Convergence rates for the quantum central limit theorem

Simon Becker

Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences University of Cambridge, Cambridge CB3 0WA, United Kingdom

Nilanjana Datta

Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WA, United Kingdom

Ludovico Lami

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, United Kingdom and

Institute of Theoretical Physics and IQST, Universität Ulm, Albert-Einstein-Allee 11D-89069 Ulm, Germany

Cambyse Rouzé

Zentrum Mathematik, Technische Universität München, 85748 Garching, Germany

(Dated: December 2019)

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, which is due to Cushen and Hudson, is of particular relevance in the framework of quantum optics. It implies that the state in any single output arm of an \( n \)-splitter, which is fed with \( n \) identical copies of a centred state \( \rho \) with finite second moments, converges to the Gaussian state with the same first and second moments as \( \rho \). In this paper, we obtain finite sample size bounds on the convergence rates in this quantum central limit theorem. We exploit the phase space formalism of characteristic functions to carry out a refined analysis of the rate of convergence. For instance, we prove that the convergence takes place at a rate \( \Theta \left( n^{-1/2} \right) \) in the Hilbert–Schmidt norm whenever the third moments of \( \rho \) are finite. Trace norm or relative entropy bounds can be obtained by leveraging the energy boundedness of the state. We provide analytical and numerical examples showing that our results are tight in many respects. A suitable 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 single-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 channel, with a rate \( \tilde{\Theta} \left( n^{-1/4} \frac{1}{\sqrt{m}} \right) \). This allows us to establish lower and upper 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 \( \chi_\rho \) is uniformly bounded by some \( \eta_\rho < 1 \) outside of any neighbourhood of the origin; even more, we point out that \( \eta_\rho \) can be made to depend only on the energy of the state \( \rho \).
I. 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 convergence. 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 generalized 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 [6–12], quantum field theory [13–15], von Neumann algebras [16, 17], free probability [18], noncommutative stochastic processes [19] and quantum information theory [20, 21]. For a more detailed list of papers on noncommutative or quantum central limit theorems (QCLT), see for example [17], [22] and references therein.

An important pair of non-commuting observables is the pair \((q, p)\) of canonically conjugate operators, which obey Heisenberg’s canonical commutation relations (CCR) \([q, p] = i\) where \(I\) denotes the identity operator. 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 \(a := (q + ip)/\sqrt{2}\) and \(a^\dagger := (q - ip)/\sqrt{2}\). 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 \(\rho\) is uniquely identified by its characteristic function, defined for all \(z \in \mathbb{C}\) by \(\chi_\rho(z) := \text{Tr} \left[ \rho e^{za^\dagger - z^*a} \right]\). The special class of Gaussian states comprises all quantum states whose characteristic function is the (classical) characteristic functions of a normal random variable on \(\mathbb{C}\). 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 \(\{(q_n, p_n); n = 1, 2, 3, \ldots\}\), 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 \(\rho\) (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 \(F_{jk} := e^{2\pi i/\sigma}^{-jk}\) is the discrete Fourier transform matrix. Passivity here means that \(U_{n\text{-split}}\) commutes with the canonical Hamiltonian of the field, i.e. \([U_{n\text{-split}}, \sum_j a_j^\dagger a_j] = 0\). When \(n\) identical copies of a state \(\rho\) 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 \(\rho\), and denoted by

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

\[2\]
The characteristic function of a random variable \(X \in \mathbb{R} \subseteq \mathbb{C}\) is defined by \(\chi_X(z) := E[e^{zX^* - z^*X}]\).
\(\rho^{\text{fin}}\). Observe that the same state can also be obtained as the output of a cascade of \(n-1\) beam splitters with suitably tuned transmissivities \(\lambda_j = j/(j+1)\) for \(j = 1, 2, \ldots n-1\) (see Figure 1a).

Cushen and Hudson’s result is that if \(\rho\) is a centred state (i.e. with zero mean) and finite second moments, its convolutions \(\rho^{\text{fin}}\) converge to the Gaussian state \(\rho_G\) with the same first and second moments as \(\rho\) in the limit \(n \to \infty\) (Theorem 5). 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 [23], or [24, 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 \(\rho^{\text{fin}}\) to the ‘Gaussification’ \(\rho_G\) of \(\rho\), obtained under suitable assumptions on \(\rho\) – 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 \(\rho^{\text{fin}} - \rho_G\) in the limit of large \(n\), as well as bounds on the relative entropy of \(\rho^{\text{fin}}\) with respect to \(\rho_G\) in the same limit. For example, we prove that if \(\rho\) has finite third-order phase space moments (Theorem 7). From this we are able to derive convergence estimates with respect to both the trace norm and the relative entropy, which read
\[
\|\rho^{\text{fin}} - \rho_G\|_2 = \mathcal{O}(n^{-1/(2m+2)}) \quad \text{and} \quad D(\rho^{\text{fin}}\|\rho_G) = \mathcal{O}(n^{-1/(2m+2)}), \quad \text{respectively (Corollary 8).}
\]

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 \(\sigma\) such that \(\text{Tr}\left[\sigma (a^\dagger a)^{1-\delta}\right]\) is finite for all \(\delta > 0\) (and infinite for \(\delta = 0\), yet \(\sigma^{\text{fin}}\) 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 \(\rho\), which is assumed to have bounded energy. We are able to show that for \(n \to \infty\) such a cascade channel converges, up to irrelevant symplectic unitaries, to a thermal attenuator channel with transmissivity \(\lambda\) and the same environment photon number as that of \(\rho\). 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, obtained by Winter [25, 26] and Shirokov [27, 28], with the known formulae expressing or estimating energy-constrained classical or quantum capacities of thermal attenuator channels [29, 32], 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 \(\rho\) 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 \(\|z\| \geq \varepsilon\) (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 states \( \rho \) with mean energy \( E \). We then prove that \( |\chi_\rho(z)| \leq 1 - \frac{c}{E} \) 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 Section II we introduce the notation and definitions used in the paper. In Section III we recall the Cushen and Hudson quantum central limit theorem. Our main results are presented in Section V. The rest of the paper is devoted to the proofs of these results. We start with the novel properties of quantum characteristic functions (Section V), which lie at the heart of our approach. Then, in Section VI we prove our quantum Berry–Esseen theorems. Section VII is devoted to the discussion of the optimality and sharpness of our results. In Section VIII we apply our quantitative non-i.i.d. extension of the Cushen–Hudson theorem to an optical fibre subjected 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.

**II. NOTATION AND DEFINITIONS**

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

Let $\mathcal{H}$ denote a separable Hilbert space, and let $B(\mathcal{H})$ denote the set of bounded linear operators acting on $\mathcal{H}$. Let $D(\mathcal{H})$ denote the set of quantum states of a system with Hilbert space $\mathcal{H}$ that is the set of density operators $\rho$ (positive semi-definite, i.e. $\rho \geq 0$, trace class operators with unit trace) acting on $\mathcal{H}$. We denote by $\| \cdot \|_{T_p(\mathcal{H})} \equiv \| \cdot \|_p$ the Schatten $p$-norm, defined as $\|X\|_p = (\text{Tr}|X|^p)^{1/p}$. The Schatten $p$-class $T_p(\mathcal{H})$ is the Banach subspace of $B(\mathcal{H})$ formed by all bounded linear operators whose Schatten $p$-norm is finite. We shall hereafter refer to $T_1(\mathcal{H})$ 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 operators defined on some domains $\text{Dom}(A), \text{Dom}(B) \subseteq \mathcal{H}$. According to [33, Definition 10.15], we write that $A \geq B$ if and only if $\text{Dom}(A^{1/2}) \subseteq \text{Dom}(B^{1/2})$ and $\|A^{1/2} \psi\|^2 \geq \|B^{1/2} \psi\|^2$ for all $\psi \in \text{Dom}(A^{1/2})$. Now, let $A$ be a positive operator, and let $\rho$ 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 $\rho$ as

$$\text{Tr}[\rho A] := \sum_{i: p_i > 0} p_i \|A^{1/2} |e_i\rangle\|^2 \in \mathbb{R} \cup \{+\infty\}, \quad (1)$$

with the convention that $\text{Tr}[\rho A] = +\infty$ if there exists an index $i$ such that $p_i > 0$ and $|e_i\rangle \not\in \text{Dom}(A^{1/2})$. 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 [33, Example 7.1]. We will say that $X$ has finite expected value on $\rho$ if $|e_i\rangle \in \text{Dom}(X_{+}^{1/2}) \cap \text{Dom}(X_{-}^{1/2})$ for all $i$ such that $p_i > 0$, and moreover the two series $\sum_i p_i \|X_{+}^{1/2} |e_i\rangle\|^2$ both converge. In this case, we call

$$\text{Tr}[\rho X] := \sum_{i: p_i > 0} p_i \|X_{+}^{1/2} |e_i\rangle\|^2 - \sum_{i: p_i > 0} p_i \|X_{-}^{1/2} |e_i\rangle\|^2 \quad (2)$$

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

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

For an $n$-linear tensor $A: \times_{i=1}^n \mathbb{C}^m \to \mathbb{C}^k$, we write $A(x^{\times n}) := A(x, ..., x)$ 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 $\dagger$ the standard transposition of vectors, and with $\dagger$ the combination of the two.

For partial derivatives with respect to complex variables $z, z^*$ we write $\partial_z$ and $\partial_{z^*}$. Consider an $m$-dimensional multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ with $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_m$. Then $\partial^{\alpha}_{z} := \partial_{z_1}^{\alpha_1} \partial_{z_2}^{\alpha_2} \cdots \partial_{z_m}^{\alpha_m}$, and analogously for $z^*$. The total derivatives of order $k$ of a function $f: \mathbb{C}^m \to \mathbb{C}$

\[3\] That is, operators $A \in B(\mathcal{H})$ for which $\| A \|_1 := \text{Tr} |A| < \infty$.\]
we denote by $D^k f := \left( \partial_z^\alpha \partial_{\bar{z}}^\beta f \right)_{|\alpha|+|\beta|=k}$. We then recall the definition of the Fréchet derivative for functions $f : \mathbb{C}^m \to \mathbb{C}$ such that $D^k f : \mathbb{C}^m \to \mathcal{B}(\mathbb{C}^m \times \ldots \times \mathbb{C}^m, \mathbb{C})$ and therefore

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

with $j_\alpha(\ell) := \min \left\{ j \in \{1, \ldots, m\} : \ell \leq \sum_{j'=1}^j \alpha_{j'} \right\}$. Let $C_0(\mathbb{C}^m)$ denote the space of continuous functions $f : \mathbb{C}^m \to \mathbb{C}$ that tend to zero as $|z| \to \infty$. We write $C_\infty^k(\mathbb{C}^m)$ 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 $\bar{\Omega}$, a function $f : \bar{\Omega} \to \mathbb{C}$, and a non-negative integer $k \in \mathbb{N}_0$, we denote by $C^k(\bar{\Omega})$ the space of functions for which the norm

$$\|f\|_{C^k(\bar{\Omega})} := \max_{|\alpha|+|\beta| \leq k} \sup_{z \in \Omega} \left| \partial_z^\alpha \partial_{\bar{z}}^\beta f(z) \right|$$

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

$$\|f\|_{C^k(\bar{\Omega})} := \|f\|_{C^{[k]}(\bar{\Omega})} + \sup_{z \in \Omega, w \neq z} \left| \frac{f(z) - f(w)}{|z - w|^{k-[k]}} \right|.$$

This extension allows us to consider the normed spaces $C^k(\bar{\Omega})$ for all $k \geq 0$. Typically, we will deal with the case where $\Omega$ is bounded, so that $C^k(\bar{\Omega})$ 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 $\|f\|^2_{L^2(\Omega)} := \int_{\Omega} |f(z)|^2 \, d^m z$ is finite.

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

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

where

$$\|X\|_{\mathcal{W}^{k,1}(H_m)} := \left\| \left( \sum_j a_j a_j^\dagger \right)^{k/4} \left( X \sum_j a_j a_j^\dagger \right)^{k/4} \right\|_1.$$ 

This yields for density operators $\rho$

$$\text{Tr} \left[ \rho \left( \sum_j a_j a_j^\dagger \right)^{k/2} \right] = \|\rho\|_{\mathcal{W}^{k,1}(H_m)}.$$

### B. Definitions

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 $H_m := L^2(\mathbb{R}^m)$ of square-integrable functions on $\mathbb{R}^m$. Let $x_j, p_j$ be the canonical position and momentum operators on the $j$th 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

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

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

$$\text{Tr}[\rho a_j] = \frac{1}{\sqrt{2}} (\text{Tr}[\rho x_j] + i \text{Tr}[\rho p_j]) = 0 \quad \forall j = 1, \ldots, m,$$  

i.e. if all expected values of the canonical operators on $\rho$, 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 $|n\rangle := (n_1! \ldots n_m!)^{-1/2} \otimes_{j=1}^m (a_j^\dagger)^{n_j} |0\rangle$, 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 $\rho$ can be defined as

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

which is well defined although possibly infinite. The relative entropy between two states $\rho$ and $\sigma$ is usually written as follows

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

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

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

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

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 $E > \inf_{\psi \in \mathcal{H}} \langle \psi | H | \psi \rangle \geq 0$, 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

$$C_H (\mathcal{N}, E) = \lim_{n \to \infty} \frac{1}{n} \chi_H (n) \left( \mathcal{N}^\otimes n, nE \right),$$

$$\chi_H (\mathcal{N}, E) := \max_{\sum_i p_i \rho_i, \sum_i p_i \rho_i^\dagger} \left\{ S \left( \sum_i p_i \rho_i \right) - \sum_i p_i S (\rho_i) \right\},$$

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 $j$th tensor factor, and tensor products with the identity operator are omitted for

\footnote{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 $\rho$. Since all terms of this sum are non-negative, the sum itself can be assigned a well-defined value, possibly $+\infty$.}

\footnote{To define it one considers the infinite sum $D(\rho\|\sigma) := \sum_{i,j} |\langle e_i | f_j \rangle|^2 (p_i \log p_i - p_i \log q_j + q_j - p_i)$, 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 $\rho$ and $\sigma$, respectively. As detailed in \[35\], the convexity of $x \mapsto x \log x$ implies that all terms of this sum are non-negative, which makes the expression well defined.}
notational simplicity. With the same notation, one can also define the energy-constrained quantum capacity of \(\mathcal{N}\), given by \(\mathcal{Q}_{H}(\mathcal{N}, E) := \lim_{n \to \infty} \frac{1}{n} Q^{(1)}_{H(n)}(\mathcal{N}^\otimes n, nE)\),

\[
Q^{(1)}_{H} (\mathcal{N}, E) := \sup_{|\Psi\rangle \in \mathcal{H}^\otimes \hat{R}} \left\{ S (\mathcal{N} (\text{Tr}_{\hat{R}} |\Psi\rangle \langle \Psi|)) - S ((\mathcal{N} \otimes I) (|\Psi\rangle \langle \Psi|)) \right\},
\]

(12)

where \(\text{Tr}_{\hat{R}}\) is the partial trace over the entirely arbitrary ancillary Hilbert space \(\hat{R}\). 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 \(H := \sum_j a_j^\dagger a_j\) of \(m\) modes. In this case, we will omit the subscripts and simply write the energy-constrained capacities as \(\mathcal{Q}(\mathcal{N}, E)\) and \(\mathcal{Q}(\mathcal{N}, E)\).

2. Phase space formalism

We define the displacement operator \(\mathcal{D}(z)\) associated with a complex vector \(z \in \mathbb{C}^m\) as

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

(13)

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

\[
\mathcal{D}(z) \mathcal{D}(w) = \mathcal{D}(z + w) e^{\frac{1}{2} (z^{t} w^{*} - z^{*} w)},
\]

(14)

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

Let \(H_{\text{quad}} = \sum_{j,k} (X_{jk} a_j^\dagger a_k + Y_{jk} a_j a_k + Y_{jk}^* a_j^\dagger a_k^\dagger)\), where \(X = X^\dagger\) is an \(m \times m\) Hermitian matrix, and \(Y = Y^\dagger\) is an \(m \times m\) complex symmetric matrix. The unitaries \(e^{-iH_{\text{quad}}}\) generated by such Hamiltonians, and products thereof, 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 \(z_I := \Im z\). 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 \(H_{\text{quad}}\) 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

\[
\chi_T(z) := \text{Tr} \left[ T \mathcal{D}(z) \right].
\]

(15)

Conversely, the operator \(T\) can be reconstructed from \(\chi_T\) via the weakly defined identity

\[
T = \int \frac{d^{2m}z}{\pi^m} \chi_T(z) \mathcal{D}(-z).
\]

(16)

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 \(\chi_T\) of a trace class operator \(T\) is

\[6\] 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 [38, p.37], combined with [38, Propositions 2.12, 2.18, and 2.19] and with the observation that the exponential Lie map of the unitary group is surjective.
bounded and uniformly continuous [40, § 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:

$$\chi_{\psi_f}(z) = e^{-iz^T z R} \int d^m x \ f^*(x) \ f(x - \sqrt{2z} R) \ e^{\sqrt{2}iz^T x},$$

(17)

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

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

(18)

$$= \frac{2^m}{\pi^m} \text{Tr} \left[ \mathcal{D}(-z) T \mathcal{D}(z) (-1)^m a_j^a_j \right].$$

(19)

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 (19) it is not difficult to see that $|W_T(z)| \leq \frac{2^m}{\pi^m} \|T\|_1$, where $\|T\|_1 = \text{Tr}[T]$ reduces to 1 when $T$ is a density operator. By taking the Fourier transform of (17), one can show that

$$W_{\psi_f}(z) = \frac{2}{\pi} \int d^m x \ f^*(x + \sqrt{2z} R) \ f(-x + \sqrt{2z} R) \ e^{2\sqrt{2}iz^T x}.$$  

(20)

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

$$\int d^{2m} z \ |z|^2 W_\rho(z) = \text{Tr} \left[ \rho \sum_j \left( a_j^a_j + I/2 \right) \right]$$

(21)

The displacement operator $\mathcal{D}(z)$ induce a translation of the Wigner function as follows

$$\chi_{\mathcal{D}(z) \rho \mathcal{D}(z)^\dagger}(u) = e^{z^T u - z^T u^*} \chi_\rho(u), \quad W_{\mathcal{D}(z) \rho \mathcal{D}(z)^\dagger}(u) = W_\rho(u - z).$$

(22)

The map $T \mapsto \chi_T$, defined for trace class operators $T$ in (15), 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

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

(23)

and therefore

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

(24)

Henceforth, we refer to (24) 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

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

(25)
The thermal state is the maximiser of the entropy among all states with a fixed maximum average energy:

$$\max \{ S(\rho) : \rho \in \mathcal{D}(\mathcal{H}_1), \quad \text{Tr}[\rho a^\dagger a] \leq N \} = S(\tau_N) = g(N)$$

(26)

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

$$g(x) := (x + 1) \log(x + 1) - x \log x.$$  

(27)

The characteristic function and Wigner function of the thermal state evaluate to [41, Eq. (4.4.21) and (4.5.31)]

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

(28)

respectively, so that $\tau_N$ is easily seen to be a centred Gaussian state.

C. Moments

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

$$M_k(\rho) := \text{Tr} \left[ \rho \left( \sum_j a_j^\dagger a_j^\dagger \right)^{k/4} \right] = \left\| \left( \sum_j a_j^\dagger a_j^\dagger \right)^{k/4} \rho \left( \sum_j a_j^\dagger a_j^\dagger \right)^{k/4} \right\|_1 < \infty.$$  

(29)

**Remark.** More explicitly, this requires the operator $\left( \sum_j a_j^\dagger a_j^\dagger \right)^{k/4} \rho \left( \sum_j a_j^\dagger a_j^\dagger \right)^{k/4}$, which is initially defined only on a dense domain, to be actually bounded, and its extension to the whole Hilbert space $\mathcal{H}_m$ to be of trace class. These conditions are fairly easy to check once the matrix representation of $\rho$ in the Fock basis is given. Namely, (29) is equivalent to

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

(30)

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

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 phase space moments, relying on the regularity of the quantum characteristic function, in the quantum setting as well.

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

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

(31)

for some $\varepsilon > 0$, where $B(0, \varepsilon) := \{ z \in \mathbb{C}^m : \| z \| \leq \varepsilon \}$, and the norm on the space $C^k(B(0, \varepsilon))$ is defined in [4] and [5].
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 Section IV.

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 [42] 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 $\rho_{\lambda z}^\varphi$. One can show that its characteristic function can be evaluated as

$$\chi_{\rho_{\lambda z}^\varphi}(z) = e^{-|z|^2} \sum_{n=2}^{\infty} \frac{\cos(2nz)}{n^2 \log n},$$

which turns out to be continuously differentiable everywhere [42]. However,

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

which implies that $\rho$ 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 $k$th order phase space moment implies the existence of the $k$th order standard moment. Again, this is in total analogy with the classical case [43, Theorem 1.8.16].

D. 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

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

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

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

Accordingly, displacement operators are transformed by

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

The beam splitter unitary can be used to define the following ($\lambda$-dependent) quantum convolution: for two $m$-mode quantum states $\rho, \sigma$ and $\lambda \in [0,1]$, their ($\lambda$-dependent) quantum convolution is given by the state $\rho \boxplus_\lambda \sigma$ which is defined according to [44] as

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

In terms of characteristic functions, this definition corresponds to

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

$\Box$ Tensor products are omitted here.
In terms of characteristic and Wigner functions, we can also write
\[ n \] iteratively defined for a positive integer \( n \), where \( \sigma \) is given by (27).

Iterating (37), using (42), shows that
\[ \text{Tr} \left[ (\sigma \om \sigma) \sum_j a_j^* a_j \right] = \lambda \text{Tr} \left[ \sigma \sum_j a_j^* a_j \right] + (1 - \lambda) \text{Tr} \left[ \sum_j a_j^* a_j \right]. \] (38)

For all \( m \)-mode quantum states \( \sigma \) and all \( \lambda \in [0,1] \), we can use the corresponding convolution to define a quantum channel \( \mathcal{N}_{\sigma,\lambda} : \mathcal{T}_1(\mathcal{H}_1) \to \mathcal{T}_1(\mathcal{H}_1) \), whose action is given by
\[ \mathcal{N}_{\sigma,\lambda}(\rho) := \rho \om \lambda \sigma. \] (39)

When \( \sigma = \tau_N \) is a thermal state, the channel \( \mathcal{N}_{\tau_N,\lambda} := \mathcal{A}_N^\lambda \) is called a thermal attenuator channel. Its action, obtained by combining (36) and (28), is given by
\[ \mathcal{A}_N^\lambda : \chi_{\rho}(z) \rightarrow \chi_{\mathcal{A}_N^\lambda(\rho)}(z) := \chi_{\rho} \left( \sqrt{\lambda} z \right) e^{-\frac{1-\lambda}{2(N+1)} |z|^2}. \] (40)

The energy-constrained classical capacity of the thermal attenuator, defined in (11), can be shown to be given by
\[ C(\mathcal{A}_N^\lambda, E) = g(\lambda E + (1 - \lambda) N) - g((1 - \lambda N), \] (41)
where \( g \) is given by (27).

In what follows, we will be interested in the symmetric quantum convolutions \( \rho_1 \om \ldots \om \rho_n \), iteratively defined for a positive integer \( n \) and states \( \rho_1, \ldots, \rho_n \), by the relations \( \rho \om \sigma := \rho \om_{1/2} \sigma \) and
\[ \rho_1 \om \ldots \om \rho_n := (\rho_1 \om \ldots \om \rho_{n-1}) \om_{1/n} \rho_n. \] (42)

We will also use the shorthand
\[ \rho^{\otimes n} := \underbrace{\rho \om \ldots \om \rho}_{n \text{ times}}. \] (43)

In terms of characteristic and Wigner functions, we can also write
\[ \chi_{\rho_1 \om \ldots \om \rho_n}(z) = \chi_{\rho_1} \left( \frac{z}{\sqrt{n}} \right) \ldots \chi_{\rho_n} \left( \frac{z}{\sqrt{n}} \right), \] (44)
\[ W_{\rho_1 \om \ldots \om \rho_n}(z) = n^m (W_{\rho_1} \ast \ldots \ast W_{\rho_n}) \left( \sqrt{n} z \right). \] (45)

Here, \( \ast \) denotes convolution, which is defined for \( n \) functions \( f_1, \ldots, f_n : \mathbb{C}^m \to \mathbb{R} \) by
\[ (f_1 \ast \ldots \ast f_n)(x) := \int d^{mn} y_1 \ldots d^{mn} y_{n-1} f_1(y_1) \ldots f_{n-1}(y_{n-1}) f_n(x - y_1 - \ldots - y_{n-1}). \] (46)

Equation (44) 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 (44), on which we will focus most of our efforts, reads
\[ \chi_{\rho^{\otimes n}}(z) = \chi_{\rho} \left( \frac{z}{\sqrt{n}} \right)^n. \] (47)

Iterating (37), using (42), shows that
\[ V(\rho^{\otimes n}) V^\dagger = (V \rho V^\dagger)^{\otimes n} \] (48)
holds for all symplectic unitaries \( V \).
III. 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 \( (\rho^{\mathbb{E}_n})_{n \in \mathbb{N}} \) converges weakly to the Gaussian state \( \rho_G \) of same first and second moments as \( \rho \):

\[
\lim_{n \to \infty} \text{Tr} \left[ \rho^{\mathbb{E}_n} X \right] = \text{Tr} \left[ \rho_G X \right], \quad \forall \, X \in \mathcal{B}(\mathcal{H}),
\]

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

**Remark.** The state \( \rho_G \) is commonly called the *Gaussification* of \( \rho \).

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 result holds [24, Lemma 4]:

**Lemma 4.** Let \( (\rho_n)_{n \in \mathbb{N}} \) be a sequence of density operators on \( \mathcal{H}_m \). The following are equivalent:

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

The above lemma together with Theorem 3 allow us to conclude the following seemingly stronger convergence:

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

\[
\lim_{n \to \infty} \left\| \rho^{\mathbb{E}_n} - \rho_G \right\|_1 = 0.
\]

IV. MAIN RESULTS

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

- Firstly, in the case in which the state \( \rho \) satisfies the conditions of the Cushen–Husdon theorem, we provide quantitative bounds on the rate at which the sequence of states \( (\rho^{\mathbb{E}_n})_{n \in \mathbb{N}} \) converges to \( \rho_G \), under the assumption of finiteness of certain phase space moments of \( \rho \). 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 (Section IV A).

- Secondly, we provide an example to show that the assumption that the second moments be finite in the Cushen–Husdon theorem cannot be weakened (Section IV B).
• Thirdly, we extend our results to the non-i.i.d. setting, i.e. we consider a scaling in the quantum convolution different from (42). This allows us to analyse the propagation of states through cascades of beam splitters with varying transmissivities (Section IV C).

• Finally, we provide a precise asymptotic analysis of the behaviour of quantum characteristic functions at zero and at infinity (Section IV D).

A. 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 $O(n^{-1/2})$ in the quantum central limit theorem under a fourth order moment condition. The rate of convergence is boosted to $O(n^{-1})$ if the third derivative of the characteristic function at zero vanishes:

**Theorem 6** (Quantum Berry–Esseen theorem, High regularity). Let $\rho$ 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

$$\|\rho^{\boxplus n} - \rho_G\|_2 = O(M'_4 (n^{-1/2})) \tag{51}$$

Here $M'_4$ is the moment defined in (31). Moreover, if $D^3 \chi_\rho(0) = 0$ then the convergence is at least with rate $O(M'_4 (n^{-1}))$.

The proof of Theorem 6 is provided in Section VI. In the next Theorem, we weaken the assumption on the moments of the state $\rho$, which leads to a slower rate of convergence.

**Theorem 7** (Quantum Berry–Esseen theorem, Low regularity). Let $\rho$ 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

$$\|\rho^{\boxplus n} - \rho_G\|_2 = O(M'_{4+\alpha} (n^{-\alpha/2})) \tag{52}$$

Here $M'_{4+\alpha}$ is the moment defined in (31).

The proof of Theorem 7 is provided in Section VI. The variable $\alpha$ allows us to obtain a convergence rate under the assumption of finiteness of 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 $\rho$ has finite third-order phase space moments. Then,

$$\|\rho^{\boxplus n} - \rho_G\|_1 = O(M'_3 (n^{-\frac{1}{2m+1}})), \quad D(\rho^{\boxplus n} \| \rho_G) = O(M'_3 (n^{-\frac{1}{2m+1}})) \tag{52}$$

where $M'_3$ is defined in (31). The above rates are replaced by $O(M'_{3+\alpha} (n^{-\alpha/(2m+2)})$ when $\rho$ only satisfies the conditions of Theorem 7.

The proof of this Corollary is given in Section VI.
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 $M_k'\theta$ given by (31), 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 (29). 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

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

(53)

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 [45] for more details): given two Banach spaces $(X_0, \|\cdot\|_{X_0})$ and $(X_1, \|\cdot\|_{X_1})$, and a parameter $0 \leq \theta \leq 1$, define the $K$-function as follows:

$$K(t, X) := \inf_{X = X_0 + X_1} \left(\|X_0\|_{X_0} + t\|X_1\|_{X_1}\right).$$

(54)

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

$$(X_0, X_1)_{\theta} := \{X \in X_0; \Phi_{\theta}(K(t, X)) < \infty\},$$

with norm $\|X\|_{(X_0, X_1)_{\theta}} = \Phi_{\theta}(K(t, X)).$

Now, given two couples of Banach spaces $X_0, X_1$ and $Y_0, Y_1$, and a map $\Psi : X_0 + X_1 \to Y_0 + Y_1$ such that $\Psi : X_0 \to Y_0$ and $\Psi : X_1 \to Y_1$ are bounded, the map $\Psi : (X_0, X_1)_{\theta} \to (Y_0, Y_1)_{\theta}$ is bounded and:

$$\|\Psi : (X_0, X_1)_{\theta} \to (Y_0, Y_1)_{\theta}\| \leq \|\Psi : X_0 \to Y_0\|^{1-\theta}\|\Psi : X_1 \to Y_1\|^\theta.$$

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 $\mathcal{W}^{k_0,1}(\mathcal{H}_m)$ and $\mathcal{W}^{k_1,1}(\mathcal{H}_m)$ form a compatible couple such that for any $m$-mode quantum state $\rho$ and $\theta \in (0, 1)$ the real interpolation norm satisfies

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

The proof of Proposition 10 is stated in Appendix B.
### B. 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 \(|\psi_f\rangle\) with wave function

\[
f(x) := \frac{1}{\sqrt{2}} \frac{1}{(1 + x^2)^{3/4}}.
\]

Then: (a) \(M_{2(1-\delta)}(\psi_f) = \langle \psi_f | (aa^\dagger)^{1-\delta} | \psi_f \rangle < \infty\) for all \(\delta > 0\); yet (b) the sequence \((|\psi_f\rangle \langle \psi_f|)^{\otimes n}\) 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 Section VII.

We now come to the second question (ii) regarding tightness of the estimates in Theorems 6 and 7 and Corollary 8. In Section VII 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 Gaussianisation \(\psi_G = \tau_3/2\). Our findings indicate that \(\|\psi^{\otimes n} - \psi_G\|_2 \sim cn^{-1/2}\), 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 Figure 4). Hence, the \(O(n^{-1/2})\) 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 \(\|\psi^{\otimes n} - \psi_G\|_{\alpha}\) for the simple case of a single-photon state \(\psi = |1\rangle \langle 1|\) and for \(\alpha = 1, 2\), and find numerical evidence that again \(\|\psi^{\otimes n} - \psi_G\|_{\alpha} \sim cn^{-1}\) for some absolute constant \(c\) (Example 6 and Figure 4). This shows that the \(O(n^{-1})\) convergence rate stated in Theorem 6 under the assumption that \(D^3\chi_{\rho}(0) = 0\), is also attained.

### C. 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 \(\lambda\) as a cascade of \(n\) beam splitters, in each of which the signal state \(\omega\) is mixed via an elementary beam splitter of transmissivity \(\lambda^{1/n}\) with a fixed state \(\rho\), modelling the environmental noise (Figure 2). Each step corresponds to the action of the channel \(\mathcal{N}_{\rho,\lambda^{1/n}}^n : \omega \mapsto \omega \boxplus \lambda^{1/n} \rho\) (cf. the definition (39)), so that the whole cascade can be represented by the \(n\)-fold composition \(\mathcal{N}_{\rho,\lambda^{1/n}}^n\). We are interested in the asymptotic expression of the output state \(\mathcal{N}_{\rho,\lambda^{1/n}}^n(\omega)\) as the number \(n\) tends to infinity, as a function of the input state \(\omega\). In other words, we want to study the asymptotic channel \(\lim_{n \to \infty} \mathcal{N}_{\rho,\lambda^{1/n}}^n\).

Let us note in passing that the above cascade architecture, in spite of some apparent resemblance, is different from that depicted in Figure 11. 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 \(\rho^{\otimes n}\) at the output.
Theorem 12 (Approximation of thermal attenuators channels by cascades of beam splitters). Let \( \rho \) be a centred \( m \)-mode quantum state with finite third-order phase space moments \( M_3 \), cf. \( [31] \), and denote by \( \rho_G \) its Gaussification. Then,

\[
\| N^\rho_{\lambda, \lambda^\mathrm{in}} - N^\rho_{\lambda^G, \lambda} \|_\diamond = \Theta_{M_3} \left( n^{-\frac{1}{2(m+1)}} \right),
\]

where \( \| . \|_\diamond \) stands for the diamond norm \( [10] \).

One can further make use of the recently derived continuity bounds under input energy constraints \( [25–28] \) in order to find bounds on capacities of the cascade channel \( N^\rho_{\lambda, \lambda^\mathrm{in}} \) in the physically relevant case where the Gaussification \( \rho_G \) of \( \rho \) is a thermal state.

Corollary 13. Consider a single-mode quantum state \( \rho \) with finite third-order phase space moments \( M_3 \) (cf. \( [31] \)) and thermal Gaussification \( \rho_G = \tau_N \) as in \( [25] \). Then, for \( \lambda \in [0, 1] \), mean photon number \( N := \text{Tr} \left[ \rho a^\dagger a \right] < \infty \), and some input energy \( E > 0 \), the energy-constrained classical and quantum capacity of the cascade channel \( N^\rho_{\lambda, \lambda^\mathrm{in}} \) relative to the canonical Hamiltonian \( a^\dagger a \) satisfy

\[
\begin{align*}
\left| \mathcal{C}(N^\rho_{\lambda, \lambda^\mathrm{in}}, E) - g(\lambda E + (1 - \lambda)N) - g((1 - \lambda)N) \right| & \leq \Delta_c(n; N, M_3', \lambda, E) \quad \text{(56)} \\
\text{and} \\
\max \left\{ 0, \log \frac{\lambda}{1 - \lambda} - g(N) \right\} - \Delta_q(n; N, M_3', \lambda, E) \leq \mathcal{Q}(N^\rho_{\lambda, \lambda^\mathrm{in}}, E) \leq F(\lambda, N) + \Delta_q(n; N, M_3', \lambda, E),
\end{align*}
\]

(57)

where \( g(x) := (x + 1) \log(x + 1) - x \log x \) (as in \( [27] \)), the function \( F \) is defined as

\[
F(\lambda, N) = \begin{cases} 
\min \left\{ \log \frac{\lambda - N(1 - \lambda)}{(1 + N)(1 - \lambda)}, \log \left[ \lambda^N(1 - \lambda) \right] - g(N) \right\} & \text{if } \lambda \geq \frac{N + 1/2}{N + 1}, \\
0 & \text{otherwise},
\end{cases}
\]

(58)

This amounts to assuming that \( \rho \) can be brought to its so-called Williamson form (see \( [67] \) of Section \( [VI] \) by a passive symplectic unitary only.
and the remainder terms are such that
\[ \Delta_c(n; N, M'_3, \lambda, E) \leq C(M'_3) n^{-1/4} \log n, \]
\[ \Delta_q(n; N, M'_3, \lambda, E) \leq C(M'_3) n^{-1/8} \log n. \]

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

The proofs of Theorem 12 and Corollary 13 are postponed to Section VII.

D. 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 Section V B.

**Proposition 14.** The quantum characteristic function of an \( m \)-mode quantum state \( \rho \) is a continuous function that is arbitrarily small in absolute value outside of a sufficiently large compact set, i.e. \( \chi_\rho \) belongs to the Banach space \( C_0(\mathbb{C}^m) \) of asymptotically vanishing functions. In particular, for any \( \varepsilon > 0 \) we have
\[ \max_{z \in \mathbb{C}^m \setminus B(0, \varepsilon)} |\chi_\rho(z)| < 1, \]
where \( B(0, \varepsilon) \) denotes a ball of radius \( \varepsilon \) centred at the origin.

The proof of Proposition 14 is given in Section V B.

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 \( \rho \) be an \( m \)-mode quantum state with finite average energy \( E := \text{Tr} \left[ \rho \sum_j (a_j^\dagger a_j + 1/2) \right] \), where we have explicitly accounted for the non-zero energy of the ground state. Then, for all \( z \in \mathbb{C}^m \) and all \( \delta \in [0, 1] \) it holds that
\[ |\chi_\rho(z)| \leq 1 - \frac{(1 - \delta)^3 \delta^{2m-1} ((2m + 1)!)^2}{6 \cdot 24^m \cdot E^{2m-1}} \min \left\{ \|z\|^2, \frac{\pi^2 \delta}{4E} \right\}. \]

The proof of Proposition 15 is given in Section V B.

V. 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 Section V B. 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 a density operator with itself naturally induces a classical random variable (Section V A).

- **Decay bounds:** We derive new decay estimates and asymptotic properties of the quantum characteristic function at infinity (Section V B).
A. Quantum–classical correspondence

In this section we show that the quantum convolution $\rho \boxplus \sigma$ of any two states $\rho$ and $\sigma$ has a non-negative Wigner function. While the mathematics behind this is known (see e.g. [16, Proposition (1.99)]), or [2, Proposition 5]), we believe that its physical implications have not been appreciated to the extent they deserve.

**Lemma 16.** Let $\rho$ and $\sigma$ be arbitrary $m$-mode quantum states on an $m$-mode system. Then the Wigner function of their convolution $\rho \boxplus \sigma$ defined by (35), with $\lambda = 1/2$, is given by

$$W_{\rho \boxplus \sigma}(z) = \frac{2^m}{\pi^m} \text{Tr} \left[ \rho \hat{\mathcal{D}}(\sqrt{2}z)J\sigma J\mathcal{D}(\sqrt{2}z)^\dagger \right],$$

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

$$W_{\rho \boxplus \sigma}(z) \geq 0 \quad \forall \ z \in \mathbb{C}^m. \quad (63)$$

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

$$W_{J\rho J^\dagger}(z) = W_{\rho}(-z) \quad (64)$$

for all $m$-mode quantum states $\rho$ and all $z \in \mathbb{C}^m$. This follows from the easily verified fact that $Ja_j J = -a_j$ for all $j$, which also implies that $J\mathcal{D}(z)J = \mathcal{D}(-z)$. In fact, using (19) we find that

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

We now compute

$$W_{\rho \boxplus \sigma}(z) = 2^m (W_{\rho} \star W_{\sigma}) \left( \sqrt{2}z \right)$$

$$= \frac{2^m}{\pi^m} \int d^{2m} u W_{\rho}(u) W_{\sigma} \left( \sqrt{2}z - u \right)$$

$$= \frac{2^m}{\pi^m} \int d^{2m} u W_{\rho}(u) W_{J\sigma J^\dagger} \left( u - \sqrt{2}z \right)$$

$$= \frac{2^m}{\pi^m} \int d^{2m} u W_{\rho}(u) W_{J\mathcal{D}(\sqrt{2}z)J\sigma J\mathcal{D}(\sqrt{2}z)^\dagger} \left( u \right)$$

$$= \frac{2^m}{\pi^m} \text{Tr} \left[ \rho \mathcal{D}(\sqrt{2}z)J\sigma J\mathcal{D}(\sqrt{2}z)^\dagger \right].$$

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

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 $\rho$, and notice that $\rho^{\boxplus 2n} = (\rho \boxplus \rho)^{\boxplus n}$. Consider the probability density function $f_X := W_{\rho \boxplus \rho} \geq 0$, where positivity holds by (63). 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 $\rho$. 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 $f_X$

$$W_{\rho}^{\boxplus 2n}(u) = W_{(\rho \boxplus \rho)^{\boxplus n}}(u)$$

$$= \frac{1}{n} W_{\rho}^{n^2} \left( \sqrt{nu} \right)$$

$$= \frac{1}{n} f_X^n \left( \sqrt{nu} \right)$$

$$\geq f_{(X_1 + \ldots + X_n)/\sqrt{n}}(u)$$
where 1 follows from \((45)\) and 2 follows from the change of variables \(u \mapsto \sqrt{n}u\). This implies by applying both the classical and quantum Plancherel theorem that
\[
\|\rho^{\otimes 2n} - \rho_G\|_2 = \pi^{-m}\|X^{\otimes 2n} - X^G\|_{L^2(\mathbb{R}^{2m})} \\
= \pi^m\|W^{\otimes 2n} - W^G\|_{L^2(\mathbb{R}^{2m})} \\
= \pi^m\|f(X_1 + \ldots + X_n)/\sqrt{n} - f_G\|_{L^2(\mathbb{R}^{2m})}
\]
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 \(p_X^{\otimes n}\) 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.

### B. 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 \(\chi_X^{cl}\) 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; see also [47, Section 3.5]. Examples include the Dirac, Bernoulli, geometric and Poisson distributions.

Knowing that \(\|\chi_X^{cl}(t)\| < 1\) for all values \(t \neq 0\) however does not imply that \(\limsup_{t \to \infty} |\chi_X^{cl}(t)| < 1\). 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** [47, 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
\[
P(X = q_n) = P(X = -q_n) = 2^{-(n+1)}.
\]
One easily shows that the characteristic function of the random variable \(X\) is given by \(\chi_X^{cl}(t) = \sum_{i=1}^{\infty} \frac{\cos(tq_i)}{2^i}\). One has \(\limsup_{t \to \infty} |\chi_X^{cl}(t)| = 1\).

We now show the surprising fact that quantum characteristic functions do not exhibit this somewhat pathological behaviour. Instead, for any quantum state \(\rho\) it holds that \(\limsup_{|z| \to \infty} |\chi_\rho(z)| = 0\), 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 \(\lim_{|z| \to \infty} \langle \psi_f | \chi_\rho(z) | \psi_f \rangle = 0\) 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 \(\chi_\psi f\) belongs to the Banach space \(C_0(\mathbb{C}^m)\), where the norm on \(C_0(\mathbb{C}^m)\) is the supremum norm.

---

9 These are discrete random variables with probability distributions supported on a lattice.
We consider smooth compactly supported functions \( f \) first. For such functions, the claim follows by combining (i) Eq. (17); (ii) the fact that \( f \) is normalised, i.e. \( \{d^m|f(x)|^2 = 1 \); 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 \( f_n \in C_0^\infty(\mathbb{R}^m) \) converging to \( f \in L^2(\mathbb{R}^m) \), so that

\[
\sup_{z \in \mathbb{C}^m} \left| \left( \chi_{\psi_{f_n}} - \chi_{\psi_f} \right)(z) \right| \\
= \sup_{z \in \mathbb{C}^m} \left| \langle \psi_{f_n}, \mathcal{D}(z) \psi_{f_n} \rangle - \langle \psi_f, \mathcal{D}(z) \psi_f \rangle \right| \\
\leq \sup_{z \in \mathbb{C}^m} \left\{ \left| \langle \psi_{f_n}, \mathcal{D}(z) \psi_{f_n} \rangle - \langle \psi_f, \mathcal{D}(z) \psi_f \rangle \right| + \left| \langle \psi_{f_n}, \mathcal{D}(z) \psi_f \rangle - \langle \psi_f, \mathcal{D}(z) \psi_f \rangle \right| \right\} \\
\leq 2 \| \psi_{f_n} - \psi_f \|_{L^2(\mathbb{R}^m)} \to_0 0.
\]

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 (60) it suffices to show that for any \( \zeta \in \mathbb{C}^m \) it follows by combining (i) Eq. (17); (ii) the fact that \( \chi_{\psi_f}(\zeta) \leq \eta(\rho, \varepsilon) \) (cf. (60)). However, the problem of characterising the quantity \( \eta(\rho, \varepsilon) \) in terms of some physically meaningful property of the state \( \rho \) 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 \( \rho_G \) of mean photon number \( \text{Tr} \left[ \rho_G \sum_j a_j^\dagger a_j \right] \leq \frac{e^2}{8\ln \frac{1}{1-\delta}} - \frac{1}{4} \) such that \( \left| \chi_{\rho_G}(z) \right| \geq 1 - \delta \).

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 defined by the characteristic function

\[
\chi_{\rho_G}(z) := e^{-\frac{1}{2}(2(Rez_1)^2 - \frac{1}{2}\eta(3z_1)^2 - \frac{1}{2}\sum_{j>1} |z_j|^2),}
\]

where we set \( \eta := \min \left\{ \frac{2}{T}, \ln \frac{1}{1-\delta}, 1 \right\} > 0 \). The mean photon number of \( \rho_G \) is well known to be given by \( \text{Tr} \left[ \rho_G \sum_j a_j^\dagger a_j \right] = \frac{1}{4} \left( \eta + \frac{1}{\eta} \right) - \frac{1}{2} \leq \frac{1}{4\eta} - \frac{1}{4} \), 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 \( \left| \chi_{\rho}(z) \right| \) at any designated point \( z \in \mathbb{C}^m \) in terms of the energy of the state \( \rho \). 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 \( \rho \), we can take \( \eta(\rho, \varepsilon) = 1 - \frac{\varepsilon}{2} \min \left\{ \|z\|^2, \rac{C}{\varepsilon} \right\} \), where \( E \) is the energy of \( \rho \), and \( c, C \) are universal constants.
Proof of Proposition 15. We write the following chain of inequalities.

\[ |\chi_{\rho}(z)|^2 \leq 2 \left| \chi_{X(\rho \boxplus \rho)}^{cl} \left( \sqrt{2}z \right) \right|^{1/2} \]

\[ = \left( 1 - \frac{(1 - \delta)^2 \delta^{2m-1} \left( (2m + 1)!! \right)^2}{3 \cdot 2^{4m} \cdot E^{2m-1}} \min \left\{ \|z\|^2, \frac{\pi \delta}{4E} \right\} \right)^{1/2} \]

\[ \leq 1 - \frac{(1 - \delta)^2 \delta^{2m-1} \left( (2m + 1)!! \right)^2}{6 \cdot 2^{4m} \cdot E^{2m-1}} \min \left\{ \|z\|^2, \frac{\pi \delta}{4E} \right\} \]

Here, 1 is an application of the quantum convolution rule (cf. the \(n = 2\) case of (47)). In 2 we introduced the classical random variable \(X(\rho \boxplus \rho) \in \mathbb{R} C^m\) 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 [43, Corollary 2.7.2]: we set \(\alpha := \sup_{z \in \mathbb{C}^m} W_{\rho \boxplus \rho}(z) \leq \frac{2m}{\pi m}\), with the latter estimate coming from (19), and \(\alpha = 2\), so that

\[ \gamma_\alpha = \gamma_2 = \int d^m z \|z\|^2 W_{\rho \boxplus \rho}(z) \left[ (\rho \boxplus \rho) \sum_j (a_j^+ a_j + 1/2) \right] \left[ \rho \sum_j (a_j^+ a_j + 1/2) \right] =: E; \]

also, we substituted \(m \rightarrow 2m\), because our phase space \(C^m\) 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 \(\sqrt{1 - x} \leq 1 - \frac{x}{2}\).

Remark. In [43, Section 2.7], several other estimates for \(\left| \chi_{X}^{cl}(t) \right|\) 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 [43 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.

VI. 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 Section IV C 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 (Section VI A).

- **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 moments of the state, cf. Lemma 18.
• **Hilbert–Schmidt convergence:** We implement the above ideas to prove Theorems 6 and 7 and Proposition 22 (Section VI B).

• **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 [48], and in relative entropy, using entropic continuity bounds [25] (Section VI C).

• **Convergence rates for cascades of beam splitters:** In the final subsection, we prove the results claimed in Section IV C, namely convergence rates for cascades of beam splitters converging to thermal attenuator channels (Section VIII).

### A. Preliminary steps

#### 1. Williamson form

Let \( \rho \) 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

\[
\rho' := V^\dagger \rho V \tag{67}
\]

satisfies

\[
\chi_{\rho'}(z) = 1 - \frac{1}{2} \sum_j \nu_j |z_j|^2 + o(\|z\|^2) \quad (z \to 0). \tag{68}
\]

With a slight abuse of terminology, we will call \( \rho' \) the **Williamson form** of \( \rho \) [49]. 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 (67) does not change the first moments of the state, so that if \( \rho \) is centred then \( \rho' \) remains centred. Finally, (iii) the same unitary \( V \) brings not only \( \rho \) but also its Gaussification \( \rho_G \) to their Williamson forms simultaneously, so that

\[
\chi_{\rho_G}(z) = \exp \left[ -\frac{1}{2} \sum_j \nu_j |z_j|^2 \right], \quad W_{\rho_G}(z) = \left( \frac{2}{\pi} \right)^m \exp \left[ -2 \sum_j \nu_j |z_j|^2 \right]. \tag{69}
\]

holds as well. Thanks to the covariance of the quantum convolution with respect to symplectic unitaries [48], we see that

\[
\| \rho^\text{fin} - \rho_G \|_2 = \| V^\dagger (\rho^\text{fin} - \rho_G) V \|_2 = \| (\rho')^\text{fin} - \rho_G' \|_2. \tag{70}
\]

Combining this with the quantum Plancherel identity (24) yields

\[
\| \rho^\text{fin} - \rho_G \|_2^2 = \int \frac{d^{2m} z}{\pi^m} \left| \chi_{(\rho')^\text{fin}}(z) - \chi_{\rho_G'}(z) \right|^2 \tag{71}
\]

In short, when estimating any unitarily invariant distance of \( \rho^\text{fin} \) from its limit \( \rho_G \), 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.
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 characteristic function around zero.

Lemma 17. Let \( \rho \) be an \( m \)-mode quantum state with finite second-order phase space moment. Without loss of generality, we assume that \( \rho \) is centred and in Williamson form, and that its Gaussification \( \rho_c \) has characteristic function as in \([69]\). Then for every \( \varepsilon > 0 \) we have that

\[
\| \rho^{m_{\Omega}} - \rho_c \|_2 \leq \frac{1}{\pi m} \int_{\|z\| \leq \sqrt{\pi} \varepsilon} d^2mz \left| \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right|^2 + O(n^{-\infty})
\]

where \( \| \rho_{m_{\Omega}} - \rho_c \|_2 \leq \sqrt{m(m+1)(m+2)} \| D^3 \chi_\rho(0) \|/\sqrt{n} + O(n^{-\infty}) \)

as \( n \to \infty \). If \( \rho \) has also finite third-order phase space moments, then

\[
\| \rho^{m_{\Omega}} - \rho_c \|_2 \leq \frac{1}{\pi m} \left( \int_{\|z\| \leq \sqrt{\pi} \varepsilon} d^2mz \left| \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \left( 1 + \frac{1}{6 \sqrt{n}} D^3 \chi_\rho(0)(z^x_3) \right) \right|^2 \right)^{1/2},
\]

where \( \| D^3 \chi_\rho(0) \| := \sup_{\|z\| \leq 1} D^3 \chi_\rho(0)(z^x_3) \).

Proof. The first identity (72) follows along the lines of the second one (73) and so we focus on verifying the latter. Using the quantum Plancherel identity (24) and the relation (44), we apply the triangle inequality and split the integration domain into two disjoint sets such that

\[
\pi^m \| \rho^{m_{\Omega}} - \rho_c \|_2 = \left( \int d^2mz \left| \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \right|^2 \right)^{1/2}
\]

\[
\leq \left( \int d^2mz \| \mathbb{I}_{B(0, \sqrt{n} \varepsilon)}(z) \| \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \left( 1 + \frac{1}{6 \sqrt{n}} D^3 \chi_\rho(0)(z^x_3) \right) \right)^{1/2}
\]

\[
+ \left( \int d^2mz \| \mathbb{I}_{\mathbb{R}^2 \setminus B(0, \sqrt{n} \varepsilon)}(z) \| \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \left( 1 + \frac{1}{6 \sqrt{n}} D^3 \chi_\rho(0)(z^x_3) \right) \right)^{1/2}
\]

where for some \( \Omega \subseteq \mathbb{C}^m \) we set \( \mathbb{I}_\Omega(z) = 1 \) if \( z \in \Omega \) and \( \mathbb{I}_\Omega(z) = 0 \) otherwise.

We then find that the last term in (74) can be estimated explicitly using spherical coordinates and the bound \( |D^3 \chi_\rho(0)(z^x_3)| \leq \| D^3 \chi_\rho(0) \| |z^3| \), yielding

\[
\int d^2mz \left| e^{-\frac{1}{2} \sum_j \nu_j |z_j|^2} \left( \frac{1}{6} D^3 \chi_\rho(0)(z^x_3) \right) \right|^2 \leq \int_0^\infty \exp(-r^2) \left( \frac{1}{6} D^3 \chi_\rho(0)(z^x_3) \right)^2 \lesssim \int_0^\infty \exp(-r^2) r^{2m+5} dr \frac{\text{vol}(S^{m-1})}{36} \frac{\| D^3 \chi_\rho(0) \|^2}{2}
\]

\[
= \Gamma(m + 3) \frac{\| D^3 \chi_\rho(0) \|^2}{72}
\]

where we used that \( \int_0^\infty \exp(-r^2) r^{2m+5} dr = \frac{\Gamma(m+3)}{2}, \) and recalled the expression \( \text{vol}(S^{N-1}) = \frac{2\pi^{N/2}}{\Gamma(N/2)} \) for the volume of the \((N-1)\)-sphere.
Furthermore, the second-to-last term in (74) is exponentially small by Proposition 14, since we can write

\[ \int d^{2m} z \mathbb{I}_{\mathbb{R}^{2m} \setminus B(0, \varepsilon \sqrt{n})}(z) \left| \chi_{\rho} \left( \frac{z}{\sqrt{n}} \right) \right|^2 e^{-\frac{1}{2} \sum \nu_j |z_j|^2} \left( 1 + \frac{1}{6 \sqrt{n}} D^3 \chi_{\rho}(0)(z \times 3) \right)^2 \]

\[ \leq 2 \int d^{2m} z \mathbb{I}_{\mathbb{R}^{2m} \setminus B(0, \varepsilon \sqrt{n})}(z) \left| \chi_{\rho} \left( \frac{z}{\sqrt{n}} \right) \right|^2 e^{-\frac{1}{2} \sum \nu_j |z_j|^2} \left( 1 + \frac{|z|^3}{6 \sqrt{n}} \| D^3 \chi_{\rho}(0) \| \right)^2 \]

\[ \leq 2 \left\| \mathbb{I}_{\mathbb{R}^{2m} \setminus B(0, \varepsilon)} \chi_{\rho} \right\|_{L^\infty(\mathbb{R}^{2m})} \int d^{2m} z \mathbb{I}_{\mathbb{R}^{2m} \setminus B(0, \varepsilon \sqrt{n})}(z) \left| \chi_{\rho} \left( \frac{z}{\sqrt{n}} \right) \right|^2 e^{-\frac{1}{2} \sum \nu_j |z_j|^2} \left( 1 + \frac{|z|^3}{6 \sqrt{n}} \| D^3 \chi_{\rho}(0) \| \right)^2 \]

where in 1 we use the fact that the $L^2$ norm of the characteristic function is at most one and switch to spherical coordinates to compute the second integral. In 2 we use Hölder’s inequality to evaluate the second integral.

Since the exponential decay exceeds the polynomial growth, it follows that

\[ \sup_{r \in (\varepsilon \sqrt{n}, \infty)} \exp \left( -\frac{1}{2} r^2 \right) r^{2(m-1)} \left( 1 + \frac{r^3 \| D^3 \chi_{\rho}(0) \|}{6 \sqrt{n}} \right)^2 = O(1) \]

which implies the result. \hfill \Box

In the subsequent lemma we record various important estimates on the behaviour of the quantum characteristic function in a neighbourhood of zero. These bounds will play a key role in the proofs of our quantum Berry–Esseen theorems.

**Lemma 18.** Let $\rho$ be a centred $m$-mode state in Williamson form. The following bounds on the quantum characteristic function hold for a centred state around zero for $z \in B := B(0, \varepsilon)$, with $\varepsilon > 0$ arbitrary, and multi-indices $\alpha \in \mathbb{N}_0^{2m}$

\[ \left| \chi_{\rho} \left( \frac{z}{\sqrt{n}} \right) - 1 \right| \leq \max_{|\alpha| = 2} \max_{y \in B} |D^\alpha \chi_{\rho}(y)| \frac{R^2}{2n} \text{ for } \chi_{\rho} \in C^2(B) \]

\[ \left| \chi_{\rho} \left( \frac{z}{\sqrt{n}} \right) - 1 + \sum \nu_j |z_j|^2 \right| \leq \max_{|\alpha| = 3} \max_{y \in B} |D^\alpha \chi_{\rho}(y)| \frac{R^3}{6n^{3/2}} \text{ for } \chi_{\rho} \in C^3(B), \text{ and } (76) \]

\[ \left| \chi_{\rho} \left( \frac{z}{\sqrt{n}} \right) - 1 + \sum \nu_j |z_j|^2 \right| - \frac{D^3 \chi_{\rho}(0)(z \times 3)}{6n^{3/2}} \leq \max_{|\alpha| = 4} \max_{y \in B} |D^\alpha \chi_{\rho}(y)| \frac{R^4}{24n^2} \text{ for } \chi_{\rho} \in C^4(B). \]

**Proof.** The estimates follow from the multi-dimensional Taylor formula. \hfill \Box
B. Proofs of convergence rates in Hilbert–Schmidt distance

We start with the proof of Theorem 6 assuming fourth order moments.

Proof of Theorem 6. Recall that we can assume that $D^2 \chi_\rho(0) = \text{diag}(\nu_1, ..., \nu_{2m})$ with $\nu_1, ..., \nu_{2m} \geq 1$ in the proof. Let $\mu < 2$ and $R$ small enough such that

$$\max_{|\alpha|=2} \max_{y \in B} |D^\alpha \chi_\rho(y)| \frac{R^2}{2n} \leq \frac{\mu}{2},$$

where $B = B(0, R)$. This implies that for $a(\mu) := \sum_{j=2}^{\infty} j^{-1} (\mu/2)^{j-2}$ and some constant $C_1 > 0$, for any $z \in B(0, \varepsilon \sqrt{n})$,

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

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

$$\leq n \left| - \sum_{j=2}^{\infty} j^{-1} \left( 1 - \chi_\rho \left( \frac{z}{\sqrt{n}} \right) \right)^j \right| + \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - 1 + \frac{\sum_j \nu_j |z_j|^2}{2n} - \frac{1}{6n\sqrt{n}} D^3 \chi_\rho(0)(z \times 3) \right| \quad (77)$$

$$\leq n \left( a(\mu) \left( \max_{|\alpha|=2} \max_{y \in B} |D^\alpha \chi_\rho(y)| \frac{\sum_j \nu_j |z_j|^2}{2n} \right)^2 + \max_{|\alpha|=4} \max_{y \in B} |D^\alpha \chi_\rho(y)| \frac{|z|^4}{24n^2} \right) \leq \frac{C_1 |z|^4}{n}. \quad \text{Here, in 1 we use the Taylor expansion of the logarithm and in 2 we use Lemma 18. Similarly, one shows using Lemma 18 for some constant } C_2 > 0$$

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

$$= n \left| - \sum_{j=2}^{\infty} j^{-1} \left( 1 - \chi_\rho \left( \frac{z}{\sqrt{n}} \right) \right)^j \right| + \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - 1 + \frac{\sum_j \nu_j |z_j|^2}{2n} \right| \quad (78)$$

$$\leq n \left( a(\mu) \left( \max_{|\alpha|=2} \max_{y \in B} |D^\alpha \chi_\rho(y)| \frac{\sum_j \nu_j |z_j|^2}{2n} \right)^2 + \max_{|\alpha|=3} \max_{y \in B} |D^\alpha \chi_\rho(y)| \frac{|z|^3}{6n\sqrt{n}} \right) \leq \frac{C_2 |z|^3}{\sqrt{n}}. \quad \text{Again, in 1 we use the Taylor expansion of the logarithm and in 2 we use Lemma 18. From the two estimates (77 and 78) stated above, we find using } |e^x - 1 - x| \leq |x|^2 e^x \text{ in 1 and } |z| \leq \varepsilon \sqrt{n} \text{ in 2 below that}$$

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

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

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

$$+ \left| \exp \left( \log \left( \chi_\rho \left( \frac{z}{\sqrt{n}} \right)^n \right) + \frac{\sum_j \nu_j |z_j|^2}{2} \right) - \left( 1 + \log \left( \chi_\rho \left( \frac{z}{\sqrt{n}} \right)^n \right) - \frac{\sum_j \nu_j |z_j|^2}{2} \right) \right) \right| \quad (79)$$

$$\leq \exp \left( - \frac{\sum_j \nu_j |z_j|^2}{2} \right) \left( C_1 \frac{|z|^4}{n} + C_2 \frac{|z|^6}{n} e^{C_2 |z|^2} \right) \quad 1 \leq \exp \left( - \frac{\sum_j \nu_j |z_j|^2}{2} \right) \left( C_1 \frac{|z|^4}{n} + C_2 \frac{|z|^6}{n} e^{C_2 |z|^2} \right) \quad 2.$$
Using Lemma 17, the result follows by inserting (79) into (73). In particular, if $D^3 \chi_\rho(0) = 0$ then by (79) the convergence rate is $\mathcal{O}(n^{-1})$.

We recall that any density operator $\rho$ has an expansion into the Fock basis such that

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

(80)

Hence, we find for the characteristic function that

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

(81)

The above expressions allows us now to give the proof of Proposition 20.

**Proof of Proposition 20.** By Theorem 6 it suffices to show that $D^3 \chi_\rho(0) = 0$ under the assumptions of the Proposition. The associated characteristic functions of the basis expansion are explicitly given by [46, Theorem 104.]

$$\chi_{|i\rangle \langle j|}(z) = e^{-|z|^2} H_{ji}(z).$$

(82)

Here, $H_{ji}$ is the complex Hermite polynomial defined as

$$H_{ji}(z) = \begin{cases} \sqrt{\frac{n}{j!}} z^j \Gamma_{i}^{-j} |z|^2, & \text{where } j \geq i \geq 0 \\ (-1)^{j-i} \sqrt{\frac{n}{j!}} \pi^{j-i} L_{j}^{-j} |z|^2, & \text{where } 0 \leq j < i. \end{cases}$$

(83)

and $L_{j}^{i}$ the generalized Laguerre polynomial

$$L_{j}^{i}(z) = \frac{(j + i)!}{(j - m)!(i + m)!} \frac{(z)^m}{m!} .$$

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 $\chi_{|i\rangle \langle j|}(z) = e^{-|z|^2} H_{ji}(z)$ for some suitable function $H_{ji}$, as in (82), that for the different possible third derivatives, we have

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

$$\frac{\partial^3}{\partial \rho^3} \chi_{|i\rangle \langle j|}(0) = -3 \partial_{\rho} H_{ji}(0) + \partial_{\rho}^3 H_{ji}(0),$$

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

and

$$\frac{\partial^3}{\partial \rho \partial z} \chi_{|i\rangle \langle j|}(0) = -\partial_{\rho} H_{ji}(0) + \partial_{\rho}^3 H_{ji}(0).$$

(84)

Therefore, the only possible non-zero contribution to the third derivative of the quantum characteristic function $\chi_\rho$ 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 $z$ appears in (82) with a power at least 4 and thus this term necessarily has to vanish. It suffices therefore to consider the case $|i - j| = 2$. If $H_{ji}$ is only differentiated once,
then it is clear that this derivative has to vanish at zero, since \( z \) appears with a 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. By the chain rule any first order derivative of the term \( L^{j_x-j}_{j_x}(|z|^2) \) vanishes at the origin, since the derivative of the generalized Laguerre polynomials is again a generalized Laguerre polynomial

\[
\frac{d^k}{dx^k} L^{(\alpha)}_n(x) = \begin{cases} 
(-1)^k L^{(\alpha+k)}_{n-k}(x) & \text{if } k \leq n \\
0 & \text{otherwise}
\end{cases}
\]

This concludes the proof.

We continue with the proof of the low-regularity QCLT that assumes finiteness of phase space moments of order up to \( 2 + \alpha \).

**Proof of Theorem 7.** Let \( \mu < 2 \), \( z \in B := B(0, R) \) and \( \alpha \in (0, 1] \), such that

\[
\max_{|\alpha| = 2} \max_{y \in B} |D^\alpha \chi_\rho(y)| \leq \frac{\mu}{2} \quad \text{and} \quad \left| \chi_\rho \left( \frac{z}{\sqrt{n}} \right) - 1 + \sum_j \nu_j |z_j|^2 \right| \leq [D^2 u]_{C^\alpha(B)} \frac{R^{2+\alpha}}{2n^{1+\alpha/2}}
\]

by fractional Taylor expansion. By a similar estimate as in (78), we have that for any \( z \in B(0, \varepsilon \sqrt{n}) \)

\[
\left| \log \left[ \chi_\rho \left( \frac{z}{\sqrt{n}} \right) \right] + \frac{\sum_j \nu_j |z_j|^2}{2} \right| \leq n \left( a(\mu) \left( \max_{|\beta| = 2} |D^\beta \chi_\rho(y)| \frac{|z|^2}{2n} \right)^2 + [D^2 u]_{C^\alpha(B)} \frac{|z|^{2+\alpha}}{2n^{1+\alpha/2}} \right)
\]

\[
\leq \frac{C_2 |z|^{2+\alpha}}{n^{\alpha/2}}.
\]

Using this estimate and the trivial bound \(|e^x - 1| \leq |x| |e^x| \) in 1 below, we find

\[
\left| \chi_\rho \left( \frac{z}{\sqrt{n}} \right)^n - \exp \left[ - \frac{1}{2} \sum_j \nu_j |z_j|^2 \right] \right| = \exp \left( - \frac{\sum_j \nu_j |z_j|^2}{2} \right) \left( \exp \left[ \left( \chi_\rho \left( \frac{z}{\sqrt{n}} \right)^n \right) \right] + \frac{\sum_j \nu_j |z_j|^2}{2} \right) - 1 \leq 1 \exp \left( - \frac{\sum_j \nu_j |z_j|^2}{2} + \frac{C_2 |z|^{2+\alpha}}{n^{\alpha/2}} \right)
\]

\[
\leq \exp \left( - \frac{\sum_j \nu_j |z_j|^2}{2} + \frac{C_2 |z|^{2+\alpha}}{n^{\alpha/2}} \right)
\]

For sufficiently small \( \varepsilon \), we thus get using that

\[
\left\| \chi_\rho \left( \frac{z}{\sqrt{n}} \right)^n - \exp \left[ - \frac{1}{2} \sum_j \nu_j |z_j|^2 \right] \right\|_{L^2(B(0, \varepsilon \sqrt{n}))} \leq \exp \left( - \frac{\sum_j \nu_j |z_j|^2}{2} + \varepsilon C_2 |z|^{2+\alpha} \right) C_2 |z|^{2+\alpha} \left\| \chi_\rho \left( \frac{z}{\sqrt{n}} \right)^n - \exp \left[ - \frac{1}{2} \sum_j \nu_j |z_j|^2 \right] \right\|_{L^2(B(0, \varepsilon \sqrt{n}))} \leq \exp \left( - \frac{\sum_j \nu_j |z_j|^2}{2} + \varepsilon C_2 |z|^{2+\alpha} \right) C_2 |z|^{2+\alpha} \left\| \chi_\rho \left( \frac{z}{\sqrt{n}} \right)^n - \exp \left[ - \frac{1}{2} \sum_j \nu_j |z_j|^2 \right] \right\|_{L^2(\mathbb{R}^{2m})}
\]

where the last line holds for \( \varepsilon \) small enough such that \( 2\varepsilon C_2 < 1 \). The claim then follows from using (88) in (72).
C. 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. Since $\rho \in D(L^2(\mathbb{R}^n))$ is assumed to have finite second moment (finite second phase space moments imply finite second standard moments), so does $\rho^{\otimes n}$. This implies the existence of $E > 0$ such that $\rho^{\otimes n}$ satisfies $\text{Tr}(\rho^{\otimes n} \sum_{i=1}^{m} a_i a_i^\dagger) \leq E$ for all $n$, and so does $\rho_G$. For any $E' > 0$, denote by $P_{E'}$ the projection onto the finite dimensional subspace generated by the eigenvectors of the multi-mode number operator $N_m := \sum_{i=1}^{m} a_i a_i^\dagger$ of eigenvalue less than $E'$. Then, by Markov’s inequality, for any $\varepsilon > 0$,

$$\text{Tr}(\rho^{\otimes n} P_{E/\varepsilon}), \text{Tr}(\rho_G P_{E/\varepsilon}) \geq 1 - \varepsilon.$$  

From the so-called ‘gentle measurement lemma’ [48, Lemma 9], we have that

$$\|\rho^{\otimes n} - P_{E/\varepsilon} \rho^{\otimes n} P_{E/\varepsilon}\|_1, \|\rho_G - P_{E/\varepsilon} \rho_G P_{E/\varepsilon}\|_1 \leq 2\sqrt{\varepsilon}.$$  

Then,

$$\|\rho^{\otimes n} - \rho_G\|_1 \leq \|\rho^{\otimes n} - P_{E/\varepsilon} \rho^{\otimes n} P_{E/\varepsilon}\|_1 + \|P_{E/\varepsilon}(\rho^{\otimes n} - \rho_G)\|_1 + \|P_{E/\varepsilon} \rho_G P_{E/\varepsilon} - \rho_G\|_1$$

$$\leq 4\sqrt{\varepsilon} + \|P_{E/\varepsilon}(\rho^{\otimes n} - \rho_G)\|_1 + \|P_{E/\varepsilon} \rho_G P_{E/\varepsilon} - \rho_G\|_2$$

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

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

We now turn to the proof of the convergence in relative entropy. Observe that, since $\rho^{\otimes n}$ and $\rho_G$ share the same first and second moments, $\text{Tr}(\rho^{\otimes n} \ln \rho_G) = \text{Tr}(\rho_G \ln \rho_G)$ and $D(\rho^{\otimes n} \| \rho_G) = S(\rho_G) - S(\rho^{\otimes n})$. The result follows directly from [25, Lemma 18].

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

A. 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 $f(x) = \frac{1}{\sqrt{3} + x^2}$. The characteristic function of this state can be computed thanks to (17), which in this case evaluates to

$$\chi_{|\psi_f\rangle\langle\psi_f|}(z) = \frac{\sqrt{2} e^{-|z|/\sqrt{2} + izR}}{\sqrt{2} + izR}. \quad (89)$$

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

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

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 [33, Proposition 5.16]. For all $r \in (0, 1)$, all positive (possibly unbounded) operators $A$, and all $|\psi\rangle \in \text{Dom } (A^{1/2})$, we have that

$$\left\| A^{r/2} |\psi\rangle \right\|^2 = \frac{\sin(\pi r)}{\pi} \int_0^\infty t^{r-1} \langle \psi | \frac{A}{tI + A} |\psi\rangle \, dt, \quad (90)$$

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

Proof of Proposition 11. We start by proving claim (a). Note that $2aa^\dagger = x^2 + p^2 + 1 \leq 2(x^2 + p^2)$, 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 (90) to $A + B$. For a generic $|\psi\rangle \in \text{Dom } (A^{r/2}) \cap \text{Dom } (B^{r/2})$, we obtain that

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

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

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

$$= \left\| A^{r/2} |\psi\rangle \right\|^2 + \left\| B^{r/2} |\psi\rangle \right\|^2,$$
where the inequality in the above derivation follows e.g. from Corollary 10.13. Now, setting \( A = x^2, B = p^2 \) and \( r = 1 - \delta \), we obtain that
\[
(a a^\dagger)^{1 - \delta} \leq |x|^{2(1 - \delta)} + |p|^{2(1 - \delta)} \leq |x|^{2(1 - \delta)} + 1 + p^2.
\]
Computing the expectation value on \( |\psi_f\rangle \) yields
\[
\langle \psi_f | (a a^\dagger)^{1 - \delta} | \psi_f \rangle \leq \langle \psi_f | \left( |x|^{2(1 - \delta)} + 1 + p^2 \right) | \psi_f \rangle = \frac{1}{\sqrt{\pi}} \Gamma \left( \frac{3}{2} - \delta \right) \Gamma(\delta) + 1 + \frac{7}{10},
\]
where the last step is by explicit computation. This proves (a). We now move on to (b). For this we evaluate the characteristic function of the convolution \( |\psi_f \rangle \otimes |\psi_f \rangle \otimes |\psi_f \rangle \) on the purely imaginary line. For \( t \in \mathbb{R} \), using (17) we obtain that
\[
\chi_{|\psi_f \rangle \otimes |\psi_f \rangle \otimes |\psi_f \rangle} (it) = \int_{-\infty}^{\infty} dx |f(x)|^2 e^{\sqrt{2} i t x} = \sqrt{2} |t| K_1 \left( \sqrt{2} |t| \right),
\]
were \( K_1 \) is a modified Bessel function of the second kind, and the last equality follows from (55) and Eq. (9.6.25)]. Therefore, for any fixed \( t > 0 \) it holds that
\[
\lim_{n \to \infty} \chi_{|\psi_f \rangle \otimes |\psi_f \rangle \otimes |\psi_f \rangle} \left( \frac{it}{\sqrt{n}} \right) = \lim_{n \to \infty} \chi_{|\psi_f \rangle \otimes |\psi_f \rangle} \left( \frac{it}{\sqrt{n}} \right) = \lim_{n \to \infty} \left( 1 + \left( c + \log t - \frac{1}{2} \log n \right) \frac{t^2}{n} + O \left( n^{-3/2} \right) \right)^n = 0,
\]
where we have used the expansion in Eq. (9.6.53) (see also Eq. (6.3.2) and (9.6.7)). Since \( \chi_{|\psi_f \rangle \otimes |\psi_f \rangle \otimes |\psi_f \rangle} (0) = 1 \) for all \( n \) because \( |\psi_f \rangle \otimes |\psi_f \rangle \otimes |\psi_f \rangle \) is a valid quantum state, the sequence of functions \( \chi_{|\psi_f \rangle \otimes |\psi_f \rangle \otimes |\psi_f \rangle} \) does not possess a continuous limit. Hence, it cannot converge to the characteristic function of any quantum state. This proves (b).

\[\Box\]

### B. 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 \( \rho \) be a 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 \( \rho^{\otimes n} \) converges at least with rate \( \Theta(n^{-1}) \) to its Gaussification
\[
\| \rho^{\otimes n} - \rho_G \|_2 = \Theta \left( n^{-1} \right).
\]
In particular, every density operator satisfying the assumptions of Theorem 6 that is diagonal in the Fock basis achieves a \( \Theta(n^{-1}) \) rate.

The proof of Proposition 20 is provided in Section VI B. The following example shows that the \( \Theta(n^{-1}) \) convergence rate stated in Proposition 20, under the assumption that \( D^3 \chi_\rho(0) = 0 \), is in fact attained.

**Example 4 \((\Theta(n^{-1})\)-rate).** By Proposition 20 we can take \( \rho = |1\rangle \langle 1| \) to obtain a convergence rate of at least \( \Theta(n^{-1}) \) in the QCLT. That the \( \Theta(n^{-1}) \) rate is actually attained is illustrated in the right figure in Figure 4. The \( \Theta(n^{-1}) \) rate is saturated both in Hilbert–Schmidt and trace norm.
FIG. 4: This plot shows the expressions $cn^\alpha \|\rho^{\otimes n} - \rho\|$ for a constant $c > 0$ such that $\lim_{n \to \infty} cn^\alpha \|\rho^{\otimes n} - \rho\| = 1$. The left figure shows that the $O(1/\sqrt{n})$ convergence rate is sharp (Theorem 7) by using the state from Ex. 5. The right figure shows that we can obtain a rate $O(1/n)$ if $D^3 \chi^\rho(0) = 0$ (Theorem 6) by using the state from Ex. 4. In both figures we write $\rho^n$ for $\rho^{\otimes n}$.

The following example shows that the $O(n^{-1/2})$ of Theorem 7 is attained.

**Example 5** ($O(n^{-1/2})$-rate). Consider the state
\[ \rho = \frac{|0\rangle + |3\rangle \langle 0| + |3\rangle \langle 3|}{\sqrt{2}}. \]

Its characteristic function is explicitly given by
\[ \chi^\rho(z) = \frac{e^{-\frac{\pi}{4} |z|^2}}{12} \left( 12 - 2i\sqrt{6}\pi z^3 - 3z^2 \right) - 18\pi |z|^2 + 9\pi^2 |z|^4 - \pi^3 |z|^6 \]
where $z = z_R + iz_L$ since $\langle 0|\rho|3\rangle \neq 0$. Now, the condition of Proposition 20 does not hold. One verifies directly that $D\chi^\rho(0) = 0$ and $D^2 \chi^\rho(0) = \left( \begin{array}{cc} -4\pi & 0 \\ 0 & -4\pi \end{array} \right)$. Let $\Phi(z) = e^{-2\pi |z|^2}$. We then find that $\|\chi^\rho - \Phi\|_{L^2(\mathbb{R}^2)}$ converges with rate $n^{-1/2}$, see Figure 4.

**VIII. CASCADE OF BEAM SPLITTERS: PROOFS**

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

**A. Generalities of the cascade channels**

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

---

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.
Lemma 21. For an \( m \)-mode quantum state \( \rho \), some \( \lambda \in [0, 1] \), and a positive integer \( n \), consider the cascade channel \( \mathcal{N}^n_{\rho, \lambda} \) (cf. \( \mathcal{N}^{(n)}_\rho \)). One has that
\[
\mathcal{N}^n_{\rho, \lambda} = \mathcal{N}_{\rho, \lambda}^n, \quad \chi_{\mathcal{N}^n_{\rho, \lambda}}(\omega)(z) = \chi_{\mathcal{N}_{\rho, \lambda}^n}(\omega) = \chi_\omega \left( \sqrt{\lambda} z \right) \chi_{\rho, \lambda}(\sqrt{1 - \lambda} z),
\]
where the effective environment state \( \rho, \lambda \) is defined via its characteristic function
\[
\chi_{\rho, \lambda}(z) := \prod_{j=1}^n \chi_{\rho, \lambda} \left( \sqrt{1 - \lambda^{1/n}} \lambda^{j-1} \lambda^{1/n} z \right).
\]

Proof. We proceed by induction. The case \( n = 1 \) follows from \( \lambda^n \). Let us assume that the claim holds for \( n - 1 \), so that
\[
\mathcal{N}^{n-1}_{\rho, \mu} = \mathcal{N}_{\rho, \mu}^{n-1}, \quad \chi_{\mathcal{N}^{n-1}_{\rho, \mu}}(\omega)(z) = \chi_{\mathcal{N}_{\rho, \mu}^{n-1}}(\omega) = \chi_\omega \left( \sqrt{\mu} z \right) \prod_{j=1}^{n-1} \chi_{\rho, \mu} \left( \sqrt{1 - \mu^{1/(n-1)}} \mu^{j-1} \mu^{1/(n-1)} z \right).
\]
By setting \( \mu = \lambda^{(n-1)/n} \) we see that
\[
\mathcal{N}^{n-1}_{\rho, \lambda^{(n-1)/n}} = \mathcal{N}_{\rho, \lambda^{(n-1)/n}}^{n-1}, \quad \chi_{\mathcal{N}^{n-1}_{\rho, \lambda^{(n-1)/n}}}(\omega)(z) = \chi_{\mathcal{N}_{\rho, \lambda^{(n-1)/n}}^{n-1}}(\omega) = \chi_\omega \left( \sqrt{\lambda} \lambda^{(n-1)/n} z \right) \prod_{j=1}^{n-1} \chi_{\rho, \lambda^{(n-1)/n}} \left( \sqrt{1 - \lambda^{1/n}} \lambda^{j-1} \lambda^{1/n} z \right).
\]
Since \( \mathcal{N}^n_{\rho, \lambda}(\omega) = \mathcal{N}_{\rho, \lambda}^n \left( \mathcal{N}^{n-1}_{\rho, \lambda^{(n-1)/n}}(\omega) \right) = \mathcal{N}^{n-1}_{\rho, \lambda^{(n-1)/n}}(\omega) \boxtimes \lambda^{1/n} \rho \), composition with the \( n \)-th copy of the channel \( \mathcal{N}^n_{\rho, \lambda} \) yields
\[
\begin{aligned}
\chi_{\mathcal{N}^n_{\rho, \lambda}}(\omega)(z) &= \chi_{\mathcal{N}^{n-1}_{\rho, \lambda^{(n-1)/n}}}(\omega) \left( \lambda^{1/2n} z \right) \chi_{\rho} \left( \sqrt{1 - \lambda^{1/n}} z \right) \\
&= \chi_\omega \left( \sqrt{\lambda} z \right) \prod_{j=1}^{n-1} \chi_{\rho} \left( \sqrt{1 - \lambda^{1/n}} \lambda^{j-1} \lambda^{1/n} z \right) \chi_{\rho} \left( \sqrt{1 - \lambda^{1/n}} \lambda^{j-1} \lambda^{1/n} z \right) \\
&= \chi_\omega \left( \sqrt{\lambda} z \right) \prod_{j=0}^{n-1} \chi_{\rho} \left( \sqrt{1 - \lambda^{1/n}} \lambda^{j} \lambda^{1/n} z \right) \\
&= \chi_\omega \left( \sqrt{\lambda} z \right) \prod_{j=1}^{n} \chi_{\rho} \left( \sqrt{1 - \lambda^{1/n}} \lambda^{j-1} \lambda^{1/n} z \right),
\end{aligned}
\]
which proves \( \text{(94)} \) and \( \text{(95)} \). Finally, one can also verify by induction that
\[
\rho(\lambda, n) = \rho \left( \lambda^{n-1/n}, n-1 \right) \boxtimes \eta(\lambda, n) \rho,
\]
where \( \eta(\lambda, n) := \frac{\lambda^{1/n}(1-\lambda^{(n-1)/n})}{1-\lambda} \in [0, 1] \), so that \( \rho(\lambda, n) \) is a legitimate quantum state for all \( \lambda \in [0, 1] \) and all \( n \).

B. On the effective environment state

Thanks to Lemma 21, the study of the the cascade channel \( \mathcal{N}^n_{\rho, \lambda} \) boils down to that of the iteratively convolved state \( \rho(\lambda, n) \) of \( \text{(95)} \). Since such a convolution is not symmetric (cf. \( \text{(47)} \)), 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 \( \mathbb{I} \) \( \mathbb{I} \), and even to sequences of correlated random variables \( \mathbb{I} \) \( \mathbb{I} \). Rates of convergence for the former case can be found for instance in \( \mathbb{I} \) \( \mathbb{I} \) (see e.g. Theorem 13.3 of \( \mathbb{I} \) \( \mathbb{I} \)).
Proposition 22. Let $\rho$ be a centred $m$-mode quantum state with finite second-order phase space moments. Then the sequence of quantum states $\rho(\lambda, n)$ defined in (95) converges to the Gaussification $\rho_G$ of $\rho$ in trace norm. Moreover, if $\rho$ has finite third-order phase space moments $M_3'$, defined in (31), then

$$\|\rho(\lambda, n) - \rho_G\|_2 = \Theta_{M_3'} \left( n^{-1/2} \right),$$

$$\|\rho(\lambda, n) - \rho_G\|_1 = \Theta_{M_3'} \left( n^{-1/(2m+2)} \right).$$

Proof. The argument is a variation of that used to prove Theorem 7 in Section VI B. To simplify the notation we write

$$\chi_{\rho,j,n}(z) := \chi_{\rho} \left( \sqrt{1 - \lambda^{1/n}} \lambda^{i-1} \frac{1}{\pi n} z \right).$$

We record that by L'Hôpital's rule,

$$\sqrt{1 - \lambda^{1/n}} = \sqrt{-\frac{\log(\lambda)}{n}} (1 + o(1)) \text{ as } n \to \infty.$$  (99)

This implies that there are $C, c > 0$ such that for all $n$ and $j \in \{1, \ldots, n\}$

$$c \sqrt{-\frac{\log(\lambda)}{n}} \lambda \leq \sqrt{1 - \lambda^{1/n} \lambda^{(j-1)/(2n)}} \leq C \sqrt{-\frac{\log(\lambda)}{n}}.$$  (100)

We also notice that the sum of the second derivatives of the characteristic functions $\chi_{\rho,j,n}$ is just given by

$$V := \sum_{j=1}^n D^2 \chi_{\rho,j,n}(0)(1 - \lambda^{1/n}) \lambda^{i-1} = D^2 \chi_{\rho}(0)(1 - \lambda^{1/n}) \sum_{j=1}^n \lambda^{i-1}$$

$$= D^2 \chi_{\rho}(0)(1 - \lambda).$$

Thus, we see that $V$ does not depend on $n$. Using the Williamson form, we may for simplicity just assume that $V = \text{diag}(\nu_1, \ldots, \nu_{2m})$ for $\nu_1, \ldots, \nu_{2m} \geq 1$ in the sequel. We then have by Taylor expansion for $z \in B := B(0, R)$ due to (100) that

$$|\chi_{\rho,j,n}(z) - 1| = \Theta \left( \max_{|\alpha|=2, y \in B} |D^\alpha \chi_{\rho}(y)| \frac{R^2}{2n} \right),$$

$$\left| \sum_{j=1}^n \chi_{\rho,j,n}(z) - 1 + \frac{\sum_j \nu_j |z_j|^2}{2n} \right| = \Theta \left( \max_{|\alpha|=3, y \in B} |D^\alpha \chi_{\rho}(y)| \frac{R^2}{6n^3} \right).$$

(102)

We can therefore estimate for $|z| \leq \sqrt{n} \varepsilon$ by the Taylor series of the logarithm in 1 (below),

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

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

$$= \left| \sum_{j=1}^n \left( \sum_{i=2}^\infty \frac{(1 - \chi_{\rho,j,n}(z))^i}{i} + \chi_{\rho,j,n}(z) - 1 + \frac{\sum_j \nu_j |z_j|^2}{2n} \right) \right|$$

$$\leq \left| \sum_{j=1}^n \sum_{i=2}^\infty \frac{(1 - \chi_{\rho,j,n}(z))^i}{i} \right| + \left| \sum_{j=1}^n \left( \chi_{\rho,j,n}(z) - 1 + \frac{\sum_j \nu_j |z_j|^2}{2n} \right) \right| \leq \Theta \left( \frac{|z|^3}{\sqrt{n}} \right),$$

(103)
where we used (102) and an estimate similar to (78) in 2. Hence, we find using (103) and the trivial bound $|e^{x} - 1| \leq |x|e^{|x|}$ in 1 (below) that for some constant $C > 0$

$$
\left| \prod_{j=1}^{n} \chi_{\rho,j,n}(z) - \exp \left[ -\frac{1}{2} \sum_{j} \lambda_{j} |z_{j}|^{2} \right] \right| \\
= \exp \left( -\sum_{j} \lambda_{j} |z_{j}|^{2} \right) \exp \left( \log \left( \prod_{j=1}^{n} \chi_{\rho,j,n}(z) \right) + \sum_{j} \lambda_{j} |z_{j}|^{2} \right) - 1
\leq n^{-1/2} \left\| \exp \left( -\sum_{j} \lambda_{j} |z_{j}|^{2} + C|z|^{2} \right) C|z|^{3} \right\|_{L^{2}(B(0,\varepsilon \sqrt{n}))}  \\
\leq n^{-1/2} \left\| \exp \left( -\sum_{j} \lambda_{j} |z_{j}|^{2} + C|z|^{2} \right) C|z|^{3} \right\|_{L^{2}(\mathbb{R}^{2m})} = O(n^{-1/2}).
$$

(104)

For sufficiently small $\varepsilon$, such that $2C\varepsilon < 1$, we thus get

$$
\left\| \prod_{j=1}^{n} \chi_{\rho,j,n}(z) - \exp \left[ -\frac{1}{2} \sum_{j} \lambda_{j} |z_{j}|^{2} \right] \right\|_{L^{2}(B(0,\varepsilon \sqrt{n}))} \\
\leq n^{-1/2} \left\| \exp \left( -\sum_{j} \lambda_{j} |z_{j}|^{2} + C|z|^{2} \right) C|z|^{3} \right\|_{L^{2}(B(0,\varepsilon \sqrt{n}))}  \\
\leq n^{-1/2} \left\| \exp \left( -\sum_{j} \lambda_{j} |z_{j}|^{2} + C|z|^{2} \right) C|z|^{3} \right\|_{L^{2}(\mathbb{R}^{2m})} = O(n^{-1/2}).
$$

(105)

By (100) there exists a sufficiently small $\delta > 0$ such that for all $|z| \geq \varepsilon \sqrt{n}$ and $j \in \{1, \ldots, n\}$ for $n$ large enough

$$
\sup_{|z| \geq \varepsilon \sqrt{n}} |\chi_{\rho}(1 - \lambda^{1/n})^{1/2} \lambda^{1/2} \frac{j-1}{n} z| \leq \sup_{|z| \geq \delta} |\chi_{\rho}(z)| < 1.
$$

(106)

Decomposing the integration range as in Lemma 17 and using (106) yields then the claim

$$
\left\| \prod_{j=1}^{n} \chi_{\rho,j,n}(z) - \exp \left( -\sum_{j} \lambda_{j} |z_{j}|^{2} / 2 \right) \right\|_{L^{2}(\mathbb{R}^{2m})} = O(n^{-1/2}),
$$

which concludes the proof.

\[\square\]

C. 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 $\rho$ 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 $\rho_{1}$ and $\rho_{2}$, and some $\lambda \in [0,1]$, the corresponding channels defined as in (39) satisfy

$$
\| \mathcal{N}_{\rho_{1},\lambda} - \mathcal{N}_{\rho_{2},\lambda} \|_{\diamond} \leq \| \rho_{1} - \rho_{2} \|_{1}.
$$

(107)
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
\[
\| (\mathcal{N}_{\rho_1, \lambda} - \mathcal{N}_{\rho_2, \lambda}) \otimes \text{id}_R(\omega) \|_1 = \| \text{Tr}_R[U_\lambda(\omega \otimes \rho_1)U_\lambda^\dagger] - \text{Tr}_R[U_\lambda(\omega \otimes \rho_2)U_\lambda^\dagger] \|_1
\]
\[
\leq \| U_\lambda(\omega \otimes \rho_1 - \omega \otimes \rho_2)U_\lambda^\dagger \|_1
\]
\[
= \| \rho_1 - \rho_2 \|_1 ,
\]
where the inequality stems from the monotonicity of trace distance under quantum channels. \qed

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

Proof of Theorem 12. Recall from Lemma 21 that $\mathcal{N}_{\rho, \lambda} = \mathcal{N}_{\rho(\lambda, n), \lambda}$, where $\rho(\lambda, n)$ is the state with characteristic function given by (95). Applying Lemma 23 and Proposition 22, we have that
\[
\left\| \mathcal{N}_{\rho(\lambda, n), \lambda} - \mathcal{N}_{\rho_G(\lambda), \lambda} \right\|_\infty \leq \left\| \rho(\lambda, n) - \rho_G \right\|_1 \leq O(m n^{-1/(2m+2)}) ,
\]
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 (57) and (59). Our aim is to apply [26, Theorem 9] to the two channels $\mathcal{N}_{\rho, \lambda}^{\lambda}$ and $\mathcal{N}_{\rho_G(\lambda), \lambda} = \mathcal{A}_\lambda^N$, for the special case $m = 1$ (cf. (40)). We set
\[
\epsilon_n := \frac{1}{2} \left\| \mathcal{N}_{\rho, \lambda}^{\lambda} - \mathcal{A}_\lambda^N \right\|_\infty \geq \frac{1}{2} \left\| \mathcal{N}_{\rho, \lambda}^{\lambda} - \mathcal{A}_\lambda^N \right\|_E ,
\]
where the energy-constrained diamond norm is defined with respect to the canonical Hamiltonian, namely the number operator $H = a^\dagger a$ (see [28, Eq. (2)] and [26, Eq. (2)]). Note that $\epsilon_n = O(M_3^2 (n^{-1})$ by Theorem 12.

The input–output energy relations can be easily determined for both channels thanks to (96) and (38), 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
\[
\text{Tr} \left[ \mathcal{N}_{\rho, \lambda}^{\lambda}(\omega) a^\dagger a \right] = \text{Tr} \left[ \mathcal{N}_{\rho(\lambda, n), \lambda}(\omega) a^\dagger a \right] = \lambda \text{Tr} \left[ \omega a^\dagger a \right] + (1 - \lambda)N ,
\]
\[
\text{Tr} \left[ \mathcal{A}_\lambda^N(\omega) a^\dagger a \right] = \lambda \text{Tr} \left[ \omega a^\dagger a \right] + (1 - \lambda)N .
\]
This means that we can set $\alpha = \lambda$ and $E_0 = (1 - \lambda)N$, and hence $\tilde{E} = \lambda E + (1 - \lambda)N$, in [26, Theorem 9]. We obtain that
\[
\left| \mathcal{Q} (\mathcal{N}_{\rho, \lambda}^{\lambda}, E) - \mathcal{Q} (\mathcal{A}_\lambda^N, E) \right| \leq 56 \sqrt{\epsilon_n} g \left( \frac{4(\lambda E + (1 - \lambda)N)}{\epsilon_n} \right) + 6 g (4 \sqrt{\epsilon_n})
\]
\[
\leq c \sqrt{\epsilon_n} \log \frac{1}{\epsilon_n}
\]
\[
\leq C(M_3^2) \frac{\log n}{n^{1/8}} .
\]
Here, in step 1 we applied [26, Theorem 9] together with the formula $S(\tau_N) = g(N)$ (see [26] and [27]); 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 $\epsilon_n \leq C(M_3^2) n^{-1/4}$ 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 $\mathcal{N}_{ρ, λ}^\text{in}$ in terms of that of the thermal attenuator $\mathcal{A}_λ^\text{N}$ of (40), in turn given by (41). Although we could use the estimates in [28], we prefer to resort to the tighter ones provided in [28]. We obtain that

$$\left| C\left(\mathcal{N}_{ρ, λ}^\text{in}, E\right) - C\left(\mathcal{A}_λ^\text{N}, E\right) \right| \leq 7\varepsilon_n \left( \log (\lambda E + (1 - \lambda)N + 1) - \log \frac{\varepsilon_n}{2} + 1 \right) + 2g \left( \frac{5\varepsilon_n}{2} \right) + 4h_2 \left( \frac{\varepsilon_n}{2} \right)$$

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

The inequality in 4 is an application of [28] Proposition 6]. To see why, let us re-write the result of Shirokov [28] Proposition 6] for one-mode channels and with respect to the canonical Hamiltonian as

$$\left| C\left(\mathcal{N}_1, E\right) - C\left(\mathcal{N}_2, E\right) \right| \leq 2\varepsilon \left( 2t + r_ε(t) \right) \log \left( E' + 1 \right) + \log (et) ) + 2g (εr_ε(t)) + 4h_2(εt).$$

Here, $\mathcal{N}_i$ ($i = 1, 2$) are two quantum channels with $\frac{1}{2} \| M_1 - M_2 \|_E \leq ε$, we picked $E'$ such that $\sup_{\rho \in \mathcal{M}_1} \| a^\dagger a - ρ \|_E = E'$, the function $r_ε$ is defined by $r_ε(t) := \frac{1 + t}{1 - t}$, and $h_2(x) := -x \log x - (1 - x) \log (1 - x)$ is the binary entropy. Setting $\mathcal{N}_1 = \mathcal{N}_{ρ, λ}^\text{in}, \mathcal{N}_2 = \mathcal{A}_λ^\text{N}$, we see that $E' = λE + (1 - λ)N$ (cf. [108] and [28] Eq. (21)]); choosing $t = 1/2$ and hence $r_ε(t) \leq r_1(t) = r_1(1/2) = 5/2$ 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.

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

**ACKNOWLEDGEMENTS**

ND would like to thank M. Jabbour for helpful discussions. LL acknowledges financial support from the European Research Council under the Starting Grant GQCO (Grant no. 637352) and from Ulm university; he is also grateful to V. Giovannetti and A. Holevo for many discussions on the implications of Lemma [16] SB thanks G. Baverez for interesting discussions on stable laws and gratefully acknowledges support by the EPSRC grant EP/L016516/1 for the University of Cambridge CDT, the CCA. CR acknowledges financial support from the TUM university Foundation Fellowship and by the DFG cluster of excellence 2111 (Munich Center for Quantum Science and Technology).

**Appendix A: Standard moments vs 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 $c_{k, m}(ε) < \infty$ such that the standard moments and phase space moments, defined by [29] and [31], respectively, satisfy $M'_k(ρ, ε) \leq c_{k, m}(ε) M_k(ρ)$ 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

$$\left( \sum_j a_j a_j^\dagger \right)^{k/2} \geq d_{k, m} \sum_j \left( |x_j|^k + |p_j|^k \right), \quad (A1)$$
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 $j^{th}$ 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 $a_ja_j^\dagger$ commute with each other, and employing standard inequalities between $p$-norms, we deduce that

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

Therefore, from now on we look at the one-mode case only. The vector space $V_m := \text{span}\{|n\rangle : n \in \mathbb{N}_0\}$ of states with a finite expansion in the Fock basis is a core for both $(aa^\dagger)^{k/2}$, 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 $h := \lfloor k/2 \rfloor \in \mathbb{N}_0$. Since the function $A \mapsto A^r$ is well known to be operator monotone [33 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 $\Pi_N$ as the projector onto span of the first $N + 1$ Fock states (from 0 to $N$), we have to show that

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

where the inequality now involves only matrices. To show that $\Pi_N$ is positive semidefinite, it suffices that show that $\Pi_N$ is diagonally dominant, i.e. for all $N$ and $0 \leq n \leq N$ it holds that

$$
\langle n|\Pi_N|n\rangle - \sum_{n'=0,\ldots,N, n' \neq n} |\langle n|\Pi_N|n'\rangle| \geq 0. \quad (A2)
$$
Writing down the left-hand side yields

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

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

\[
\geq (1 + n)^h - d_h \sum_{n': |n' - n| \leq 2h} | \langle n | x^{2h} | n' \rangle |
\]

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

\[
\geq (1 + n)^h - d_h F_h (n + 2h)^h
\]

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 \((6)\) to expand

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

In 3 we applied standard estimates for factorials: for example, when \( n \leq n' \leq n + 2h \) we used the fact that \( n! \leq (n')^{n'-n}/2 \leq (n')^{n'-n+2} \leq (2h)^{2h} \leq (n + 2h)^{2h} \); moreover, we defined \( F_h := \sum_{\ell, \ell' \geq 0, \ell + \ell' \leq 2h} \left( |f_{h, \ell, \ell'}|^2 + |f_{h, \ell, \ell'}^*|^2 \right) \). Finally, 4 follows by choosing e.g. \( d_h^{-1} = (2h + 1)^h \).

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 \( (A + B)^k \geq d_k (A^k + B^k) \) 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 \( \lambda^k \); while that of \( A^k + B^k = A + B \) is clearly \( \lambda \). 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 \( \rho \) with finite \( k \)-moments \( M_k(\rho) \), as
defined by \[^{27}\] , also satisfies

\[
M_k(\rho, \varepsilon) = \| \chi_\rho \|_{C^k(\mathcal{B}(0,\varepsilon))} \leq c_{k,m}(\varepsilon) M_k(\rho). \tag{A3}
\]

In particular, according to \[^{31}\] , \( \rho \) has finite phase space moments of order up to \( k \).

**Proof.** Let \( \rho \) be an \( m \)-mode quantum state. We start by considering the modified state \( \sigma := \rho \boxplus |0\rangle \langle 0| \) that is obtained by convolving it with the (multi-mode) vacuum state according to the rule \[^{35}\] (for \( \lambda = 1/2 \)). A first important observation is that the moments of \( \sigma \) and \( \rho \) are related. Namely,

\[
M_k(\sigma) \leq M_k(\rho) \quad \forall k \in [0, \infty). \tag{A4}
\]

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

\[
\langle n|\sigma|n \rangle = \langle n|\Delta(\rho \boxplus |0\rangle \langle 0|)|n \rangle \\
= \sum_{\ell \in \mathbb{N}_0^m} \langle \ell|\rho|\ell \rangle \langle n|(|\ell\rangle \langle \ell| \boxplus |0\rangle \langle 0|)|n \rangle \\
= \sum_{\ell \geq n} \langle \ell|\rho|\ell \rangle 2^{-|\ell|} \binom{|\ell|}{n}.
\]

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 \( \Delta \) commutes with the quantum convolution operator, in the sense that \( \Delta(\omega_1 \boxplus \omega_2) = \Delta(\omega_1) \boxplus \Delta(\omega_2) \) for all \( m \)-mode quantum states \( \omega_1, \omega_2 \). To show this, first exploit linearity and factorisation of \( \Delta \) 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 \( U \) commutes with the total Hamiltonian \( a^\dagger a + b^\dagger b \). 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 \( \ell \), employing the relations \[^{33}\]. Computing the \( k \)-th moment of \( \sigma \) then yields

\[
M_k(\sigma) \overset{4}{=} \sum_{n \in \mathbb{N}_0^m} (m + |n|)^{k/2} \langle n|\sigma|n \rangle \\
\overset{5}{=} \sum_{\ell \in \mathbb{N}_0^m} \langle \ell|\rho|\ell \rangle \sum_{n \leq \ell} (m + |n|)^{k/2} 2^{-|\ell|} \binom{|\ell|}{n} \\
\overset{6}{=} \sum_{\ell \in \mathbb{N}_0^m} \langle \ell|\rho|\ell \rangle (m + |\ell|)^{k/2} \\
= M_k(\rho).
\]

Here, 4 and 7 follows from the representation in \[^{30}\] ; 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 \( P_{t}(n) := 2^{-|\ell|} \binom{|\ell|}{n} \) form
a probability distribution over the set of multi-indices \( n \in \mathbb{N}^m_0 \) with \( n \leq \ell \). This proves that the \( k^{\text{th}} \) moments of \( \sigma \) are upper bounded by those of \( \rho \).

The state \( \sigma \) is also useful because its characteristic function is a close relative of that of \( \rho \). Namely, according to (36) we have that
\[
\chi_{\sigma}(z) = \chi_{\rho}(z/\sqrt{2}) e^{-\|z\|^2/4},
\]
and hence
\[
\|\chi_{\rho}\|_{C^k(B(0,\varepsilon))} \leq g_{k,m}(\varepsilon) \|\chi_{\sigma}\|_{C^k(B(0,\varepsilon))} \tag{A6}
\]
for some constants \( g_{k,m}(\varepsilon) \). Thus, it suffices to find a suitable upper estimate for the norm \( \|\chi_{\sigma}\|_{C^k(B(0,\varepsilon))} \).

By Lemma [16] the Fourier transform of \( \chi_{\sigma} \), i.e. the Wigner function \( W_{\sigma} \) of \( \sigma \), is everywhere non-negative. Hence, \( \chi_{\sigma} \) can be seen as the characteristic function of a classical random variable \( Z \) over \( \mathbb{C}^m \), with probability density function \( W_{\sigma} \). If we show that \( Z \) has finite absolute moments of order \( k \), then thanks to [43] Theorem 1.8.15 we deduce that \( \chi_{\sigma} \) is \( k \)-fold differentiable everywhere, and since
\[
\left| \partial_\alpha \partial_\beta \chi_{\sigma}(z) \right| = \left| \int d^m u \left( \prod_j (-u_j)^{\beta_j} (u_j^*)^{\alpha_j} \right) W_{\sigma}(u) e^{z^\top u^* - z^\top u} \right| \leq \int d^m u \|u^{\alpha + \beta}\| W_{\sigma}(u) \tag{A7}
\]
for all multi-indices \( \alpha, \beta \in \mathbb{N}^m_0 \), we in fact have that
\[
\|\chi_{\sigma}\|_{C^k(B(0,\varepsilon))} \leq 1 + \int d^m u \|u\| W_{\sigma}(u) =: 1 + L_k(\sigma) \tag{A8}
\]

Therefore, we now look at the quantity \( L_k(\sigma) \). For a vector \( u = u_R + i u_I \in \mathbb{C}^m \), with \( u_R, u_I \in \mathbb{R}^m \), we observe that
\[
\|u\|^k = \left( \sum_j (u_{Rj}^2 + u_{Ij}^2) \right)^{k/2} \leq (2m)^{\max\left\{ \frac{k}{2} - 1, 0 \right\}} \left( \sum_j \left| u_{Rj} \right|^k + \left| u_{Ij} \right|^k \right).
\]
Thus,
\[
L_k(\sigma) \leq (2m)^{\max\left\{ \frac{k}{2} - 1, 0 \right\}} \int d^m u_R d^m u_I \sum_j \left( \left| u_{Rj} \right|^k + \left| u_{Ij} \right|^k \right) W_{\sigma}(u)
\]
\[
= (2m)^{\max\left\{ \frac{k}{2} - 1, 0 \right\}} \left( \int d^m u_R \left( \sum_j \left| u_{Rj} \right|^k \right) \right) \int d^m u_I W_{\sigma}(u)
\]
\[
+ \int d^m u_I \left( \sum_j \left| u_{Ij} \right|^k \right) \int d^m u_R W_{\sigma}(u)
\]
\[
\leq (2m)^{\max\left\{ \frac{k}{2} - 1, 0 \right\}} 2^{-\frac{2}{2}(m+k-1)} \left( \text{Tr} \left[ \sigma \left( \sum_j \left| x_j \right|^k \right) \right] + \text{Tr} \left[ \sigma \left( \sum_j \left| p_j \right|^k \right) \right] \right)
\]
\[
\leq (2m)^{\max\left\{ \frac{k}{2} - 1, 0 \right\}} 2^{-\frac{2}{2}(m+k-1)} d_{k,m}^{-1} \text{Tr} \left[ \sigma \left( \sum_j a_j a_j^\top \right)^k \right]^{1/2}
\]
\[
\leq c_{k,m} M_k(\sigma). \tag{10}
\]

In the above derivation, the identity in 8 can be verified by first reducing to the case of a pure \( \sigma \), which can be done by linearity and by multiple applications of Tonelli’s theorem, and by subsequently remembering that for a pure state \( |\psi\rangle \) with wave function \( f \in L^2(\mathbb{R}^m) \) it holds e.g. that
\[
\int d^m u R_W_{|\psi\rangle\langle\psi|}(u) = \sqrt{2} \| f(\sqrt{2} u_R) \|^2. \tag{9}
\]
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 (A6), (A8), and (A4), we deduce that
\[
M_k(\rho, \varepsilon) = \|\chi_{\rho}\|_{C^k(B(0,\varepsilon))} \leq g_{k,m}(\varepsilon) \|\chi_{\sigma}\|_{C^k(B(0,\varepsilon))} \leq g_{k,m}(\varepsilon)(1 + L_k(\sigma)) \leq 2 g_{k,m}(\varepsilon)c_{k,m} \leq 2 g_{k,m}(\varepsilon)c_{k,m} M_k(\rho) =: c_{k,m}(\varepsilon) M_k(\rho),
\]
which concludes the proof.
Appendix B: Standard moments vs phase space moments: the fractional case

In the last section, we showed that the \( k \)-th phase space moment was controlled by the \( k \)-th 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 \([45, \text{Theorem } 5.4.1]\) satisfy

\[
(L^p(w_0), L^p(w_1))_\theta = L^p(w_\theta)
\]

where \( w_\theta := w_0^{-\theta} w_1^{\theta} \). This would allow us to extend the estimate in (A7) 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 \( \rho \) be an \( m \)-mode quantum state and \( k \geq 0 \). If \( \text{Tr} \left[ \rho \left( \sum_j a_j a_j^\dagger \right)^{k/2} \right] < \infty \), then

\[
\| \chi_\rho \|_{C^k(B(0,\varepsilon))} < \infty \text{ for some } \varepsilon > 0.
\]

Moreover,

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

for some constant \( C_\varepsilon > 0 \).

We have seen in Appendix A that the map \( \rho \to \chi_\rho \) is bounded from \( \mathcal{W}^{k,1}(\mathcal{H}_m) \) 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

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

for some positive constant \( C_\varepsilon \) that comes from the bounds derived in Section A for \( k_0 \) and \( k_1 \). It only remains to prove that the interpolated norms \( \| \rho \|_{(\mathcal{W}^{k_0,1}(\mathcal{H}_m), \mathcal{W}^{k_1,1}(\mathcal{H}_m))_\theta} \) can further be bounded above by \( \| \rho \|_{\mathcal{W}^{(1-\theta)k_0 + \theta k_1,1}} \). First, we recall a useful technical lemma [74, 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

\[
\| T_{21} \|_1 \leq \frac{1}{2} (\| T_{11} \|_1 + \| T_{22} \|_1).
\]

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

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

(B2)

First, we restrict attention to states \( \rho \) that are orthogonal in the Fock basis. We then write \( \Pi_E \) for the spectral projection onto the Fock states of energy at most \( E \), that is \( \Pi_E := \sum_{j\in\mathcal{E}} |j\rangle \langle j| \).

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

\[
\forall n \leq N_0(t) : \gamma_n \leq t^{-1} \text{ and } \gamma_{N_0(t)+1} \geq t^{-1},
\]
Let then \( S \) and \( S^\dagger \)
where \( \alpha_n \) is the upper bound
\[
\|\rho\|_{(\mathcal{F}_k^0(\mathcal{H}_1),\mathcal{F}_k^1(\mathcal{H}_1))_\theta} \leq \sup_{t>0} \sum_{n=0}^{\infty} \langle n | \rho \rangle \| n \| (n + 1)^{k_0/2} t^{-\theta} (\alpha_n + t_{\gamma_n} \beta_n).
\]
We now recall that for \( \gamma_n \leq t^{-1} \) we have \( \alpha_n = 0 \) and \( \beta_n = 1 \) such that
\[
t^{-\theta} (\alpha_n + t_{\gamma_n} \beta_n) = t^{-\theta} t_{\gamma_n} = t^{1-\theta} \gamma_n 
\leq \gamma_n.
\]
For \( \gamma_n > t^{-1} \) we have \( \alpha_n = 1 \) and \( \beta_n = 0 \) such that
\[
t^{-\theta} (\alpha_n + t_{\gamma_n} \beta_n) = t^{-\theta} \leq \gamma_n.
\]
Thus, in either case, we have the estimate
\[
\|\rho\|_{(\mathcal{F}_k^0(\mathcal{H}_1),\mathcal{F}_k^1(\mathcal{H}_1))_\theta} \leq \sum_{n=0}^{\infty} \langle n | \rho \rangle \| n \| (n + 1)^{(1-\theta)k_0 + \theta k_1}/2.
\]
This shows that for arbitrary density operators
\[
\|\rho\|_{(\mathcal{F}_k^0(\mathcal{H}_1),\mathcal{F}_k^1(\mathcal{H}_1))_\theta} \leq \|\rho\|_{(\mathcal{F}(1-\theta)k_0 + \theta k_1(\mathcal{H}_1))}.
\]
To extend the bound to a density operator \( \rho \) that is not diagonal in the Fock basis, and not only for the diagonal \( \rho_{\text{diag}}(t) \), we partition \( \rho \) as
\[
\rho = \begin{pmatrix} X_0(t) & \rho_{21}(t) \end{pmatrix} \begin{pmatrix} \rho_{21}(t) & X_1(t) \end{pmatrix}
\]
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
\[
T^{(k)} := S^{(k)} \rho S^{(k)} = \begin{pmatrix} S_1^{(k)}(t)X_0(t)S_1^{(k)}(t) & S_1^{(k)}(t)\rho_{21}(t)S_2^{(k)}(t) \\ S_1^{(k)}(t)\rho_{21}(t)S_2^{(k)}(t) & S_2^{(k)}(t)\rho_{22}(t)S_2^{(k)}(t) \end{pmatrix}.
\]
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
\[
\|S_1^{(k)}(t)\rho_{21}S_1^{(k)}(t)\|_1 \leq \frac{1}{2} \left( \|S_1^{(k)}(t)\rho_{11}S_1^{(k)}(t)\|_1 + \|S_1^{(k)}(t)\rho_{22}S_2^{(k)}(t)\|_1 \right).
\]
From here, we examine three cases separately:
Theorem 28. Let \(|T^{(k_1)}_{11}| \leq |T^{(k_1)}_{22}|_1\). In this case, we find from choosing \(X_0 := \rho_{21}\) and \(X_1 := 0\) in (54) that
\[
K(t, \rho_{21}) \leq \|\rho_{21}(t)\|_{\mathcal{F}^{k_{0,1}}(\mathcal{H}_1)} \leq \|X_0(t)\|_{\mathcal{F}^{k_{0,1}}(\mathcal{H}_1)}.
\]

- Case 2: \(|T^{(k_0)}_{22}|_1 \geq |T^{(k_0)}_{11}|_1\) In this case, we find from choosing \(X_0 := 0\) and \(X_1 := \rho_{21}\) in (54) that
\[
K(t, \rho_{21}) \leq t \|\rho_{21}\|_{\mathcal{F}^{k_{1,1}}(\mathcal{H}_1)} \leq t \|X_1(t)\|_{\mathcal{F}^{k_{1,1}}(\mathcal{H}_1)}.
\]

- Case 3: \(|T^{(k_0)}_{22}|_1 \leq |T^{(k_0)}_{11}|_1\) and \(|T^{(k_1)}_{22}|_1 \geq |T^{(k_1)}_{11}|_1\) In this case, we find from choosing \(X_0 = \rho_{21}/2\) and \(X_1 = \rho_{21}/2\) in (54) that
\[
K(t, \rho_{21}) \leq \frac{\|\rho_{21}\|_{\mathcal{F}^{k_{0,1}}(\mathcal{H}_1)} + t \|\rho_{21}\|_{\mathcal{F}^{k_{1,1}}(\mathcal{H}_1)}}{2}
\leq \frac{1}{2} \left(\|X_0(t)\|_{\mathcal{F}^{k_{0,1}}(\mathcal{H}_1)} + t \|X_1(t)\|_{\mathcal{F}^{k_{1,1}}(\mathcal{H}_1)}\right).
\]

Hence, we have altogether that
\[
K(t, \rho(t)) \leq K(t, \rho_{\text{diag}}(t)) + 2K(t, \rho_{21}(t))
\leq 3 \left(\|X_0(t)\|_{\mathcal{F}^{k_{0,1}}(\mathcal{H}_1)} + t \|X_1(t)\|_{\mathcal{F}^{k_{1,1}}(\mathcal{H}_1)}\right)
\]

which implies that
\[
\|\rho\|_{(\mathcal{F}^{k_{0,1}}(\mathcal{H}_1), \mathcal{F}^{k_{1,1}}(\mathcal{H}_1))_\theta} \leq 3\|\rho\|_{(\mathcal{F}^{(1-\theta)k_0 + \theta k_1,1}(\mathcal{H}_1))}.
\]

The result follows from the interpolation bound (B1).

\[\Box\]

Appendix C: Standard moments vs phase space moments: a partial converse

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

**Theorem 28.** Let \(\rho\) be an \(m\)-mode quantum state such that its characteristic function \(\chi_\rho\) is \(2k\) times totally differentiable at \(z = 0\) for some integer \(k\), then the \(2k\)th standard moment is finite as well.

**Proof.** For simplicity, we restrict attention to \(m = 1\). Let \(H^{+}_{\text{lin}} := a + a^\dagger\) and \(H^{-}_{\text{lin}} := (-i)(a - a^\dagger)\) 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 \(\mu_{e_i}\) such that
\[
\langle e_i | f(H^{+}_{\text{lin}}) | e_i \rangle = \int_{\sigma(H^{+}_{\text{lin}})} f(\lambda) \ d\mu_{e_i}(\lambda) \text{ for all } f \text{ bounded measurable.}
\]

We then define the new probability measure \(\mu_{\rho} := \sum_{i=1}^{\infty} \lambda_i \mu_{e_i}\) such that
\[
\text{Tr} \left[ \rho f(H^{+}_{\text{lin}}) \right] = \int_{\sigma(H^{+}_{\text{lin}})} f(\lambda) \ d\mu_{\rho}(\lambda) \text{ for all } f \text{ bounded measurable.}
\]
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 \( \psi : \mathbb{R} \to \mathbb{C} \) as

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

which is by assumption \( 2k \) times differentiable at zero and let \( u(t) = \Re \psi(t) \). Then, \( u \) is also \( 2k \) times differentiable at zero. Since \( \psi^{2k}(0) \) exists, for \( t \in (-\varepsilon, \varepsilon) \), with sufficiently small \( \varepsilon > 0 \), the function \( t \mapsto \psi^{(2k-1)}(t) \) exists and is continuous.

We record that Taylor’s formula implies that for \( t \in (-\varepsilon, \varepsilon) \)

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

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

\[
\text{Tr} \left[ \rho f_k(t H_{\text{lin}}^{\pm}) H_{\text{lin}}^{\pm 2k} \right] = \int_{\sigma(H_{\text{lin}}^{\pm})} f_k(t \lambda) \lambda^{2k} d\mu_\psi(\lambda)
\]

\[
= \frac{(-1)^k (2k)!}{t^{2k}} \left( u(t) - \sum_{i=0}^{k-1} (-1)^i \frac{t^{2i}}{(2i)!} \text{Tr} \left[ \rho (H_{\text{lin}}^{\pm})^{2i} \right] \right)
\]

\[
\leq 2k \sup_{\theta \in (0,1)} \frac{|u^{2k-1}(\theta t)|}{|\theta t|} =: g_k(t),
\]

(C1)

Then, we have from Fatou’s lemma

\[
\text{Tr} \left[ \rho \left( H_{\text{lin}}^{\pm} \right)^{2k} \right] = \text{Tr} \left[ \rho f_k(0) \left( H_{\text{lin}}^{\pm} \right)^{2k} \right]
\]

\[
\leq \liminf_{t \to 0} \int_{\sigma(H_{\text{lin}}^{\pm})} f_k(t \lambda) \lambda^{2k} d\mu_\psi(\lambda)
\]

\[
= \liminf_{t \to 0} g_k(t) = 2k |u^{2k}(0)| < \infty.
\]

(C2)

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

[1] W. Feller. An introduction to probability theory and its applications, volume II. John Wiley & Sons, Inc., New York-London-Sydney, 2nd edition, 1971.
[2] C.D. Cushen and R.L. Hudson. A quantum-mechanical Central Limit Theorem. J. Appl. Probab., 8(3):454–469, 1971.
[3] K. Hepp and E.H. Lieb. Phase-transitions in reservoir-driven open systems with applications to lasers and superconductors. Helv. Phys. Acta, 46(5):573–603, 1974.
[4] K. Hepp and E.H. Lieb. On the superradiant phase transition for molecules in a quantized radiation field: the Dicke Maser model. Ann. Phys., 76(2):360–404, 1973.
[5] N. Giri and W. von Waldenfels. An algebraic version of the Central Limit Theorem. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 42(2):129–134, 1978.

[6] D. Goderis and P. Vets. Central Limit Theorem for mixing quantum systems and the CCR-algebra of fluctuations. *Commun. Math. Phys.*, 122(2):249–265, 1989.

[7] D. Goderis, A. Verbeure, and P. Vets. About the mathematical theory of quantum fluctuations. Leuven University Press, Leuven, Belgium, 1989.

[8] T. Matsui. Bosonic Central Limit Theorem for the one-dimensional XY model. *Rev. Math. Phys.*, 14(07n08):675–700, 2002.

[9] M. Cramer and J. Eisert. A quantum Central Limit Theorem for non-equilibrium systems: exact local relaxation of correlated states. *New J. Phys.*, 12(5):055020, 2010.

[10] D. Goderis, A. Verbeure, and P. Vets. About the exactness of the linear response theory. *Commun. Math. Phys.*, 136(2):265–283, 1991.

[11] V. Jakšić, Y. Pautrat, and C.-A. Pillet. Central Limit Theorem for locally interacting Fermi gas. *Commun. Math. Phys.*, 285(1):175–217, 2009.

[12] G.B. Arous, K. Kirkpatrick, and B. Schlein. A central limit theorem in many-body quantum dynamics. *Commun. Math. Phys.*, 321(2):371–417, 2013.

[13] J. Dereziński. Boson free fields as a limit of fields of a more general type. *Rep. Math. Phys.*, 21(3):405–417, 1985.

[14] R.F. Streater. Entropy and the Central Limit Theorem in quantum mechanics. *J. Phys. A*, 20(13):4321–4330, 1987.

[15] M. Hayashi. Quantum estimation and the Quantum Central Limit theorem. *American Mathematical Society Translations Series*, 2(227):95–123, 2009.

[16] M. Hayashi. *Quantum Information: An Introduction*. Springer Berlin Heidelberg, 2006.

[17] R. Lenczewski. Quantum Central Limit Theorems, pages 299–314. Springer US, Boston, MA, 1995.

[18] E.B. Davies. Quantum stochastic processes. *Commun. Math. Phys.*, 15(4):277–304, 1969.

[19] L. Accardi and Y.G. Lu. Quantum central limit theorems for weakly dependent maps. II. *Acta Math. Hungar.*, 63(3):249–282, 1994.

[20] M. Rosati, A. Mari, and V. Giovannetti. Narrow bounds for the quantum capacity of thermal attenuators. *Nat. Commun.*, 9(1):4339, 2018.

[21] A. Winter. Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. *Commun. Math. Phys.*, 347(1):291–313, 2016.
[34] H. Umegaki. Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Math. Sem. Rep.*, 14(2):59–85, 1962.
[35] G. Lindblad. Entropy, information and quantum measurements. *Commun. Math. Phys.*, 33(4):305–322, 1973.
[36] A. S. Holevo. Entanglement-assisted capacities of constrained quantum channels. *Theory Probab. Appl.*, 48(2):243–255, 2004.
[37] A.S. Holevo and M.E. Shirokov. On classical capacities of infinite-dimensional quantum channels. *Problems Inform. Transmission*, 49(1):15–31, 2013.
[38] M.A. de Gosson. *Symplectic Geometry and Quantum Mechanics*. Operator Theory: Advances and Applications. Birkhäuser Basel, 2006.
[39] A. Serafini. *Quantum Continuous Variables: A Primer of Theoretical Methods*. CRC Press, Taylor & Francis Group, 2017.
[40] A.S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. Publications of the Scuola Normale Superiore. Scuola Normale Superiore, 2011.
[41] S. Barnett and P.M. Radmore. *Methods in Theoretical Quantum Optics*. Oxford Series in Optical and Imaging Sciences. Clarendon Press, 2002.
[42] A. Zygmund. A remark on characteristic functions. *Ann. Math. Statist.*, 18(2):272–276, 06 1947.
[43] N.G. Ushakov. *Selected Topics in Characteristic Functions*. Modern Probability and Statistics. de Gruyter, 2011.
[44] R. König and G. Smith. The entropy power inequality for quantum systems. *IEEE Trans. Inf. Theory*, 60(3):1536–1548, 2014.
[45] J. Bergh and J. Lőström. *Interpolation spaces: an introduction*, volume 223. Springer Science & Business Media, 2012.
[46] G.B. Folland. *Harmonic Analysis in Phase Space*. Princeton University Press, 1989.
[47] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 5 edition, 2019.
[48] A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inf. Theory*, 45(7):2481–2485, 1999.
[49] J. Williamson. On the algebraic problem concerning the normal forms of linear dynamical systems. *Am. J. Math.*, 58(1):141–163, 1936.
[50] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. Dover Publications, 1965.
[51] J.W. Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.*, 15(1):211–225, 1922.
[52] W. Bryc. A remark on the connection between the large deviation principle and the central limit theorem. *Statist. Probab. Lett.*, 18(4):253–256, 1993.
[53] R.N. Bhattacharya and R.R. Rao. *Normal approximation and asymptotic expansions*, volume 64. SIAM, 1986.
[54] J.-C. Bourin and E.-Y. Lee. Unitary orbits of hermitian operators with convex or concave functions. *Bulletin of the London Mathematical Society*, 44(6):1085–1102, 2012.