show that $D^3\chi_\rho(0) = 0$ under the assumptions of the Proposition. We start by recalling that any density operator ρ has an expansion into the Fock basis such that

$$\rho = \sum_{i,j=0}^{\infty} \langle i|\rho|j\rangle\,|i\rangle\langle j|\,. \tag{88}$$

Hence, we find for the characteristic function that

$$\chi_\rho(z) = \sum_{i,j=0}^{\infty} \langle i|\rho|j\rangle \, \chi_{|i\rangle\langle j|}(z). \tag{89}$$

Using a finite-rank approximation of the density operator ρ, it suffices then by Theorem 9 to analyse the component-wise derivatives in (89). The functions $\chi_{|i\rangle\langle j|}$ are explicitly given by [59, Eq. (4.4.46) and (4.4.47)]

$$\chi_{|i\rangle\langle j|}(z) = \begin{cases} \sqrt{\frac{i!}{j!}} \, (-z)^{j-i} e^{-\frac{|z|^2}{2}} L_j^{j-i}\left(|z|^2\right) & \text{if } i \leq j, \\ \sqrt{\frac{j!}{i!}} \, (z^*)^{i-j} e^{-\frac{|z|^2}{2}} L_i^{i-j}\left(|z|^2\right) & \text{if } i > j. \end{cases} \tag{90}$$

Here, $L_n^k(x) := \frac{d^k}{dx^k} L_n(x) = (-1)^k \sum_{\ell=0}^{n-k} \binom{n}{\ell+k} \frac{(-x)^\ell}{\ell!}$ are the associated Laguerre polynomials. By assumption, it suffices to consider the case where $|i-j|$ is even or $|i-j|$ is odd and at least 5. We find that by writing the characteristic function in the form $\chi_{|i\rangle\langle j|}(z) := e^{-\frac{|z|^2}{2}} H_{ji}(z)$ for some suitable function H_{ji}, as in (90), that for the different possible third derivatives, we have

$$\partial_z^3 \chi_{|i\rangle\langle j|}(0) = -3\partial_z H_{ji}(0) + \partial_z^3 H_{ji}(0),$$
$$\partial_{z^*}^3 \chi_{|i\rangle\langle j|}(0) = -3\partial_{z^*} H_{ji}(0) + \partial_{z^*}^3 H_{ji}(0),$$
$$\partial_z^2 \partial_{z^*} \chi_{|i\rangle\langle j|}(0) = -\partial_{z^*} H_{ji}(0) + \partial_z^2 \partial_{z^*} H_{ji}(0),$$
$$\partial_{z^*}^2 \partial_z \chi_{|i\rangle\langle j|}(0) = -\partial_z H_{ji}(0) + \partial_{z^*}^2 \partial_z H_{ji}(0).$$

Therefore, the only possible non-zero contribution to the third derivative of the quantum characteristic function χ_ρ at zero could be due to terms that contain either one or three derivatives of functions H_{ji} evaluated at zero.

If $|i-j| \geq 4$ then z and z^* appear in (90) with a joint power of at least 4; thus, this term's contribution necessarily has to vanish. It suffices therefore to consider the case where $|i-j| = 2$. If H_{ji} is only differentiated once, then it is clear that this derivative has to vanish at zero, since z, z^* appear with a joint power of at least two.

If H_{ji} is differentiated three times, then the term $|z|^2$ causes the derivative to vanish at zero unless this term is differentiated precisely two times. This, however, implies that the Laguerre polynomial is differentiated exactly once. However, by the chain rule any first order derivative of the term $L_j^{|j-i|}(|z|^2)$ vanishes at the origin. This concludes the proof. □

The following example shows that the $\mathcal{O}(n^{-1})$ convergence rate stated in Proposition 20, under the assumption that $D^3 \chi_\rho(0) = 0$, is in fact attained.

Example 4 ($\mathcal{O}(n^{-1})$-rate). By Proposition 20 we can take $\rho = |1\rangle\langle 1|$ to obtain a convergence rate of at least $\mathcal{O}(n^{-1})$ in the QCLT. That the $\mathcal{O}(n^{-1})$ rate is actually attained is illustrated in the right figure in Fig. 4. The $\mathcal{O}(n^{-1})$ rate is saturated both in Hilbert–Schmidt and trace norm.

The following example shows that the $\mathcal{O}(n^{-1/2})$ convergence rate of Theorem 7 is attained.
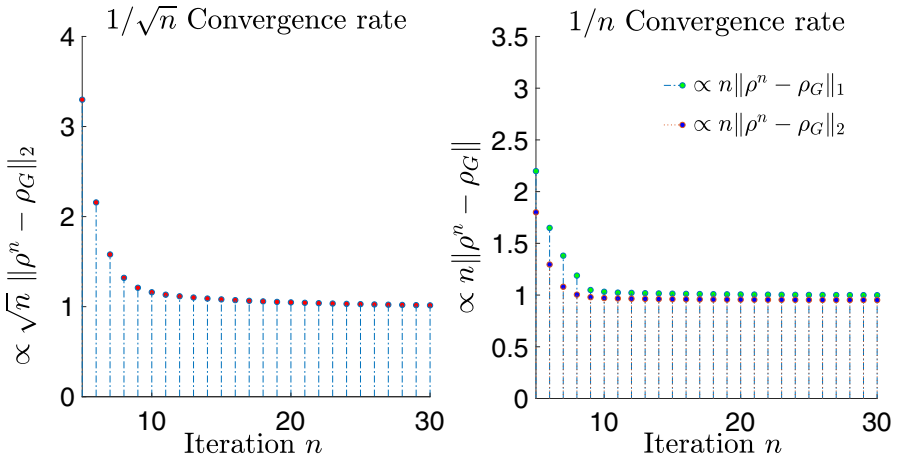
Fig. 4. This plot shows the expressions $cn^{\alpha}\|\rho^{\boxplus n} - \rho\|$ for a constant $c > 0$ such that $\lim_{n\to\infty} cn^{\alpha}\|\rho^{\boxplus n} - \rho\| = 1$. The left figure shows that the $\mathcal{O}(1/\sqrt{n})$ convergence rate is sharp (Theorem 7) by using the state from Example 5. The right figure shows that we can obtain a rate $\mathcal{O}(1/n)$ if $D^3\chi_\rho(0) = 0$ (Theorem 6) by using the state from Example 4. In both figures we write ρ^n for $\rho^{\boxplus n}$

Example 5 ($\mathcal{O}(n^{-1/2})$-rate). Consider the state[11]

$$\rho = \frac{|0\rangle + |3\rangle}{\sqrt{2}} \frac{\langle 0| + \langle 3|}{\sqrt{2}}.$$

Its characteristic function is explicitly given by (90)

$$\chi_\rho(z) = \frac{1}{12} e^{-\frac{|z|^2}{2}} \left(12 - 18|z|^2 + \sqrt{6}\left(z^3 - (z^*)^3\right) + 9|z|^4 - |z|^6\right)$$

Now, since $\langle 0|\rho|3\rangle \neq 0$ we see that the condition of Proposition 20 does not hold. One verifies directly that $\chi_\rho(z) = 1 - 2|z|^2 + o\left(|z|^2\right)$, so that ρ is already in Williamson form (cf. (64)). Letting $\Phi(z) = e^{-2|z|^2}$, we then find that $\|\chi_\rho - \Phi\|_{L^2(\mathbb{R}^2)}$ converges with rate $n^{-1/2}$, see Fig. 4.

The following example shows that the $\mathcal{O}(n^{-\alpha/2})$ convergence rate of Theorem 7 is attained at least for $\alpha = 1/2$.

Example 6. Consider the probability density function p on $\mathbb{R}$ given by

$$p(t) := \frac{\Gamma(7/4)}{\sqrt{\pi}\,\Gamma(5/4)} \frac{1}{(1+t^2)^{7/4}}.$$

Its Fourier transform reads

$$\widehat{p}(u) := \int_{-\infty}^{+\infty} dt\, e^{-itu} p(t) = \frac{|u|^{5/4}}{2^{1/4}\Gamma(5/4)} K_{5/4}(|u|), \tag{91}$$

---

[11] We use states $|0\rangle$ and $|3\rangle$ rather than $|0\rangle$ and $|1\rangle$ because the latter choice does not lead to a centred state.

where $K_\nu(z)$ is again the modified Bessel function of the second kind, and (91) follows from [75, Eq. (9.6.25)]. Define the single-mode quantum state

$$\rho = \int_{-\infty}^{+\infty} dt \, p(t) \, |t\rangle\langle t| \, ,$$

where

$$|t\rangle := e^{-t^2} \sum_{n=0}^{\infty} \frac{t^n}{\sqrt{n!}} \, |n\rangle$$

is a so-called coherent state [76–79]. The characteristic function of ρ can be easily computed as

$$\chi_\rho(z) = \int_{-\infty}^{+\infty} dt \, p(t) \, e^{-|z|^2/2 - 2itz_I} = \widehat{p}(2z_I) \, e^{-|z|^2/2} = \frac{2|z_I|^{5/4}}{\Gamma(5/4)} \, K_{5/4}(2|z_I|) \, e^{-|z|^2/2} \, ,$$

which leads us to

$$\chi_{\rho^{\boxplus n}}(z) = \chi_\rho \left(z/\sqrt{n} \right)^n = \frac{2^n |z_I|^{5n/4}}{\Gamma(5/4)^n n^{5n/8}} \, K_{5/4} \left(\frac{2|z_I|}{\sqrt{n}} \right)^n e^{-|z|^2/2} \, .$$

On the other hand, a little thought confirms that ρ has vanishing first moments and second moments given by $\mathrm{Tr}[\rho x^2] = 9/2$ and $\mathrm{Tr}[\rho p^2] = 1/2$. Its Gaussification then reads

$$\rho_{\mathrm{G}} = \int_{-\infty}^{+\infty} dt \, \frac{1}{2\sqrt{\pi}} \, e^{-t^2/4} \, |t\rangle\langle t| \, , \qquad \chi_{\rho_{\mathrm{G}}}(z) = e^{-4|z_I|^2 - |z|^2/2} \, .$$

We also observe that: (a) ρ has finite standard moments of order up to $5/2 - \delta$, for all $\delta > 0$; but (b) it has no well-defined phase space moments (nor standard moments) of order $5/2$.

To prove claim (a), start by setting $\beta := 5/4 - \delta/2$. Assuming that $\delta \leq 1/2$ so that $\beta \geq 1$, for all $t \in \mathbb{R}$ we have that

$$\langle t|(a^\dagger a)^\beta|t\rangle \overset{1}{=} e^{-t^2} \sum_{n=0}^{\infty} \frac{t^{2n}}{n!} \, n^\beta$$

$$= e^{-t^2} \sum_{n=1}^{\infty} \frac{t^{2n}}{(n-1)!} \, n^{\beta-1}$$

$$= t^2 e^{-t^2} \sum_{n=0}^{\infty} \frac{t^{2n}}{n!} \, (n+1)^{\beta-1}$$

$$\overset{2}{\leq} t^2 \left(\sum_{n=0}^{\infty} \frac{t^{2n} e^{-t^2}}{n!} (n+1) \right)^{\beta-1}$$

$$\overset{3}{=} t^2 \left(1 + t^2 \right)^{\beta-1}$$

$$\leq \left(1 + t^2 \right)^\beta \, ,$$

where 1 is just the definition of coherent state, 2 comes from the concavity of the function $x \mapsto x^{\beta-1}$ and from the fact that $q_n = \frac{t^{2n} e^{-t^2}}{n!}$ is a probability distribution over $\mathbb{N}$, and
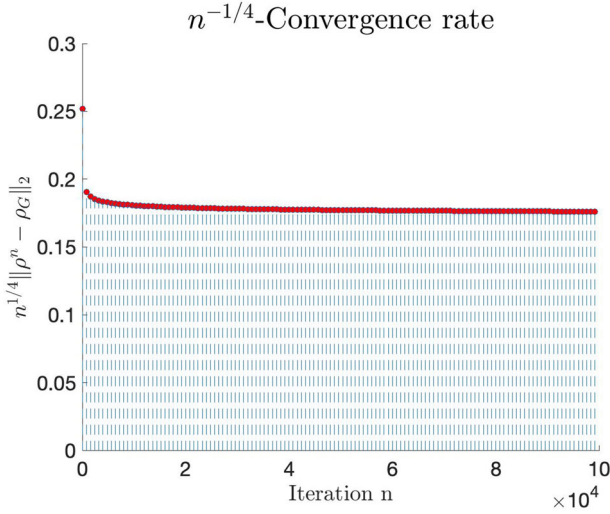
Fig. 5. This plot shows the expression $n^{-1/4}\|\rho^{\boxplus n} - \rho_G\|$, with $\rho^{\boxplus n}$ and ρ_G as in (92). In particular, this figure shows that the $\mathcal{O}(1/n^{1/4})$ convergence rate is sharp (Theorem 7) for $\alpha = 1/2$

finally in 3 we used the formula $\sum_{n=0}^{\infty} \frac{x^n}{n!}(n+1) = (1+x)e^x$. From the above calculation we now deduce that

$$\operatorname{Tr}\left[\rho\left(a^\dagger a\right)^\beta\right] \le \int_{-\infty}^{+\infty} dt \, p(t)\left(1+t^2\right)^\beta = \frac{\Gamma(7/4)}{\sqrt{\pi}\,\Gamma(5/4)}\int_{-\infty}^{+\infty}\frac{dt}{(1+t^2)^{(1+\delta)/2}} < \infty,$$

as claimed.

To prove claim (b), it suffices to use [75, Eq. (9.6.10) and (9.6.11)] in order to write $z^\nu K_\nu(z) = A(z) + z^{2\nu}\ln(z)\,B(z)$, with $\nu > 0$, A, B analytic functions, and $B(0) \ne 0$. Setting $\nu = 5/4$ shows that the phase space moment of ρ of order $5/2$, as constructed in Definition 2, is not well defined, formally $M'_{5/2}(\rho, \varepsilon) = +\infty$ for all $\varepsilon > 0$.

We now present numerical evidence hinting at the fact that $\left\|\rho^{\boxplus n} - \rho_G\right\|_2 = \mathcal{O}(n^{-1/4})$ for our choice of ρ. Note that

$$\left\|\rho^{\boxplus n} - \rho_G\right\|_2^2 = \int \frac{d^2z}{\pi}\left| e^{-4|z_I|^2 - |z|^2/2} - \frac{2^n|z_I|^{5n/4}}{\Gamma(5/4)^n n^{5n/8}}K_{5/4}\left(\frac{2|z_I|}{\sqrt{n}}\right)^n e^{-|z|^2/2}\right|^2$$

$$= \frac{2}{\sqrt{\pi}}\int_0^{+\infty} du \left| e^{-9u^2/2} - \frac{2^n u^{5n/4}}{\Gamma(5/4)^n n^{5n/8}}K_{5/4}\left(\frac{2u}{\sqrt{n}}\right)^n e^{-u^2/2}\right|^2.$$

$$(92)$$

The above integral can be evaluated numerically to a high degree of precision. Plotting the function $-\ln\left\|\rho^{\boxplus n} - \rho_G\right\|_2$ against $\ln n$ shows that $\left\|\rho^{\boxplus n} - \rho_G\right\|_2$ decays as $\mathcal{O}(n^{-1/4})$, cf. Figure 5. By what we have learnt above, Theorem 7 predicts a convergence at least as fast as $\mathcal{O}(n^{-1/4+\delta})$ for every fixed $\delta > 0$, and is therefore tight at least for $\alpha = 1/2$.

8. Cascade of Beam Splitters: Proofs

In this section, we prove the results claimed in Sect. 4.3, namely convergence rates for cascades of beam splitters converging to thermal attenuator channels.

8.1. Generalities of the cascade channels. In order to study the convergence of the cascade channel, we start by proving the following elementary equivalence.

Lemma 21. *For an m-mode quantum state ρ, some $\lambda \in [0, 1]$, and a positive integer n, consider the cascade channel $\mathcal{N}^n_{\rho, \lambda^{1/n}}$ (cf. (41)). One has that*

$$\mathcal{N}^n_{\rho, \lambda^{1/n}} = \mathcal{N}_{\rho(\lambda, n), \lambda} : \chi_\omega(z) \longmapsto \chi_{\mathcal{N}^n_{\rho, \lambda^{1/n}}(\omega)}(z) = \chi_\omega\left(\sqrt{\lambda}\, z\right) \chi_{\rho(\lambda, n)}\left(\sqrt{1 - \lambda}\, z\right),$$
(93)

where the effective environment state $\rho(\lambda, n)$ *is defined via its characteristic function*

$$\chi_{\rho(\lambda, n)}(z) := \prod_{\ell=1}^n \chi_\rho\left(\sqrt{\frac{1 - \lambda^{1/n}}{1 - \lambda}}\, \lambda^{\frac{\ell-1}{2n}} z\right).$$
(94)

Proof. We proceed by induction. The case $n = 1$ follows from (38). Let us assume that the claim holds for $n - 1$, so that

$$\mathcal{N}^{n-1}_{\rho, \mu^{1/(n-1)}} = \mathcal{N}_{\rho(\mu, n-1), \mu}, \qquad \chi_{\mathcal{N}^{n-1}_{\rho, \mu^{1/(n-1)}}(\omega)}(z) = \chi_\omega\left(\sqrt{\mu}\, z\right) \prod_{\ell=1}^{n-1} \chi_\rho\left(\sqrt{1 - \mu^{1/(n-1)}}\, \mu^{\frac{\ell-1}{2(n-1)}} z\right).$$

By setting $\mu = \lambda^{(n-1)/n}$ we see that

$$\mathcal{N}^{n-1}_{\rho, \lambda^{1/n}} = \mathcal{N}_{\rho(\lambda^{(n-1)/n}, n-1), \lambda^{(n-1)/n}}, \qquad \chi_{\mathcal{N}^{n-1}_{\rho, \lambda^{1/n}}(\omega)}(z) = \chi_\omega\left(\lambda^{\frac{n-1}{2n}} z\right) \prod_{\ell=1}^{n-1} \chi_\rho\left(\sqrt{1 - \lambda^{1/n}}\, \lambda^{\frac{\ell-1}{2n}} z\right).$$

Since $\mathcal{N}^n_{\rho, \lambda^{1/n}}(\omega) = \mathcal{N}_{\rho, \lambda^{1/n}}\left(\mathcal{N}^{n-1}_{\rho, \lambda^{1/n}}(\omega)\right) = \mathcal{N}^{n-1}_{\rho, \lambda^{1/n}}(\omega) \boxplus_{\lambda^{1/n}} \rho$, composition with the nth copy of the channel $\mathcal{N}_{\rho, \lambda^{1/n}}$ yields

$$\chi_{\mathcal{N}^n_{\rho, \lambda^{1/n}}(\omega)}(z) = \chi_{\mathcal{N}^{n-1}_{\rho, \lambda^{1/n}}(\omega)}\left(\lambda^{1/2n} z\right) \chi_\rho\left(\sqrt{1 - \lambda^{1/n}}\, z\right)$$

$$= \chi_\omega\left(\sqrt{\lambda}\, z\right) \left(\prod_{\ell=1}^{n-1} \chi_\rho\left(\sqrt{1 - \lambda^{1/n}}\, \lambda^{\frac{\ell}{2n}} z\right)\right) \chi_\rho\left(\sqrt{1 - \lambda^{1/n}}\, z\right)$$

$$= \chi_\omega\left(\sqrt{\lambda}\, z\right) \prod_{\ell=0}^{n-1} \chi_\rho\left(\sqrt{1 - \lambda^{1/n}}\, \lambda^{\frac{\ell}{2n}} z\right)$$

$$= \chi_\omega\left(\sqrt{\lambda}\, z\right) \prod_{\ell=1}^n \chi_\rho\left(\sqrt{1 - \lambda^{1/n}}\, \lambda^{\frac{\ell-1}{2n}} z\right),$$

which proves (93) and (94). Finally, one can also verify by induction that

$$\rho(\lambda, n) = \rho\left(\lambda^{\frac{n-1}{n}}, n-1\right) \boxplus_{\eta(\lambda, n)} \rho,$$
(95)

where $\eta(\lambda, n) := \frac{\lambda^{1/n}\left(1 - \lambda^{(n-1)/n}\right)}{1 - \lambda} \in [0, 1]$, so that $\rho(\lambda, n)$ is a legitimate quantum state for all $\lambda \in [0, 1]$ and all n. $\square$

8.2. On the effective environment state. Thanks to Lemma 21, the study of the cascade channel $\mathcal{N}^n_{\rho,\,\lambda^{1/n}}$ boils down to that of the iteratively convolved state $\rho(\lambda, n)$ of (94). Since such a convolution is *not* symmetric (cf. (49)), to proceed further we need to extend our quantum Berry–Esseen results to a non-i.i.d. scenario. Note that the classical central limit theorem has indeed been extended to sequences of independent, non-identically distributed random variables [1,80], and even to sequences of correlated random variables [81]. Rates of convergence for the former case can be found for instance in [82] (see e.g. Theorem 13.3 of [82]).

Proposition 22. *Let ρ be a centred m-mode quantum state with finite second-order phase space moments. Then the sequence of quantum states $\rho(\lambda, n)$ defined via (94) converges to the Gaussification ρ_G of ρ in trace norm. Moreover, if ρ has finite third-order phase space moments then*

$$\|\rho(\lambda, n) - \rho_G\|_2 = \mathcal{O}_{\lambda, M'_3}\left(n^{-1/2}\right), \tag{96}$$

$$\|\rho(\lambda, n) - \rho_G\|_1 = \mathcal{O}_{\lambda, M'_3}\left(n^{-1/(2m+2)}\right). \tag{97}$$

Here, $M'_3 = M'_3(\rho, \varepsilon)$ is defined by (33), and $\varepsilon > 0$ is sufficiently small.

Proof. The argument is a variation of that used to prove Theorem 7 in Sect. 6.2. First of all, reasoning as in Sect. 6.1.1, we can assume without loss of generality that ρ is in its Williamson form. To simplify the notation, we introduce the re-scaled vectors $w_\ell := \sqrt{\frac{1-\lambda^{1/n}}{1-\lambda}}\,\lambda^{\frac{\ell-1}{2n}}\, z \in \mathbb{C}^m$, where $\ell \in \{1, \ldots, n\}$. Then clearly $\chi_{\rho(\lambda,n)}(z) = \prod_{\ell=1}^n \chi_\rho(w_\ell)$. Note that $|w_\ell| \leq \sqrt{\frac{\log(1/\lambda)}{n(1-\lambda)}}\,|z|$; substituting $z \mapsto w_\ell$ into (72) and (74), we see that whenever $|z| \leq \sqrt{\frac{n(1-\lambda)}{\log(1/\lambda)}}\,\varepsilon$ it holds that

$$\left|\chi_\rho(w_\ell) - 1\right| \leq \frac{m(2m+1)}{2}\,M'_2(\rho, \varepsilon)\,\frac{\log\frac{1}{\lambda}}{1-\lambda}\,\frac{|z|^2}{n}, \tag{98}$$

$$\left|\chi_\rho(w_\ell) - 1 + \frac{1}{2}\frac{1-\lambda^{1/n}}{1-\lambda}\,\lambda^{\frac{\ell-1}{n}}\sum_j \nu_j|z_j|^2\right| \leq \frac{m(m+1)(2m+1)}{9}\,M'_3(\rho, \varepsilon)\left(\frac{\log\frac{1}{\lambda}}{1-\lambda}\right)^{3/2}\frac{|z|^3}{n^{3/2}}. \tag{99}$$

We start by choosing $\varepsilon > 0$ small enough so that (76) holds for some $\mu \in (0, 2)$. We can now mimic the calculations in (79), obtaining

$$\left|\log \chi_{\rho(\lambda,n)}(z) + \frac{1}{2}\sum_j \nu_j|z_j|^2\right|$$

$$= \left|\log \prod_{\ell=1}^n \chi_\rho(w_\ell) + \frac{1}{2}\sum_j \nu_j|z_j|^2\right|$$

$$\leq \sum_{\ell=1}^n \left|\log \chi_\rho(w_\ell) + \frac{1}{2}\frac{1-\lambda^{1/n}}{1-\lambda}\,\lambda^{\frac{\ell-1}{n}}\sum_j \nu_j|z_j|^2\right|$$

$$\leq \sum_{\ell=1}^n \left\{\left|\log \chi_\rho(w_\ell) + 1 - \chi_\rho(w_\ell)\right| + \left|\chi_\rho(w_\ell) - 1 + \frac{1}{2}\frac{1-\lambda^{1/n}}{1-\lambda}\,\lambda^{\frac{\ell-1}{n}}\sum_j \nu_j|z_j|^2\right|\right\}$$

$$\overset{2}{\leq} \sum_{\ell=1}^{n} \left\{ \frac{m^2(2m+1)^2}{4} \frac{\log \frac{1}{\lambda}}{1-\lambda} M_2'(\rho, \varepsilon)^2 \, a(\mu) \, \frac{|z|^4}{n^2} + \frac{m(m+1)(2m+1)}{9} M_3'(\rho, \varepsilon) \left(\frac{\log \frac{1}{\lambda}}{1-\lambda} \right)^{3/2} \frac{|z|^3}{n^{3/2}} \right\}$$

$$\overset{3}{\leq} \frac{C_6 |z|^3}{\sqrt{n}}, \tag{100}$$

Here, in 1 we observed that $\sum_{\ell=1}^{n} \frac{1-\lambda^{1/n}}{1-\lambda} \lambda^{\frac{\ell-1}{n}} = 1$ and applied the triangle inequality. To deduce 2, instead, we proceeded as for (78). Namely, on the first addend we used the identity $\log\left(1 - \frac{x}{2}\right) + \frac{x}{2} = -\frac{x^2}{4} a(x)$ satisfied by the function $a(x)$ defined by (77), we set $x = 2\left(1 - \chi_\rho(w_\ell)\right)$, we noted that $|x| \leq \mu$ implies that $|a(x)| \leq a(\mu)$, and lastly we employed (98). The second addend, instead, has been estimated thanks to (99). Finally, for fixed m the constant introduced in 3 depends only on M_3' and λ (again, $M_2' \leq M_3'$ by construction).

Proceeding as usual, we continue to estimate

$$e^{\frac{1}{2} \sum_j \nu_j |z_j|^2} \left| \chi_{\rho(\lambda,n)}(z) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right| = e^{\frac{1}{2} \sum_j \nu_j |z_j|^2} \left| \prod_{\ell=1}^{n} \chi_\rho(w_\ell) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right|$$

$$= \left| \exp\left(\log\left(\prod_{\ell=1}^{n} \chi_\rho(w_\ell) + \frac{1}{2} \sum_j \nu_j |z_j|^2 \right) \right) - 1 \right|$$

$$\overset{4}{\leq} \frac{C_6 |z|^3}{\sqrt{n}} e^{C_6 |z|^3/\sqrt{n}}$$

$$\overset{5}{\leq} \frac{C_6 |z|^3}{\sqrt{n}} e^{\frac{1}{4}|z|^2}.$$

Note that in 4 we applied the elementary inequality $|e^u - 1| \leq |u|e^{|u|}$, observed that $\mathbb{R} \ni x \mapsto xe^x$ is a monotonically increasing function, and leveraged the bound in (100). In 5, instead, we wrote $\frac{C_6|z|^3}{\sqrt{n}} \leq C_6 \sqrt{\frac{1-\lambda}{\log \frac{1}{\lambda}}} \, \varepsilon \, |z|^2 \leq \frac{1}{4}|z|^2$, where the last estimate holds provided that $\varepsilon > 0$ is small enough.

Remembering that $\nu_1, \dots, \nu_m \geq 1$, we can massage the above relation so as to get

$$\left| \chi_{\rho(\lambda,n)}(z) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right| \leq \frac{C_6 |z|^3}{\sqrt{n}} e^{-\frac{1}{4}|z|^2}. \tag{101}$$

Now, we can repeat the steps that led to (68). We obtain that

$$\pi^m \|\rho(\lambda, n) - \rho_G\|_2^2$$

$$= \int d^{2m}z \left| \chi_{\rho(\lambda,n)}(z) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right|^2$$

$$\overset{6}{\leq} \int_{|z| \leq \sqrt{n}\varepsilon} d^{2m}z \left| \chi_{\rho(\lambda,n)}(z) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right|^2 + \int_{|z| > \sqrt{n}\varepsilon} d^{2m}z \left| \chi_{\rho(\lambda,n)}(z) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right|^2$$

$$\overset{7}{\leq} \int_{|z| \leq \sqrt{n}\varepsilon} d^{2m}z \, \frac{C_6 |z|^6}{n} e^{-\frac{1}{2}|z|^2} + 2 \int_{|z| > \sqrt{n}\varepsilon} d^{2m}z \, |\chi_{\rho(\lambda,n)}(z)|^2 + 2 \int_{|z| > \sqrt{n}\varepsilon} d^{2m}z \, e^{-\sum_j \nu_j |z_j|^2}$$

$$\overset{8}{\leq} \frac{C_7}{n} + 2 \left(\sup_{|z| > \sqrt{n}\varepsilon} \prod_{\ell=2}^{n} |\chi_\rho(w_\ell)|^2 \right) \int_{|z| > \sqrt{n}\varepsilon} d^{2m}z \, |\chi_\rho(w_1)|^2 + 2e^{-\frac{\varepsilon^2}{2}n} \int_{|z| \geq \sqrt{n}\varepsilon} e^{-\frac{1}{2}|z|^2}$$

$$\leq \frac{9}{n}\frac{C_7}{} + 2 \left(\sup_{|v| > \frac{1}{\sqrt{2}} \sqrt{\lambda \log(1/\lambda)}\,\varepsilon} \left|\chi_\rho(v)\right| \right)^{2(n-1)} \left(\frac{2n(1-\lambda)}{\log \frac{1}{\lambda}} \right)^m + 2(2\pi)^m e^{-\frac{\varepsilon^2}{2}n} \tag{102}$$

The justification of the above steps is as follows. The estimate in 6 is just an application of the triangle inequality. In 7 we used (101) and the elementary fact that $|u + v|^2 \leq 2|u|^2 + 2|v|^2$ on the second addend. As for 8, we: (i) performed the integral and introduced a constant C_7 that depends on m only on the first addend; (ii) decomposed $\chi_{\rho(\lambda,n)}(z) = \chi_\rho(w_1) \cdot \prod_{\ell=2}^{n} \chi_\rho(w_\ell)$ on the second; and (iii) used the fact that $e^{-\sum_j v_j |z_j|^2} \leq e^{-|z|^2} < e^{-\frac{\varepsilon^2}{2}n} e^{-\frac{1}{2}|z|^2}$ in the prescribed range on the third. Finally, in 9 we noted that if $|z| > \sqrt{n}\,\varepsilon$ then eventually in n

$$|w_\ell| = \sqrt{\frac{1 - \lambda^{1/n}}{1 - \lambda}} \lambda^{\frac{\ell-1}{2n}} |z| > \frac{1}{\sqrt{2}} \sqrt{\lambda \log \frac{1}{\lambda}}\,\varepsilon$$

for all $\ell \in \{1, \ldots, n\}$; moreover, we used the fact that $\int d^{2m}u\, |\chi_\rho(u)|^2 \leq 1$ to evaluate the integral in the second addend.

Since the second term in the rightmost side of (102) decays faster than any inverse power of n as $n \to \infty$ thanks to Proposition 14, the proof of (96) is complete. Lastly, (97) follows similarly to Corollary 8. □

8.3. Approximating cascade channels. With this convergence at hand, we provide a quantitative bound on the approximation of thermal attenuator channels by cascades of beam splitters (with possibly non-Gaussian environment states). Recall that, to an environment state ρ one can associate an attenuator channel $\mathcal{N}_{\rho,\lambda}(\omega) := \omega \boxplus_\lambda \rho$ of transmissivity $0 \leq \lambda \leq 1$. The following simple lemma is crucial to convert the above state approximation result (Proposition 22) into a statement about approximations of attenuator channels.

Lemma 23. *Given any two m-mode quantum states ρ_1 and ρ_2, and some $\lambda \in [0, 1]$, the corresponding channels defined as in (41) satisfy*

$$\left\| \mathcal{N}_{\rho_1,\lambda} - \mathcal{N}_{\rho_2,\lambda} \right\|_\diamond \leq \left\| \rho_1 - \rho_2 \right\|_1. \tag{103}$$

Proof. Let R be any reference system, and let $\omega \in \mathcal{D}(\mathcal{H}_{AR})$ be a state on the bipartite system AR. Then

$$\begin{aligned}
\left\| (\mathcal{N}_{\rho_1,\lambda} - \mathcal{N}_{\rho_2,\lambda}) \otimes \mathrm{id}_R(\omega) \right\|_1 &= \left\| \mathrm{Tr}_R[U_\lambda(\omega \otimes \rho_1)U_\lambda^\dagger] - \mathrm{Tr}_R[U_\lambda(\omega \otimes \rho_2)U_\lambda^\dagger] \right\|_1 \\
&\leq \left\| U_\lambda(\omega \otimes \rho_1 - \omega \otimes \rho_2)U_\lambda^\dagger \right\|_1 \\
&= \left\| \rho_1 - \rho_2 \right\|_1,
\end{aligned}$$

where the inequality stems from the monotonicity of trace distance under quantum channels. □

With this lemma at hand, we are ready to prove Theorem 12.

Proof of Theorem 12. Recall from Lemma 21 that $\mathcal{N}_{\rho,\,\lambda^{1/n}}^n = \mathcal{N}_{\rho(\lambda,n),\,\lambda}$, where $\rho(\lambda, n)$ is the state with characteristic function given by (94). Applying Lemma 23 and Proposition 22, we have that

$$\left\| \mathcal{N}_{\rho,\,\lambda^{1/n}}^n - \mathcal{N}_{\rho_G,\lambda} \right\|_\diamond \overset{(93)}{=} \left\| \mathcal{N}_{\rho(\lambda,n),\,\lambda} - \mathcal{N}_{\rho_G,\lambda} \right\|_\diamond \overset{(103)}{\leq} \| \rho(\lambda, n) - \rho_G \|_1 \overset{(97)}{=} \mathcal{O}_{M_3'}\left(n^{-1/(2m+2)} \right),$$

concluding the proof. □

Proof of Corollary 13. We now move on to Corollary 13. Let us start by proving the statement on quantum capacities, namely (56) and (57). Our aim is to apply [34, Theorem 9] to the two channels $\mathcal{N}_{\rho,\,\lambda^{1/n}}^n$ and $\mathcal{N}_{\rho_G,\lambda} = \mathcal{E}_{N,\lambda}$, for the special case $m = 1$ (cf. (42)). We set

$$\varepsilon_n := \frac{1}{2} \left\| \mathcal{N}_{\rho,\,\lambda^{1/n}}^n - \mathcal{E}_{N,\lambda} \right\|_\diamond \geq \frac{1}{2} \left\| \mathcal{N}_{\rho,\,\lambda^{1/n}}^n - \mathcal{E}_{N,\lambda} \right\|_{\diamond E},$$

where the energy-constrained diamond norm is defined with respect to the canonical Hamiltonian, namely the number operator $H = a^\dagger a$ (see [36, Eq. (2)] and [34, Eq. (2)]). Note that $\varepsilon_n = \mathcal{O}_{M_3'}\left(n^{-1/4} \right)$ by Theorem 12.

The input–output energy relations can be easily determined for both channels thanks to (95) and (40), which together show that $\text{Tr}\left[\rho(\lambda, n)\, a^\dagger a \right] = \text{Tr}\left[\rho\, a^\dagger a \right] = N = \text{Tr}\left[\tau_N a^\dagger a \right]$. One obtains that

$$\begin{aligned} \text{Tr}\left[\mathcal{N}_{\rho,\,\lambda^{1/n}}^n(\omega)\, a^\dagger a \right] &= \text{Tr}\left[\mathcal{N}_{\rho(\lambda,n),\,\lambda}(\omega)\, a^\dagger a \right] = \lambda\, \text{Tr}\left[\omega\, a^\dagger a \right] + (1 - \lambda)N, \\ \text{Tr}\left[\mathcal{E}_{N,\lambda}(\omega)\, a^\dagger a \right] &= \lambda\, \text{Tr}\left[\omega\, a^\dagger a \right] + (1 - \lambda)N. \end{aligned} \tag{104}$$

This means that we can set $\alpha = \lambda$ and $E_0 = (1 - \lambda)N$, and hence $\widetilde{E} = \lambda E + (1 - \lambda)N$, in [34, Theorem 9]. We obtain that

$$\begin{aligned} \left| Q\left(\mathcal{N}_{\rho,\,\lambda^{1/n}}^n, E \right) - Q\left(\mathcal{E}_{N,\lambda}, E \right) \right| &\overset{1}{\leq} 56\sqrt{\varepsilon_n}\, g\left(\frac{4(\lambda E + (1 - \lambda)N)}{\sqrt{\varepsilon_n}} \right) + 6\, g\left(4\sqrt{\varepsilon_n} \right) \\ &\overset{2}{\leq} c\, \sqrt{\varepsilon_n} \log \frac{1}{\varepsilon_n} \\ &\overset{3}{\leq} C(M_3') \frac{\log n}{n^{1/8}}. \end{aligned}$$

Here, in step 1 we applied [34, Theorem 9] together with the formula $S(\tau_N) = g(N)$ (see (28) and (29)); the inequality in 2 holds eventually in n for some universal constant $c \leq 57 + 24 \log e$, as can be seen by combining the bounds $g(x) \leq \log(x + 1) + \log e$ (tight for large x) and $g(x) \leq -2x \log x$ (valid for sufficiently small x); finally, in 3 we used the fact that $\varepsilon_n \leq C(M_3')\, n^{-1/4}$ eventually in n by the already proven Theorem 12, together with the observation that $x \mapsto -x \log x$ is an increasing function for sufficiently small $x > 0$.

To complete the first part of the proof we need to estimate the classical capacity of $\mathcal{N}_{\rho,\,\lambda^{1/n}}^n$ in terms of that of the thermal attenuator $\mathcal{E}_{N,\lambda}$ of (42), in turn given by (43). Although we could use the estimates in [34], we prefer to resort to the tighter ones provided in [36]. We obtain that

$$\left| \mathscr{C}\left(\mathscr{N}_{\rho,\,\lambda^{1/n}}^{n}, E \right) - \mathscr{C}\left(\mathscr{E}_{N,\lambda}, E \right) \right| \overset{4}{\leq} 7\varepsilon_n \left(\log\left(\lambda E + (1-\lambda)N + 1\right) - \log\frac{\varepsilon_n}{2} + 1 \right) + 2g\left(\frac{5\varepsilon_n}{2}\right) + 4h_2\left(\frac{\varepsilon_n}{2}\right)$$

$$\overset{5}{\leq} c'\varepsilon_n \log\frac{1}{\varepsilon_n} \overset{6}{\leq} C'(M_3') \frac{\log n}{n^{1/4}}.$$

The inequality in 4 is an application of [36, Proposition 6]. To see why, let us re-write the result of Shirokov [36, Proposition 6] for one-mode channels and with respect to the canonical Hamiltonian as

$$\left| \mathscr{C}\left(\mathscr{N}_1, E \right) - \mathscr{C}\left(\mathscr{N}_2, E \right) \right| \leq 2\epsilon \left(2t + r_\epsilon(t)\right)\left(\log(E'+1) + 1 - \log(\epsilon t)\right) + 2g\left(\epsilon r_\epsilon(t)\right) + 4h_2(\epsilon t).$$

Here, $\mathscr{N}_i$ ($i = 1, 2$) are two quantum channels with $\frac{1}{2}\|\mathscr{N}_1 - \mathscr{N}_2\|_{\diamond E} \leq \epsilon$, we picked E' such that $\sup_{\rho:\, \mathrm{Tr}[\rho\, a^\dagger a] \leq E} \mathrm{Tr}\left[\mathscr{N}_i(\rho)\, a^\dagger a\right] \leq E'$, the function r_ϵ is defined by $r_\epsilon(t) := \frac{1+t/2}{1-\epsilon t}$, and $h_2(x) := -x\log x - (1-x)\log(1-x)$ is the binary entropy. Setting $\mathscr{N}_1 = \mathscr{N}_{\rho,\,\lambda^{1/n}}^{n}$, $\mathscr{N}_2 = \mathscr{E}_{N,\lambda}$, we see that $E' = \lambda E + (1-\lambda)N$ (cf. (104) and [36, Eq. (21)]); choosing $t = 1/2$ and hence $r_\varepsilon(t) \leq r_1(t) = r_1(1/2) = 5/2$ yields the above relation 4, as claimed. The inequality in 5 holds for all sufficiently large n and for some absolute constant $c' \leq 15$. Finally, 6 is analogous to 3 above. $\quad\square$

Remark. Let us stress that the threshold in n above which the inequalities in the above proof hold true depends on both $\lambda E + (1-\lambda)N$ and M_3' (which dictates the rate of convergence of $\varepsilon_n \to 0$). Although this is a minor point from the point of view of the mathematical derivation, it may be important for applications.

Remark. An analytical formula for the quantum capacity of the thermal attenuator that appears in Corollary 13 is currently not known. The best lower bound to date reads [45, Eq. (9)]

$$Q\left(\mathscr{E}_{N,\lambda}, E\right) \geq \max_{0 \leq x \leq 1} x \left\{ g\left(\lambda\frac{E}{x} + (1-\lambda)N\right) - g\left(\frac{1}{2}\left(D_{\lambda,N}(x) + (1-\lambda)\left(\frac{E}{x} - N\right) - 1\right)\right) \right.$$
$$\left. - g\left(\frac{1}{2}\left(D_{\lambda,N}(x) - (1-\lambda)\left(\frac{E}{x} - N\right) - 1\right)\right) \right\}, \tag{105}$$

where

$$D_{\lambda,N}(x) := \sqrt{\left((1+\lambda)\frac{E}{x} + (1-\lambda)N + 1\right)^2 - 4\lambda\frac{E}{x}\left(\frac{E}{x} + 1\right)}. \tag{106}$$

The best upper bound to date, instead, can be obtained by combining the results of [40, Eq. (23)–(25)] (see also [41, Section 8]) with those of [44, Theorem 9] and [43, Theorem 46], in turn derived by refining a technique introduced in [42]. We look at the case where $\lambda \geq \frac{N+1/2}{N+1}$, because below that value of λ the channel $\mathscr{E}_{N,\lambda}$ becomes 2-extendable [83] (that is, anti-degradable [84–86]) and therefore $Q\left(\mathscr{E}_{N,\lambda}, E\right) = 0$.

$$Q\left(\mathscr{E}_{N,\lambda}, E\right) \leq \max\left\{F_1(N, \lambda; E), F_2(N, \lambda; E)\right\}, \tag{107}$$

$$F_1(N, \lambda, E) := g\left(\lambda E + (1-\lambda)N\right) - g\left(\frac{(1-\lambda)(N+1)}{\lambda - (1-\lambda)N}\left(\lambda E + (1-\lambda)N\right)\right), \tag{108}$$

$$F_2(N, \lambda; E) := -\log\left((1-\lambda)\lambda^N\right) - g(N). \tag{109}$$

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

Appendix A: Standard Moments Versus Phase Space Moments: The Integer Case

In this appendix we prove that a state with finite standard moments of order up to k also has finite phase space moments of order up to k, i.e. Theorem 10. More precisely, we show that its characteristic function is differentiable k times, and that there are constants $c_{k,m}(\varepsilon) < \infty$ such that the standard moments and phase space moments, defined by (31) and (33), respectively, satisfy $M'_k(\rho, \varepsilon) \leq c_{k,m}(\varepsilon) M_k(\rho)$ for all m-mode quantum states ρ. We start with the following lemma.

Lemma 24. *For all positive integers m and real numbers $k \in [0, \infty)$, there is a universal constant $d_{k,m} > 0$ such that*

$$\left(\sum_j a_j a_j^\dagger \right)^{k/2} \geq d_{k,m} \sum_j \left(|x_j|^k + |p_j|^k \right), \tag{A1}$$

where $x_j := (a_j + a_j^\dagger)/\sqrt{2}$ and $p_j := -i(a_j - a_j^\dagger)/\sqrt{2}$ are the position and momentum quadratures of the jth mode.

Proof. First of all, it suffices to consider the one-mode case. Indeed, assume that $(aa^\dagger)^{k/2} \geq d_k(|x|^k + |p|^k)$ for some $d_k > 0$. Then, leveraging the fact that the operators $a_j a_j^\dagger$ commute with each other, and employing standard inequalities between p-norms, we deduce that

$$\left(\sum_j a_j a_j^\dagger \right)^{k/2} \geq m^{\min\{k/2-1, 0\}} \sum_j (a_j a_j^\dagger)^{k/2} \geq m^{\min\{k/2-1, 0\}} d_k \sum_j \left(|x_j|^k + |p_j|^k \right).$$

Therefore, from now on we look at the one-mode case only. The vector space $V_m := \mathrm{span}\{|n\rangle : n \in \mathbb{N}_0\}$ of states with a finite expansion in the Fock basis is a core for both $\left(aa^\dagger\right)^{k/2}$, as well as $|x|^k$ and $|p|^k$. Thus, it suffices for us to prove the inequality (A1) on states in V_m.

It is enough to show that $(aa^\dagger)^{k/2} \geq d_k |x|^k$ for some constants $d_k > 0$, as the other inequality $(aa^\dagger)^{k/2} \geq d_k |p|^k$ is obtained by performing a phase space rotation of an angle $\pi/2$, i.e. by conjugating both sides by the unitary operator $e^{i \frac{\pi}{2} aa^\dagger}$.

We now prove that the inequality $(aa^\dagger)^{k/2} \geq d_k|x|^k$ holds for some $d_k > 0$ on all vectors in V_m. Write $k = 2rh$, where $r \in (0, 1]$ and $h := \lceil k/2 \rceil \in \mathbb{N}_0$. Since the function $A \mapsto A^r$ is well known to be operator monotone [46, Proposition 10.14], it suffices to show that $(aa^\dagger)^h \geq d_h x^{2h}$ for all non-negative integers $h \in \mathbb{N}_0$. To this end, let us take advantage of our restriction to states with a finite expansion in the Fock basis. Defining Π_N as the projector onto the span of the first $N + 1$ Fock states (from 0 to N), we have to show that

$$A_N := \sum_{n=0}^{N}(1+n)^h \, |n\rangle\langle n| - d_h \Pi_N x^{2h} \Pi_N^\dagger = \Pi_N \left((1 + a^\dagger a)^h - d_h x^{2h}\right) \Pi_N^\dagger \geq 0 \,,$$

where the inequality now involves only matrices. Thanks to Gershgorin's circle theorem [87,88], in order to show that A_N is positive semi-definite, it suffices to prove that A_N is diagonally dominant, i.e. that for all N and $0 \leq n \leq N$ the inequality

$$\langle n|A_N|n\rangle - \sum_{\substack{n'=0,\ldots,N, \\ n' \neq n}} |\langle n|A_N|n'\rangle| \geq 0 \tag{A2}$$

holds true. Writing down the left-hand side yields

$$\langle n|A_N|n\rangle - \sum_{\substack{n'=0,\ldots,N, \\ n' \neq n}} |\langle n|A_N|n'\rangle|$$

$$\geq (1+n)^h - d_h \sum_{n'=0}^{N} \left|\langle n|x^{2h}|n'\rangle\right|$$

$$\overset{1}{\geq} (1+n)^h - d_h \sum_{n':|n-n'|\leq 2h} \left|\langle n|x^{2h}|n'\rangle\right|$$

$$= (1+n)^h - d_h \sum_{n':n\leq n'\leq n+2h} \left|\langle n|x^{2h}|n'\rangle\right| - d_h \sum_{n':\max\{n-2h,0\}\leq n' < n} \left|\langle n|x^{2h}|n'\rangle\right|$$

$$\overset{2}{\geq} (1+n)^h - d_h \sum_{n':n\leq n'\leq n+2h} \sum_{\ell=0}^{h-(n'-n)/2} |f_{h,\ell,n'-n}| \left|\langle n|(a^\dagger a)^\ell a^{n'-n}|n'\rangle\right|$$

$$- d_h \sum_{n':\max\{n-2h,0\}\leq n' < n} \sum_{\ell=0}^{h-(n-n')/2} |f'_{h,\ell,n-n'}| \left|\langle n|(a^\dagger a)^\ell (a^\dagger)^{n-n'}|n'\rangle\right|$$

$$= (1+n)^h - d_h \sum_{n':n\leq n'\leq n+2h} \sum_{\ell=0}^{h-(n'-n)/2} |f_{h,\ell,n'-n}| n^\ell \sqrt{\frac{n'!}{n!}}$$

$$- d_h \sum_{n':\max\{n-2h,0\}\leq n' < n} \sum_{\ell=0}^{h-(n-n')/2} |f'_{h,\ell,n-n'}| n^\ell \sqrt{\frac{n!}{n'!}}$$

$$\overset{3}{\geq} (1+n)^h - d_h F_h(n + 2h)^h$$

$$\overset{4}{\geq} 0 \,.$$

Here, in 1 we extended the sum over n' to all values that yield a non-vanishing result, i.e. those that satisfy $|n - n'| \leq 2h$. In 2 we used the canonical commutation relations (7) to expand

$$x^{2h} = 2^{-h}(a + a^\dagger)^{2h} = \sum_{\ell, \ell' \geq 0, \, \ell + \ell' \leq 2h} \left(f_{h, \ell, \ell'} (a^\dagger a)^\ell a^{\ell'} + f'_{h, \ell, \ell'} (a^\dagger a)^\ell (a^\dagger)^{\ell'} \right).$$

In 3 we applied standard estimates for factorials: for example, when $n \leq n' \leq n + 2h$ we used the fact that $n^\ell \sqrt{\frac{n'!}{n!}} \leq n^\ell (n')^{(n'-n)/2} \leq (n')^{\ell + (n'-n)/2} \leq (n')^{2h} \leq (n + 2h)^{2h}$; moreover, we defined $F_h := \sum_{\ell, \ell' \geq 0, \, \ell + \ell' \leq 2h} \left(|f_{h, \ell, \ell'}| + |f'_{h, \ell, \ell'}| \right)$. Finally, 4 follows by choosing e.g. $d_h^{-1} = (2h + 1)^h F_h$. Since (A2) holds for all n, we conclude that $A_N \geq 0$ for all N, which completes the proof. $\square$

Remark. The inequality in Lemma 24 depends critically on the special properties of the canonical operators. In fact, there is no universal constant $d_k > 0$ that makes the general relation $(A + B)^k \geq d_k \left(A^k + B^k \right)$ true for all positive matrices $A, B \geq 0$. To see why this is the case, it suffices to consider two pure states $A = |\psi\rangle\langle\psi|$ and $B = |\phi\rangle\langle\phi|$. Setting $\lambda := 1 - |\langle\psi|\phi\rangle| \in [0, 1]$, it can be shown that the minimal eigenvalue of $(A + B)^k$ is λ^k, while that of $A^k + B^k = A + B$ is clearly λ. By Weyl's principle, the conjectured matrix inequality would imply that $\lambda^k \geq d_k \lambda$ for all $\lambda \in [0, 1]$, absurd.

Proposition 25. *Let $k \geq 0$ and $m \geq 1$ be integers; also, let $\varepsilon > 0$ be given. Then, there is a constant $c_{k,m}(\varepsilon) < \infty$ such that every m-mode quantum state ρ with finite k-moments $M_k(\rho)$, as defined by (31), also satisfies*

$$M'_k(\rho, \varepsilon) = \left\| \chi_\rho \right\|_{C^k(B(0,\varepsilon))} \leq c_{k,m}(\varepsilon) M_k(\rho).$$

In particular, according to (33), ρ has finite phase space moments of order up to k.

Proof. Let ρ be an m-mode quantum state. We start by considering the modified state $\sigma := \rho \boxplus |0\rangle\langle0|$ that is obtained by convolving it with the (multi-mode) vacuum state according to the rule (37) (for $\lambda = 1/2$). A first important observation is that the moments of σ and ρ are related. Namely,

$$M_k(\sigma) \leq M_k(\rho) \qquad \forall \, k \in [0, \infty). \tag{A3}$$

To see why, we pick a multi-index $n \in \mathbb{N}_0^m$ and evaluate the nth diagonal entries of σ with respect to the Fock basis. We obtain that

$$\begin{aligned}
\langle n|\sigma|n\rangle &\overset{1}{=} \langle n|\Delta(\rho \boxplus |0\rangle\langle0|)|n\rangle \\
&\overset{2}{=} \langle n|\Delta(\rho) \boxplus |0\rangle\langle0| \, |n\rangle \\
&= \sum_{\ell \in \mathbb{N}_0^m} \langle\ell|\rho|\ell\rangle \, \langle n| \, (|\ell\rangle\langle\ell| \boxplus |0\rangle\langle0|) \, |n\rangle \\
&\overset{3}{=} \sum_{\ell \geq n} \langle\ell|\rho|\ell\rangle \, 2^{-|\ell|} \binom{\ell}{n}.
\end{aligned}$$

Here, in 1 we introduced the dephasing operator in the Fock basis, whose action is given by $\Delta(X) := \sum_{k \in \mathbb{N}_0^m} |k\rangle\langle k| \, X \, |k\rangle\langle k|$. In 2 we observed that $\Delta(\omega \boxplus \delta) = \Delta(\omega) \boxplus \delta$

for all m-mode quantum states ω whenever $\delta = \Delta(\delta)$ is already diagonal in the Fock basis. To show this, first exploit linearity and factorisation of Δ to reduce to the one-mode case. Then, use the representation $\Delta(X) = \int_0^{2\pi} \frac{d\varphi}{2\pi} e^{i\varphi\, a^\dagger a} X \, e^{-i\varphi a^\dagger a}$, valid for bounded X and where the integrals are as usual weakly converging, and remember that $e^{i\varphi\left(a^\dagger a + b^\dagger b\right)} = e^{i\varphi a^\dagger a} \otimes e^{i\varphi b^\dagger b}$ is a function of the total Hamiltonian and thus commutes with the action of the beam splitter. The identity in 3 follows from the formula

$$|\ell\rangle\langle\ell| \boxplus |0\rangle\langle 0| = 2^{-|\ell|} \sum_{\ell' \leq \ell} \binom{\ell}{\ell'} |\ell'\rangle\langle\ell'|$$

for the convolution of a Fock state with the vacuum. Here, $\ell, \ell' \in \mathbb{N}_0^m$ are multi-indices, ordered entry-wise, and $\binom{\ell}{\ell'} := \prod_j \binom{\ell_j}{\ell'_j}$. The above expression can be obtained easily e.g. by first reducing to the one-mode case, and then by induction on ℓ, employing the relations (35). Computing the kth moment of σ then yields

$$M_k(\sigma) \overset{4}{=} \sum_{n \in \mathbb{N}_0^m} (m + |n|)^{k/2} \langle n|\sigma|n\rangle$$

$$\overset{5}{=} \sum_{\ell \in \mathbb{N}_0^m} \langle\ell|\rho|\ell\rangle \sum_{n \leq \ell} (m + |n|)^{k/2} 2^{-|\ell|} \binom{\ell}{n}$$

$$\overset{6}{\leq} \sum_{\ell \in \mathbb{N}_0^m} \langle\ell|\rho|\ell\rangle \, (m + |\ell|)^{k/2}$$

$$\overset{7}{=} M_k(\rho).$$

Here, 4 and 7 follow from the representation in (32); in 5 we rearranged a double series of non-negative terms, and in 6 we observed that for a given $\ell \in \mathbb{N}_0^m$ the coefficients $P_\ell(n) := 2^{-|\ell|} \binom{\ell}{n}$ form a probability distribution over the set of multi-indices $n \in \mathbb{N}_0^m$ with $n \leq \ell$. This proves that the kth moments of σ are upper bounded by those of ρ. The state σ is also useful because its characteristic function is a close relative of that of ρ. Namely, according to (38) we have that $\chi_\sigma(z) = \chi_\rho(z/\sqrt{2}) \, e^{-\|z\|^2/4}$, and hence

$$\left\| \chi_\rho \right\|_{C^k(B(0,\varepsilon))} \leq g_{k,m}(\varepsilon) \left\| \chi_\sigma \right\|_{C^k(B(0,\varepsilon))} \tag{A4}$$

for some constants $g_{k,m}(\varepsilon)$. Thus, it suffices to find a suitable upper estimate for the norm $\left\| \chi_\sigma \right\|_{C^k(B(0,\varepsilon))}$. By Lemma 16, the Fourier transform of χ_σ, i.e. the Wigner function W_σ of σ, is everywhere non-negative. Hence, χ_σ can be seen as the characteristic function of a classical random variable Z over $\mathbb{C}^m$, with probability density function W_σ. If we show that Z has finite absolute moments of order k, then thanks to [61, Theorem 1.8.15] we deduce that χ_σ is k-fold differentiable everywhere, and since

$$\left| \partial_z^\alpha \partial_{z^*}^\beta \chi_\sigma(z) \right| = \left| \int d^{2m}u \left(\prod_j (-u_j)^{\beta_j} (u_j^*)^{\alpha_j} \right) W_\sigma(u) \, e^{z^\intercal u^* - z^\dagger u} \right| \leq \int d^{2m}u \, \|u\|^{|\alpha|+|\beta|} \, W_\sigma(u) \tag{A5}$$

for all multi-indices $\alpha, \beta \in \mathbb{N}_0^m$, we in fact have that

$$\left\| \chi_\sigma \right\|_{C^k(B(0,\varepsilon))} \leq 1 + \int d^{2m}u \, \|u\|^k \, W_\sigma(u) =: 1 + L_k(\sigma). \tag{A6}$$

Therefore, we now look at the quantity $L_k(\sigma)$. For a vector $u = u_R + iu_I \in \mathbb{C}^m$, with $u_R, u_I \in \mathbb{R}^m$, we observe that

$$\|u\|^k = \left(\sum_j (u_{Rj}^2 + u_{Ij}^2)\right)^{k/2} \leq (2m)^{\max\left\{\frac{k}{2}-1, 0\right\}} \left(\sum_j \left(|u_{Rj}|^k + |u_{Ij}|^k\right)\right).$$

Thus,

$$L_k(\sigma) \leq (2m)^{\max\left\{\frac{k}{2}-1, 0\right\}} \int d^m u_R \, d^m u_I \sum_j \left(|u_{Rj}|^k + |u_{Ij}|^k\right) W_\sigma(u)$$

$$= (2m)^{\max\left\{\frac{k}{2}-1, 0\right\}} \left(\int d^m u_R \left(\sum_j |u_{Rj}|^k\right) \int d^m u_I \, W_\sigma(u)\right.$$

$$\left. + \int d^m u_I \left(\sum_j |u_{Ij}|^k\right) \int d^m u_R \, W_\sigma(u)\right)$$

$$\overset{8}{=} (2m)^{\max\left\{\frac{k}{2}-1, 0\right\}} 2^{-\frac{1}{2}(m+k-1)} \left(\mathrm{Tr}\left[\sigma \left(\sum_j |x_j|^k\right)\right] + \mathrm{Tr}\left[\sigma \left(\sum_j |p_j|^k\right)\right]\right)$$

$$\overset{9}{\leq} (2m)^{\max\left\{\frac{k}{2}-1, 0\right\}} 2^{-\frac{1}{2}(m+k-1)} d_{k,m}^{-1} \mathrm{Tr}\left[\sigma \left(\sum_j a_j a_j^\dagger\right)^{k/2}\right]$$

$$\overset{10}{\leq} c'_{k,m} M_k(\sigma).$$

In the above derivation, the identity in 8 can be verified by first reducing to the case of a pure σ, which can be done by linearity and by multiple applications of Tonelli's theorem, and by subsequently remembering that for a pure state $|\psi_f\rangle$ with wave function $f \in L^2(\mathbb{R}^m)$ it holds e.g. that $\int d^m u_I \, W_{|\psi_f\rangle\langle\psi_f|}(u) = \sqrt{2}\left|f(\sqrt{2}\,u_R)\right|^2$. The inequality in 9 is just an application of Lemma 24. Finally, in 10 we introduced a suitable constant $c'_{k,m} \geq 1$.

Combining the above estimate with (A4), (A6), and (A3), we deduce that

$$M_k(\rho, \varepsilon) = \|\chi_\rho\|_{C^k(B(0,\varepsilon))} \leq g_{k,m}(\varepsilon) \|\chi_\sigma\|_{C^k(B(0,\varepsilon))} \leq g_{k,m}(\varepsilon)(1 + L_k(\sigma))$$

$$\leq 2\, g_{k,m}(\varepsilon) c'_{k,m} M_k(\sigma) \leq 2\, g_{k,m}(\varepsilon) c'_{k,m} M_k(\rho) =: c_{k,m}(\varepsilon) M_k(\rho),$$

which concludes the proof.   □

Appendix B: Standard Moments Versus Phase Space Moments: The Fractional Case

In the last section, we showed that the kth phase space moment was controlled by the kth standard moment in the case of an integer constant k.

Here, we show that this fact still holds when k is a positive real number by an interpolation argument. In principle, we could conclude this fact from the setting of Proposition 25, using that for $L^p(w_0)$ spaces with weight function w_0 and $L^p(w_1)$ spaces with weight function w_1, the real interpolation spaces [63, Theorem 5.4.1] satisfy

$$\left(L^p(w_0), L^p(w_1)\right)_\theta = L^p(w_\theta)$$

where $w_\theta := w_0^{1-\theta} w_1^\theta$. This would allow us to extend the estimate in (A5) to fractional powers as well. However, we want to establish the stronger result that shows that the moments themselves naturally induce an interpolating family of normed spaces. That is, we show the following:

Proposition 26. *Let ρ be an m-mode quantum state and $k \geq 0$. If* $\mathrm{Tr}\left[\rho \left(\sum_j a_j a_j^\dagger\right)^{k/2}\right] <$
∞, then $\|\chi_\rho\|_{C^k(B(0,\varepsilon))} < \infty$ for some $\varepsilon > 0$. Moreover,

$$\|\chi_\rho\|_{C^k(B(0,\varepsilon))} \leq C_\varepsilon \, \mathrm{Tr}\left[\rho \left(\sum_j a_j a_j^\dagger\right)^{k/2}\right],$$

for some constant $C_\varepsilon > 0$.

We have seen in Appendix 8.3 that the map $\rho \mapsto \chi_\rho$ is bounded from $\mathscr{W}^{k,1}(\mathcal{H}_m)$ to $C^k(B(0,\varepsilon))$ for any k integer. Since the spaces $C^k(B(0,\varepsilon))$ form an interpolation family, meaning that for any $k_0, k_1 \in \mathbb{N}_0$ with $k_1 := k_0 + 1$, $C^{(1-\theta)k_0+\theta k_1}(B(0,\varepsilon)) = (C^{k_0}(B(0,\varepsilon)), C^{k_1}(B(0,\varepsilon)))_\theta$, we have from the previously mentioned interpolation method that

$$\|\chi_\rho\|_{C^{(1-\theta)k_0+\theta k_1}(B(0,\varepsilon))} \leq C_\varepsilon \, \|\rho\|_{(\mathscr{W}^{k_0,1}(\mathcal{H}_m), \mathscr{W}^{k_1,1}(\mathcal{H}_m))_\theta}, \tag{B1}$$

for some positive constant C_ε that comes from the bounds derived in Sect. 8.3 for k_0 and k_1. It only remains to prove that the interpolated norms $\|\rho\|_{(\mathscr{W}^{k_0,1}(\mathcal{H}_m), \mathscr{W}^{k_1,1}(\mathcal{H}_m))_\theta}$ can further be bounded above by $\|\rho\|_{\mathscr{W}^{(1-\theta)k_0+\theta k_1,1}}$. First, we recall a useful technical lemma [89, Lemma 3.4].

Lemma 27. *Let $T = \begin{pmatrix} T_{11} & T_{21} \\ T_{21}^\dagger & T_{22} \end{pmatrix}$ be a positive semi-definite trace class operator such that $T_{11} : \mathbb{C}^d \to \mathbb{C}^d$, then*

$$\|T_{21}\|_1 \leq \frac{1}{2} \left(\|T_{11}\|_1 + \|T_{22}\|_1\right).$$

Proof of Proposition 26. We provide the proof only for $m = 1$, since the general case follows similarly. First, observe that

$$\mathrm{Tr}\left[\rho(aa^\dagger)^{k/2}\right] = \sum_{n=0}^\infty \langle n|\rho|n\rangle \, (n+1)^{k/2}. \tag{B2}$$

First, we restrict attention to states ρ that are orthogonal in the Fock basis. We then write Π_E for the spectral projection onto the Fock states of energy at most E, that is $\Pi_{\lfloor E \rfloor} := \sum_{j \leq \lfloor E \rfloor} |j\rangle\langle j|$. Next, we fix two parameters $0 \leq k_0 \leq k_1$ and introduce the quantity $\gamma_n := (n+1)^{(k_1-k_0)/2}$, fix a parameter $t > 0$, define $N_0(t) \in \mathbb{N}$ such that

$$\forall n \leq N_0(t) : \gamma_n \leq t^{-1} \text{ and } \gamma_{N_0(t)+1} \geq t^{-1},$$

the two operators

$$X_0(t) := (I - \Pi_{N_0(t)})\rho\,(I - \Pi_{N_0(t)}) \geq 0 \text{ and } X_1(t) := \Pi_{N_0(t)}\rho\,\Pi_{N_0(t)} \geq 0$$

and $\rho \equiv \rho_{\mathrm{diag}}(t) := X_0(t) + X_1(t)$. Using these two operators we start estimating

$$K(t, \rho_{\text{diag}}(t)) := \inf_{\rho_{\text{diag}}(t) = X_0 + X_1} \|X_0\|_{\mathscr{W}^{k_0,1}(\mathcal{H}_1)} + t \|X_1\|_{\mathscr{W}^{k_1,1}(\mathcal{H}_1)}$$

$$\leq \|X_0(t)\|_{\mathscr{W}^{k_0,1}(\mathcal{H}_1)} + t \|X_1(t)\|_{\mathscr{W}^{k_1,1}(\mathcal{H}_1)}$$

$$= \left(\sum_{n > N_0(t)} \langle n|\rho_{\text{diag}}(t)|n\rangle (n+1)^{k_0/2} + t \sum_{n \leq N_0(t)} \langle n|\rho_{\text{diag}}(t)|n\rangle (n+1)^{k_0/2} \gamma_n \right)$$

$$= \sum_{n=0}^{\infty} \langle n|\rho_{\text{diag}}(t)|n\rangle (n+1)^{k_0/2} (\alpha_n + t\gamma_n \beta_n)$$

where $\alpha_n = \delta_{n > N_0(t)}$ and $\beta_n = \delta_{n \leq N_0(t)}$ with Kronecker delta δ. Thus, we obtain for the norm $\|\rho\|_{(\mathscr{W}^{k_0,1}(\mathcal{H}_1), \mathscr{W}^{k_1,1}(\mathcal{H}_1))_\theta}$ the upper bound

$$\|\rho\|_{(\mathscr{W}^{k_0,1}(\mathcal{H}_1), \mathscr{W}^{k_1,1}(\mathcal{H}_1))_\theta} \leq \sup_{t > 0} \sum_{n=0}^{\infty} \langle n|\rho_{\text{diag}}(t)|n\rangle (n+1)^{k_0/2} t^{-\theta} (\alpha_n + t\gamma_n \beta_n).$$

We now recall that for $\gamma_n \leq t^{-1}$ we have $\alpha_n = 0$ and $\beta_n = 1$ such that

$$t^{-\theta} (\alpha_n + t\gamma_n \beta_n) = t^{-\theta} t \gamma_n = t^{1-\theta} \gamma_n^{1-\theta} \gamma_n^{\theta} \leq \gamma_n^{\theta}.$$

For $\gamma_n > t^{-1}$ we have $\alpha_n = 1$ and $\beta_n = 0$ such that

$$t^{-\theta} (\alpha_n + t\gamma_n \beta_n) = t^{-\theta} \leq \gamma_n^{\theta}.$$

Thus, in either case, we have the estimate

$$\|\rho\|_{(\mathscr{W}^{k_0,1}(\mathcal{H}_1), \mathscr{W}^{k_1,1}(\mathcal{H}_1))_\theta} \leq \sum_{n=0}^{\infty} \langle n|\rho|n\rangle (n+1)^{((1-\theta)k_0 + \theta k_1)/2}.$$

This shows that for arbitrary density operators

$$\|\rho\|_{(\mathscr{W}^{k_0,1}(\mathcal{H}_1), \mathscr{W}^{k_1,1}(\mathcal{H}_1))_\theta} \leq \|\rho\|_{\mathscr{W}^{(1-\theta)k_0 + \theta k_1,1}(\mathcal{H}_1)}.$$

To extend the bound to a density operator ρ that is not diagonal in the Fock basis, and not only for the diagonal $\rho_{\text{diag}}(t)$, we partition ρ as

$$\rho = \begin{pmatrix} X_0(t) & \rho_{21}(t)^\dagger \\ \rho_{21}(t) & X_1(t) \end{pmatrix}$$

and a self-adjoint diagonal operator $S^{(k)}(t) := \text{diag}\left(S_1^{(k)}(t), S_2^{(k)}(t) \right)$ where $S_1^{(k)}(t) := \Pi_{N_0(t)}(aa^\dagger)^{k/4}\Pi_{N_0(t)}$ and $S_2^{(k)}(t) := (I - \Pi_{N_0(t)})(aa^\dagger)^{k/4}(I - \Pi_{N_0(t)})$. This implies that

$$T^{(k)} := S^{(k)}\rho S^{(k)} = \begin{pmatrix} S_1^{(k)}(t)X_0(t)S_1^{(k)}(t) & \left(S_1^{(k)}(t)\rho_{21}(t)S_2^{(k)}(t) \right)^\dagger \\ S_1^{(k)}(t)\rho_{21}(t)S_2^{(k)}(t) & S_2^{(k)}(t)\rho_{22}(t)S_2^{(k)}(t) \end{pmatrix}.$$

Let then $S_1^{(k)}(t) := \Pi_{N_0(t)}(aa^\dagger)^{k/4}\Pi_{N_0(t)}$ and $S_2^{(k)}(t) := (I - \Pi_{N_0(t)})(aa^\dagger)^{k/4}(I - \Pi_{N_0(t)})$. The previous Lemma 27 then shows that

$$\|S_1^{(k)}(t)\rho_{21}S_2^{(k)}(t)\|_1 \leq \frac{1}{2}\left(\|S_1^{(k)}(t)\rho_{11}S_1^{(k)}(t)\|_1 + \|S_2^{(k)}(t)\rho_{22}S_2^{(k)}(t)\|_1 \right).$$

From here, we examine three cases separately:

- Case 1: $\|T_{11}^{(k_1)}\|_1 \geq \|T_{22}^{(k_1)}\|_1$. In this case, we find from choosing $X_0 := \rho_{21}$ and $X_1 := 0$ in (53)

$$K(t, \rho_{21}) \leq \|\rho_{21}(t)\|_{\mathscr{W}^{k_0,1}(\mathcal{H}_1)} \leq \|X_0(t)\|_{\mathscr{W}^{k_0,1}(\mathcal{H}_1)}.$$

- Case 2: $\|T_{22}^{(k_0)}\|_1 \geq \|T_{11}^{(k_0)}\|_1$ In this case, we find from choosing $X_0 := 0$ and $X_1 := \rho_{21}$ in (53)

$$K(t, \rho_{21}) \leq t\,\|\rho_{21}\|_{\mathscr{W}^{k_1,1}(\mathcal{H}_1)} \leq t\,\|X_1(t)\|_{\mathscr{W}^{k_1,1}(\mathcal{H}_1)}.$$

- Case 3: $\|T_{22}^{(k_0)}\|_1 \leq \|T_{11}^{(k_0)}\|_1$ and $\|T_{22}^{(k_1)}\|_1 \geq \|T_{11}^{(k_1)}\|_1$ In this case, we find from choosing $X_0 = \rho_{21}/2$ and $X_1 = \rho_{21}/2$ in (53) that

$$
\begin{aligned}
K(t, \rho_{21}) &\leq \frac{\|\rho_{21}\|_{\mathscr{W}^{k_0,1}(\mathcal{H}_1)} + t\,\|\rho_{21}\|_{\mathscr{W}^{k_1,1}(\mathcal{H}_1)}}{2} \\
&\leq \frac{1}{2}\left(\|X_0(t)\|_{\mathscr{W}^{k_0,1}(\mathcal{H}_1)} + t\,\|X_1(t)\|_{\mathscr{W}^{k_1,1}(\mathcal{H}_1)}\right).
\end{aligned}
\tag{B3}
$$

Hence, we have altogether that

$$
\begin{aligned}
K(t, \rho(t)) &\leq K(t, \rho_{\mathrm{diag}}(t)) + 2K(t, \rho_{21}(t)) \\
&\leq 3\left(\|X_0(t)\|_{\mathscr{W}^{k_0,1}(\mathcal{H}_1)} + t\,\|X_1(t)\|_{\mathscr{W}^{k_1,1}(\mathcal{H}_1)}\right)
\end{aligned}
$$

which implies that

$$\|\rho\|_{(\mathscr{W}^{k_0,1}(\mathcal{H}_1),\,\mathscr{W}^{k_1,1}(\mathcal{H}_1))_\theta} \leq 3\|\rho\|_{\mathscr{W}^{(1-\theta)k_0+\theta k_1,1}(\mathcal{H}_1)}.$$

The result follows from the interpolation bound (B1).

$\square$

Appendix C: Standard Moments Versus Phase Space Moments: A Partial Converse

We now show that at least for even integers k, the existence of kth order phase space moments implies the existence of standard moments of the same order.

Theorem 28. *Let ρ be an m-mode quantum state such that its characteristic function χ_ρ is $2k$ times totally differentiable at $z = 0$ for some integer k, then the $2k^{\text{th}}$ standard moment is finite as well.*

Proof. For simplicity, we restrict attention to $m = 1$. Let $H_{\mathrm{lin}}^+ := a + a^\dagger$ and $H_{\mathrm{lin}}^+ := (-i)(a - a^\dagger)$ be two Hamiltonians, and consider the spectral decomposition of the density operator $\rho = \sum_{i=1}^\infty \lambda_i\,|e_i\rangle\langle e_i|$. Then, there exist unique probability measures μ_{e_i} such that

$$\langle e_i|f(H_{\mathrm{lin}}^\pm)|e_i\rangle = \int_{\sigma(H_{\mathrm{lin}}^\pm)} f(\lambda)\,d\mu_{e_i}(\lambda) \quad \text{for all } f \text{ bounded measurable.}$$

We then define the new probability measure $\mu_\rho := \sum_{i=1}^\infty \lambda_i \mu_{e_i}$ such that

$$\mathrm{Tr}\left[\rho\,f(H_{\mathrm{lin}}^\pm)\right] = \int_{\sigma(H_{\mathrm{lin}}^\pm)} f(\lambda)\,d\mu_\rho(\lambda) \quad \text{for all } f \text{ bounded measurable.}$$

We now proceed with an induction argument. Start by noting that for $k = 0$ the result holds. For $k \geq 1$, define the auxiliary function $\varphi : \mathbb{R} \to \mathbb{C}$ as

$$\varphi(t) := \text{Tr} \left[\rho \, e^{it H_{\text{lin}}^{\pm}} \right],$$

which is by assumption $2k$ times differentiable at zero and let $u(t) = \Re\varphi(t)$. Then, u is also $2k$ times differentiable at zero. Since $\varphi^{2k}(0)$ exists, for $t \in (-\varepsilon, \varepsilon)$, with sufficiently small $\varepsilon > 0$, the function $t \mapsto \varphi^{(2k-1)}(t)$ exists and is continuous.
We record that Taylor's formula implies that for $t \in (-\varepsilon, \varepsilon)$

$$\left| u(t) - \sum_{i=0}^{k-1} u^{(2i)}(0) \frac{t^{2i}}{(2i)!} \right| \leq \frac{|t|^{2k-1}}{(2k-1)!} \sup_{\theta \in (0,1]} \left| u^{(2k-1)}(\theta t) \right|,$$

where odd derivatives vanish at zero, since u is even.
We then define a positive continuous function $f_k : \mathbb{R} \to [0, \infty)$ with $f_k(0) = 1$ and for $t \neq 0$ as

$$f_k(t) := (-1)^k (2k)! \, t^{-2k} \left(\cos(t) - \sum_{i=0}^{k-1} (-1)^i \frac{t^{2i}}{(2i)!} \right).$$

From Taylor's formula above we obtain the following estimate for t sufficiently small

$$
\begin{aligned}
\text{Tr} \left[\rho \, f_k(t H_{\text{lin}}^{\pm}) H_{\text{lin}}^{\pm \, 2k} \right] &= \int_{\sigma(H_{\text{lin}}^{\pm})} f_k(t\lambda) \lambda^{2k} d\mu_\rho(\lambda) \\
&= \frac{(-1)^k (2k)!}{t^{2k}} \left(u(t) - \sum_{i=0}^{k-1} (-1)^i \frac{\text{Tr} \left[\rho \, (t H_{\text{lin}}^{\pm})^{2i} \right]}{(2i)!} \right) \\
&\leq 2k \sup_{\theta \in (0,1]} \frac{|u^{2k-1}(\theta t)|}{\theta |t|} =: g_k(t).
\end{aligned}
$$

Then, we have from Fatou's lemma

$$
\begin{aligned}
\text{Tr} \left[\rho \left(H_{\text{lin}}^{\pm} \right)^{2k} \right] &= \text{Tr} \left[\rho \, f_k(0) \left(H_{\text{lin}}^{\pm} \right)^{2k} \right] \\
&\leq \liminf_{t \downarrow 0} \int_{\sigma(H_{\text{lin}}^{\pm})} f_k(t\lambda) \lambda^{2k} d\mu_\rho(\lambda) \\
&= \liminf_{t \downarrow 0} g_k(t) = 2k |u^{2k}(0)| < \infty.
\end{aligned}
$$

Using integration by parts and standard estimates only, it is straightforward to verify that the finiteness of both $\text{Tr} \left[\rho \, H_{\text{lin}^{\pm}}^{2k} \right]$ implies the finiteness of $\text{Tr} \left[\rho (a a^{\dagger})^k \right]$. $\quad \square$

References

1. Feller, W.: An Introduction to Probability Theory and Its Applications, vol. II, 2nd edn. Wiley, New York (1971)
2. Cushen, C.D., Hudson, R.L.: A quantum-mechanical Central Limit theorem. J. Appl. Probab. **8**(3), 454–469 (1971)
3. Hepp, K., Lieb, E.H.: Phase-transitions in reservoir-driven open systems with applications to lasers and superconductors. Helv. Phys. Acta **46**(5), 573–603 (1974)
4. Hepp, K., Lieb, E.H.: On the superradiant phase transition for molecules in a quantized radiation field: the Dicke Maser model. Ann. Phys. **76**(2), 360–404 (1973)
5. Giri, N., von Waldenfels, W.: An algebraic version of the Central Limit theorem. Z. Wahrscheinlichkeitstheorie verw. Gebiete **42**(2), 129–134 (1978)
6. Goderis, D., Vets, P.: Central limit theorem for mixing quantum systems and the CCR-algebra of fluctuations. Commun. Math. Phys. **122**(2), 249–265 (1989)
7. Goderis, D., Verbeure, A., Vets, P.: About the mathematical theory of quantum fluctuations. Leuven University Press, Leuven, Belgium (1989)
8. Matsui, T.: Bosonic Central Limit Theorem for the one-dimensional XY model. Rev. Math. Phys. **14**(07n08), 675–700 (2002)
9. Cramer, M., Eisert, J.: A quantum Central Limit Theorem for non-equilibrium systems: exact local relaxation of correlated states. New J. Phys. **12**(5), 055020 (2010)
10. Goderis, D., Verbeure, A., Vets, P.: About the exactness of the linear response theory. Commun. Math. Phys. **136**(2), 265–283 (1991)
11. Jakšić, V., Pautrat, Y., Pillet, C.-A.: Central Limit Theorem for locally interacting Fermi gas. Commun. Math. Phys. **285**(1), 175–217 (2009)
12. Arous, G.B., Kirkpatrick, K., Schlein, B.: A central limit theorem in many-body quantum dynamics. Commun. Math. Phys. **321**(2), 371–417 (2013)
13. Brandão, F.G.S.L., Cramer, M.: Equivalence of statistical mechanical ensembles for non-critical quantum systems (2015)
14. Brandão, F.G.S.L., Gour, G.: Reversible framework for quantum resource theories. Phys. Rev. Lett. **115**, 070503 (2015)
15. Dereziński, J.: Boson free fields as a limit of fields of a more general type. Rep. Math. Phys. **21**(3), 405–417 (1985)
16. Streater, R.F.: Entropy and the Central Limit Theorem in quantum mechanics. J. Phys. A **20**(13), 4321–4330 (1987)
17. Michoel, T., Nachtergaele, B.: Central Limit Theorems for the large-spin asymptotics of quantum spins. Probab. Theory Relat. Fields **130**(4), 493–517 (2004)
18. Goderis, D., Verbeure, A., Vets, P.: Non-commutative central limits. Probab. Theory Related Fields **82**(4), 527–544 (1989)
19. Jakšić, V., Pautrat, Y., Pillet, C.-A.: A quantum central limit theorem for sums of independent identically distributed random variables. J. Math. Phys. **51**(1), 015208 (2010)
20. Voiculescu, D.V., Dykema, K.J., Nica, A.: Free Random Variables, CRM Monograph Series. American Mathematical Society, Providence (1992)
21. Accardi, L., Lu, Y.G.: Quantum central limit theorems for weakly dependent maps. II. Acta Math. Hungar. **63**(3), 249–282 (1994)
22. Hayashi, M.: Quantum Information: An Introduction. Springer, Berlin Heidelberg (2006)
23. Hayashi, M.: Quantum estimation and the Quantum Central Limit theorem. Am. Math. Soci. Transl. Ser. **2**(227), 95–123 (2009)
24. Campbell, E.T., Genoni, M.G., Eisert, J.: Continuous-variable entanglement distillation and noncommutative central limit theorems. Phys. Rev. A **87**, 042330 (2013)
25. Lenczewski, R.: Quantum Central Limit Theorems, pp. 299–314. Springer, Boston (1995)
26. Dimi, A., Daki, B.: On the central limit theorem for unsharp quantum random variables. New J. Phys. **20**, 063051 (2018)
27. Davies, E.B.: Quantum stochastic processes. Commun. Math. Phys. **15**(4), 277–304 (1969)
28. Lami, L., Sabapathy, K.K., Winter, A.: All phase-space linear bosonic channels are approximately Gaussian dilatable. New J. Phys. **20**(11), 113012 (2018)
29. Carolan, J., Harrold, C., Sparrow, C., Martín-López, E., Russell, N.J., Silverstone, J.W., Shadbolt, P.J., Matsuda, N., Oguma, M., Itoh, M., Marshall, G.D., Thompson, M.G., Matthews, J.C.F., Hashimoto, T., O'Brien, J.L., Laing, A.: Universal linear optics. Science **349**(6249), 711–716 (2015)
30. Rohde, P.P., Dowling, J.P.: The on-ramp to the all-optical quantum information processing highway. Science **349**(6249), 696 (2015)
31. Holevo, A.S., Shirokov, M.E.: Continuous ensembles and the capacity of infinite-dimensional quantum channels. Theory Probab. Its Appl. **50**(1), 86–98 (2006)

32. Wilde, M.M., Qi, H.: Energy-constrained private and quantum capacities of quantum channels. IEEE Trans. Inf. Theory **64**(12), 7802–7827 (2018)
33. Winter, A.: Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. Commun. Math. Phys. **347**(1), 291–313 (2016)
34. Winter, A.: Energy-constrained diamond norm with applications to the uniform continuity of continuous variable channel capacities (2017). arXiv:1712.10267
35. Shirokov, M.E.: Tight uniform continuity bounds for the quantum conditional mutual information, for the Holevo quantity, and for capacities of quantum channels. J. Math. Phys. **58**(10), 102202 (2017)
36. Shirokov, M.E.: On the energy-constrained diamond norm and its application in quantum information theory. Prob. Inform. Transm. **54**(1), 20–33 (2018)
37. Giovannetti, V., García-Patrón, R., Cerf, N.J., Holevo, A.S.: Ultimate classical communication rates of quantum optical channels. Nat. Photonics **8**(10), 796–800 (2014)
38. Giovannetti, V., Holevo, A.S., García-Patrón, R.: A solution of Gaussian optimizer conjecture for quantum channels. Commun. Math. Phys. **334**(3), 1553–1571 (2015)
39. Holevo, A.S., Werner, R.F.: Evaluating capacities of bosonic Gaussian channels. Phys. Rev. A **63**, 032312 (2001)
40. Pirandola, S., Laurenza, R., Ottaviani, C., Banchi, L.: Fundamental limits of repeaterless quantum communications. Nat. Commun. **8**(1), 15043 (2017)
41. Wilde, M.M., Tomamichel, M., Berta, M.: Converse bounds for private communication over quantum channels. IEEE Trans. Inf. Theory **63**(3), 1792–1817 (2017)
42. Rosati, M., Mari, A., Giovannetti, V.: Narrow bounds for the quantum capacity of thermal attenuators. Nat. Commun. **9**(1), 4339 (2018)
43. Sharma, K., Wilde, M.M., Adhikari, S., Takeoka, M.: Bounding the energy-constrained quantum and private capacities of phase-insensitive bosonic Gaussian channels. New J. Phys. **20**(6), 063025 (2018)
44. Noh, K., Albert, V.V., Jiang, L.: Quantum capacity bounds of Gaussian thermal loss channels and achievable rates with Gottesman–Kitaev–Preskill codes. IEEE Trans. Inf. Theory **65**(4), 2563–2582 (2019)
45. Noh, K., Pirandola, S., Jiang, L.: Enhanced energy-constrained quantum communication over bosonic Gaussian channels. Nat. Commun. **11**(1), 457 (2020)
46. Schmuedgen, K.: Unbounded Self-adjoint Operators on Hilbert Space. Springer, Dordrecht (2012)
47. Umegaki, H.: Conditional expectation in an operator algebra. IV. Entropy and information. Kodai Math. Sem. Rep. **14**(2), 59–85 (1962)
48. Lindblad, G.: Entropy, information and quantum measurements. Commun. Math. Phys. **33**(4), 305–322 (1973)
49. Holevo, A.S.: On capacity of a quantum communications channel. Probl. Pered. Inform. **15**(4):3–11 (1979). (English translation: Probl. Inf. Transm. 15(4):247–253)
50. Hausladen, P., Jozsa, R., Schumacher, B., Westmoreland, M., Wootters, W.K.: Classical information capacity of a quantum channel. Phys. Rev. A **54**, 1869–1876 (1996)
51. Schumacher, B., Westmoreland, M.D.: Sending classical information via noisy quantum channels. Phys. Rev. A **56**, 131–138 (1997)
52. Holevo, A.S.: The capacity of the quantum channel with general signal states. IEEE Trans. Inf. Theory **44**(1), 269–273 (1998)
53. Lloyd, S.: Capacity of the noisy quantum channel. Phys. Rev. A **55**, 1613–1622 (1997)
54. Shor, P.: Lecture Notes. MSRI Workshop on Quantum Computation (2002)
55. Devetak, I.: The private classical capacity and quantum capacity of a quantum channel. IEEE Trans. Inf. Theory **51**(1), 44–55 (2005)
56. de Gosson, M.A.: Symplectic Geometry and Quantum Mechanics. Operator Theory: Advances and Applications. Birkhäuser, Basel (2006)
57. Serafini, A.: Quantum Continuous Variables: A Primer of Theoretical Methods. CRC Press, Taylor & Francis Group (2017)
58. Holevo, A.S.: Probabilistic and Statistical Aspects of Quantum Theory. Publications of the Scuola Normale Superiore, Scuola Normale Superiore (2011)
59. Barnett, S., Radmore, P.M.: Methods in Theoretical Quantum Optics Oxford Series in Optical and Imaging Sciences. Clarendon Press, Oxford (2002)
60. Zygmund, A.: A remark on characteristic functions. Ann. Math. Stat. **18**(2), 272–276 (1947)
61. Ushakov, N.G.: Selected Topics in Characteristic Functions. Modern Probability and Statistics. de Gruyter, Berlin (2011)
62. König, R., Smith, G.: The entropy power inequality for quantum systems. IEEE Trans. Inf. Theory **60**(3), 1536–1548 (2014)
63. Bergh, J., Löfström, J.: Interpolation Spaces: An Introduction, vol. 223. Springer, Berlin (2012)
64. Sabapathy, K.K., Winter, A.: Non-Gaussian operations on bosonic modes of light: photon-added Gaussian channels. Phys. Rev. A **95**, 062309 (2017)

65. O'Brien, J.L., Furusawa, A., Vučković, J.: Photonic quantum technologies. Nat. Photonics **3**(12), 687–695 (2009)
66. Politi, A., Matthews, J.C.F., Thompson, M.G., O'Brien, J.L.: Integrated quantum photonics. IEEE J. Sel. Top. Quantum Electron. **15**(6), 1673–1684 (2009)
67. Folland, G.B.: Harmonic Analysis in Phase Space. Princeton University Press, Princeton (1989)
68. Jagannathan, R., Simon, R., Sudarshan, E.C.G., Vasudevan, R.: Dynamical maps and nonnegative phase-space distribution functions in quantum mechanics. Phys. Lett. A **120**(4), 161–164 (1987)
69. Durrett, R.: Probability: Theory and Examples. Cambridge Series in Statistical and Probabilistic Mathematics, 5th edn. Cambridge University Press, Cambridge (2019)
70. Kennard, E.H.: Zur Quantenmechanik einfacher Bewegungstypen. Z. Phys. **44**(4), 326–352 (1927)
71. Stoler, D.: Equivalence classes of minimum uncertainty packets. Phys. Rev. D **1**, 3217–3219 (1970)
72. Yuen, H.P.: Two-photon coherent states of the radiation field. Phys. Rev. A **13**, 2226–2243 (1976)
73. Winter, A.: Coding theorem and strong converse for quantum channels. IEEE Trans. Inf. Theory **45**(7), 2481–2485 (1999)
74. Williamson, J.: On the algebraic problem concerning the normal forms of linear dynamical systems. Am. J. Math. **58**(1), 141–163 (1936)
75. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables. Applied mathematics series. Dover Publications, Mineola (1965)
76. Schrödinger, E.: Der stetige Übergang von der Mikro- zur Makromechanik. Naturwissenschaften **14**(28), 664–666 (1926)
77. Klauder, J.R.: The action option and a Feynman quantization of spinor fields in terms of ordinary c-numbers. Ann. Phys. (N.Y) **11**(2), 123–168 (1960)
78. Glauber, R.J.: Coherent and incoherent states of the radiation field. Phys. Rev. **131**, 2766–2788 (1963)
79. Sudarshan, E.C.G.: Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. Phys. Rev. Lett. **10**, 277–279 (1963)
80. Lindeberg, J.W.: Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. Math. Z. **15**(1), 211–225 (1922)
81. Bryc, W.: A remark on the connection between the large deviation principle and the central limit theorem. Statist. Probab. Lett. **18**(4), 253–256 (1993)
82. Bhattacharya, R.N., Rao, R.R.: Normal approximation and asymptotic expansions, vol. 64. SIAM, Philadelphia (1986)
83. Lami, L., Khatri, S., Adesso, G., Wilde, M.M.: Extendibility of bosonic Gaussian states. Phys. Rev. Lett. **123**, 050501 (2019)
84. Devetak, I., Shor, P.W.: The capacity of a quantum channel for simultaneous transmission of classical and quantum information. Commun. Math. Phys. **256**(2), 287–303 (2005)
85. Caruso, F., Giovannetti, V., Holevo, A.S.: One-mode bosonic Gaussian channels: a full weak-degradability classification. New J. Phys. **8**(12), 310 (2006)
86. Wolf, M.M., Pérez-García, D., Giedke, G.: Quantum capacities of bosonic channels. Phys. Rev. Lett. **98**, 130501 (2007)
87. Geršgorin, S.: Über die Abgrenzung der Eigenwerte einer Matrix. Izv. Akad. Nauk. S.S.S.R **7**, 749–754 (1931)
88. Varga, R.S.: Geršgorin and his circles. Springer, Berlin (2010)
89. Bourin, J.-C., Lee, E.-Y.: Unitary orbits of Hermitian operators with convex or concave functions. Bull. Lond. Math. Soc. **44**(6), 1085–1102 (2012)