A stable quantum Darmois-Skitovich theorem

Cite as: J. Math. Phys. 61, 022201 (2020); https://doi.org/10.1063/1.5122955
Submitted: 01 August 2019 . Accepted: 13 January 2020 . Published Online: 03 February 2020

Javier Cuesta

ARTICLES YOU MAY BE INTERESTED IN

Locality and renormalization: Universal properties and integrals on trees
Journal of Mathematical Physics 61, 022301 (2020); https://doi.org/10.1063/1.5116381

Degenerations of Filippov algebras
Journal of Mathematical Physics 61, 021701 (2020); https://doi.org/10.1063/1.5119393

Mass splitting in the Thomas–Fermi–Dirac–von Weizsäcker model with background potential
Journal of Mathematical Physics 61, 021502 (2020); https://doi.org/10.1063/1.5130565
A stable quantum Darmois-Skitovich theorem

Cite as: J. Math. Phys. 61, 022201 (2020); doi: 10.1063/1.5122955
Submitted: 1 August 2019 • Accepted: 13 January 2020 •
Published Online: 3 February 2020

Javier Cuesta

AFFILIATIONS
Department of Mathematics, Technische Universität München, 85748 Garching, Germany and Munich Center for Quantum Science and Technology (MCQST), Munich, Germany

Author to whom correspondence should be addressed: j.cuesta@tum.de

ABSTRACT

The Darmois-Skitovich theorem is a simple characterization of the normal distribution in terms of the independence of linear forms. We present here a non-commutative version of this theorem in the context of Gaussian bosonic states and show that this theorem is strongly stable under small errors in its underlying conditions. An explicit estimate of the stability constants which depend on the physical parameters of the problem is given.

I. INTRODUCTION

Among all characterizations of the normal distribution, the ones concerning the independence of linear forms stand out because of their simplicity. The landmark result of such classical characterizations is due to Darmois and Skitovich. Their theorem is a generalization to \( n \)-random variables and arbitrary coefficients of the following fact: if \( X, Y \) are independent real-valued random variables with \( X + Y \) and \( X - Y \) independent, then \( X \) and \( Y \) are normally distributed with the same variance (see Theorem 4). We will be interested in studying a quantum (read as non-commutative) version of the Darmois-Skitovich (DS) theorem, which we now write shortly as the DS theorem. In this case, the role of the normal distribution is taken by Gaussian bosonic states: quantum states whose statistics are completely determined by the knowledge of the first and second moments and whose canonical observables obey the bosonic commutation relations.

Particularly noteworthy is the fact that the quantum DS theorem has a clear physical realization. Consider an arbitrary product state that passes through a beam splitter as in Fig. 1. If the output state of the beam splitter is also a product state, then the input states are Gaussian bosonic states with the same second moments. This is the content of the quantum DS-theorem. Mathematically, the content of the quantum DS theorem is given on Theorem 7.

That there does not exist any two copies of identical non-Gaussian states fulfilling this is by no means trivial since the action of a beam splitter does not create second-moment cross correlations for an identical product of quantum states (Gaussian and non-Gaussian). This operational characterization of Gaussian states was already known, however, without a direct reference to the DS theorem. We show this characterization for a general \( n \)-mode Gaussian bosonic state by means of the DS theorem. This has the advantage of a much clearer statistical interpretation and a simpler proof. Additionally, we show that a beam splitter is the only non-trivial linear operation that can have a factorizable output for all identical input states (Lemma 6). The latter places the beam splitter as the basic element for detecting non-Gaussianity.

Of course in real life, we cannot completely guarantee that two states are totally independent. Therefore, it is crucial to study how stable the DS theorem is. This means how does the conclusion of the quantum DS theorem change, when we assume that the output state is not exactly a product, but is approximately close to a product state. Our main result is a proof of the stability of the quantum DS theorem for quantum states whose all statistical moments in position and momentum, including mixed moments, exist and are finite. Such states are described by the set of Schwartz density operators. For independent input states whose output from a beam splitter is close in trace norm to a product state, we show that they are close in Hilbert-Schmidt (HS) norm to their respective Gaussian counterpart (i.e., the Gaussian state which has the same first and second moments). Moreover, the corresponding second moments of the input states have to be approximately close as well. A precise mathematical statement of this result is given in Theorem 9. We make an effort to present explicit...
FIG. 1. Quantum Darmois-Skitovich theorem: let $U_S$ be the unitary operator corresponding to the action of a non-trivial beam splitter transformation. The output state $\rho_{ab} := U_S(\rho_1 \otimes \rho_2)U_S^*$ is a product state if and only if $\rho_1$ and $\rho_2$ are Gaussian bosonic states with the same moments.

A. Notation and preliminaries

We will be entirely concerned with continuous-variable systems with a discrete number $n$ of modes. We denote by

$$R := (Q_1, P_1, \ldots, Q_n, P_n),$$ (1)

the vector of canonical operators for a quantum system and $R_k, k = 1, \ldots, 2n$ its components. Here, $Q_l, P_l, l = 1, \ldots, n$ act on the $l$-tensor factor of the Fock space $\mathcal{H} = \bigotimes_{l=1}^n L^2(\mathbb{R})$, where $L^2(\mathbb{R})$ denotes the space of Lebesgue square integrable functions on $\mathbb{R}$. The canonical commutation relations (CCR) are defined by

$$[R_k, R_l] = i\sigma_{kl},$$ (2)

where $\sigma_{kl}$ are the entries of the symplectic matrix

$$\sigma = \bigoplus_{i=1}^n \omega$$ with $$\omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$ (3)

We frequently use the shorthand notation $R_\xi := \xi \cdot \sigma R$, $\xi \in \mathbb{R}^{2n}$. The phase-space description of a quantum state $\rho$ is determined by the characteristic function $\chi : \mathbb{R}^{2n} \to \mathbb{C}$ defined by

$$\chi(\xi) := \text{Tr}[W_\xi \rho],$$ (4)

where $W_\xi = e^{i\xi \cdot \sigma R}$ is the so-called Weyl operator. The CCR are encoded in the Weyl relation

$$[R_\xi, R_\eta] = i\omega_{\xi \eta},$$ (5)

where $\omega_{\xi \eta}$ are the entries of the symplectic matrix

$$\omega = \sigma = \bigoplus_{i=1}^n \omega$$ with $$\omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$ (6)

These explicit constants contain information of how the problem can become unstable and have not been estimated before in either the classical or quantum case. The robustness of the quantum DS theorem depends on the transmittivity of the beam splitter, the number of modes, and the largest fourth moment of the output state.

The layout of the paper is as follows: In Subsection I A, we give the basic definitions and results in continuous-variable quantum information that will be used. Section II contains the main result. We give a simple proof for the characterization of Gaussian bosonic states using the DS theorem and then proceed to state the stability of the DS theorem. Section II B introduces and summarizes the main properties of Schwartz operators. In Sec. III C, we give a full proof of the stability of the DS theorem. Finally, in Sec. III D, we show some auxiliary lemmas, and in Appendices A and B, some explicit bounds are rigorously computed using the properties of Schwartz operators. These bounds will be used for the estimate of the constants appearing in the stability of DS Theorem 9.
The name of the characteristic function for the map in Eq. (1) comes from an analogy with the classical characteristic function that is the Fourier transform of a probability distribution. In fact by taking a fixed direction in phase space, we recover the classical characteristic function, and from there, we can “import” all the known results of the classical world. This indeed will be used in order to give a simple proof of the characterization of Gaussian bosonic states. The condition for a function \( \chi : \mathbb{R}^{2n} \to \mathbb{C} \) to be a bona-fide quantum characteristic function is the property of sigma-positiveness. For clarity, we state these results and refer the reader to Ref. 8, Sec. 5.4 for a proof.

**Theorem 1** (Quantum Bochner-Khinchin). For \( \chi : \mathbb{R}^{2n} \to \mathbb{C} \) to be a characteristic function of a quantum state, the following conditions are necessary and sufficient:
1. \( \chi(0) = 1 \) and \( \chi \) is continuous at \( \xi = 0 \), and
2. \( \chi \) is \( \sigma \)-positive definite, i.e., for any \( m \in \mathbb{N}, \) any set \( \{\xi_1, \xi_2, \ldots, \xi_m\} \) of vectors in \( \mathbb{R}^{2n} \), and any set \( \{c_1, c_2, \ldots, c_m\} \) of complex numbers

\[
\sum_{k,l=1}^{m} c_k \overline{c_l} \chi(\xi_k - \xi_l) e^{\frac{i}{2}(\xi_k - \xi_l) \cdot d} \geq 0. 
\]

**Corollary 2** (Classical Marginals). Let \( \chi(\xi) \) be the characteristic function of a quantum state. Then, for every fixed \( \xi \in \mathbb{R}^{2n} \), the function \( \mathbb{R} \ni t \mapsto \chi(t\xi) \) is a classical characteristic function, i.e., the Fourier transform of a classical probability distribution.

As in the classical case, the characteristic function is a moment generating function. The displacement vector is defined by the entries \( d_\xi := \text{Tr}[\rho R_\xi] \), and we say that the state is centered if \( d = 0 \). The covariance matrix (CM) is defined by the matrix entries \( \Gamma_{kl} = \text{Tr}[\rho (R_k - d_k)(R_l - d_l)] \). In order that \( \Gamma \) corresponds to a genuine quantum CM, the CCR impose the further condition \( \Gamma + i\sigma \geq 0 \), which is nothing but the uncertainty principle expressed in a coordinate-free form.

A Gaussian bosonic state is defined as a state with a Gaussian characteristic function

\[
\chi(\xi) = \exp \left[ -\frac{\xi \cdot F \xi}{4} + i\xi \cdot d \right].
\]

We write \( \chi_\rho \) to emphasize that \( \chi \) is the characteristic function of the state \( \rho \). We denote by \( M(2n, \mathbb{R}) \) and \( Sp(4n, \mathbb{R}) \) the set of \( 2n \times 2n \) matrices with real entries and the group of \( 4n \times 4n \) symplectic matrices with real entries, respectively. The latter is defined as the group of matrices \( S \in M(2n, \mathbb{R}) \) such that \( \sigma S^T = \sigma \).

Unitary Gaussian operations, i.e., unitary evolutions coming from quadratic Hamiltonians in \( P \) and \( Q \), are described by symplectic transformations. These operations have the property that

\[
\chi(U_{U_2}U_1^\dagger) \xi = \chi_\rho (S^T \xi),
\]

where \( U_2 \) is a unitary operation associated with the symplectic transformation \( S \) (strictly speaking \( U_2 \) is determined up to a phase; however, this ambiguity disappears in the conjugation \( U_2 \cdot U_5^* \)). The unitary evolution of a Gaussian state is completely determined by the new displacement vector \( d' = Sd \) and CM, \( \Gamma' = S\Gamma S^T \).

A one mode non-trivial beam splitter transformation is the one corresponding to the symplectic transformation,

\[
S = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad \theta \equiv m\pi/2, \quad m = 0, 1, 2, \ldots.
\]

It corresponds to a unitary evolution where the Hamiltonian is

\[
H = \frac{\theta}{4}(a_1^*a_2 + a_2^*a_1),
\]

with \( a_j = (Q_j + iP_j)/\sqrt{2}, a_j^* = (Q_j - iP_j)/\sqrt{2}, j = 1, 2 \) being the creation and annihilation operators. A local transformation acts in each separated mode and corresponds therefore to transformations that can be written as \( S = \bigotimes_{k=1}^{\infty} S_k \). In the context of quantum optics, examples of local transformations are phase-shifts and one-mode squeezing transformations.
The Wigner phase space distribution is defined to be the (symplectic) Fourier transform of the characteristic function

$$W(\eta) = \frac{1}{(2\pi)^n} \int e^{i \eta \cdot \xi} \rho(\xi) \, d\xi.$$  \hspace{1cm} (9)

Its importance lies in the fact that, due to Corollary 2, all one-dimensional marginals are positive distributions in phase space, which can be associated with the usual probability distributions, e.g., on position and momentum of a state \( \rho \).

We write \( A^* \) for the adjoint operator of \( A \) and \( \| \cdot \| \) for the uniform norm. The trace norm is defined as \( \| A \|_1 = \text{Tr} \sqrt{A^* A} \) and the Hilbert-Schmidt (HS) norm \( \| A \|_2 = (\text{Tr}[A^* A])^{1/2} \). We have the order \( \| \cdot \|_1 \leq \| \cdot \|_2 \leq \| \cdot \|_1 \). These norms are in fact unitarily invariant and \( \| A^* \|_p = \| A \|_p \) for \( p = 1, 2 \). The usual norm in \( L^2(\mathbb{R}^{2n}) \) will be denoted by \( \| \cdot \|_{L^2(\mathbb{R}^{2n})} \). We use sometimes the Dirac notation for a vector \( |\psi\rangle \in \mathcal{H} \) and the inner product notation \( \langle \phi | \psi \rangle \). The commutator and anticommutator are written as \([ , ]\) and \(\{ , \}\) respectively. The space of bounded operators on the Hilbert space \( \mathcal{H} \) is denoted by \( \mathcal{B}(\mathcal{H}) \).

The inverse relation of Eq. (4) is called the Weyl transform

$$T = \frac{1}{(2\pi)^n} \int \text{Tr}[W_\xi T] W_{-\xi} \, d\xi,$$  \hspace{1cm} (10)

where the integral converges weakly for any Hilbert-Schmidt operator \( T \). This is a consequence of the quantum Parseval theorem, which due to its importance, we state here.

**Theorem 3** (Quantum Parseval relation). Let \( \{ W_\xi \} \) be a strongly continuous and irreducible Weyl systems acting on the Hilbert space \( \mathcal{H} \) with respective phase space \( X = \mathbb{R}^{2n} \). Then, \( T \mapsto \text{Tr}[W_\xi T] \) extends uniquely to an isometric map from the Hilbert space of Hilbert-Schmidt class operators on \( \mathcal{H} \) onto \( L^2(X) \) such that

$$\text{Tr} T_1^* T_2 = \frac{1}{(2\pi)^n} \int \text{Tr}[W_\xi T_1^*] \text{Tr}[W_\xi T_2] \, d\xi.$$  \hspace{1cm} (11)

This theorem also implies that Eq. (4) is also valid for \( T \) Hilbert-Schmidt. The map \( \xi \mapsto \text{Tr} W_\xi T \) is called the inverse Weyl transform of \( T \); being the characteristic function, the special case \( T \) is a density operator.

We will be using repeatedly the following trace inequalities: If \( B \) is a bounded operator and \( T \) a trace-class operator, then a particular case of Hölder’s inequality states

$$\text{Tr} BT \leq \| B \| \| T \|_1.$$  

Let \( T_1, T_2 \) be two Hilbert-Schmidt operators. The trace operator version of the Cauchy-Schwarz inequality is

$$\text{Tr} T_1 T_2 \leq \| T_1 \|_1 \| T_2 \|_2.$$  

While it is true that the Hilbert-Schmidt norm is often used out of convenience, it also has operational interpretation, making it preferable for some tasks. This includes equality testing and state discrimination with fixed or random measurements. Furthermore, the Hilbert-Schmidt norm can be a good measure to quantify the difference between two quantum states in quantum optics. There, the Wigner functions of infinite dimensional quantum states are accessible by means of tomography. The difference between two quantum states is quantified by the Hilbert-Schmidt distance of the respective Wigner functions.

**II. MAIN RESULT**

In Subsection II A, we present the quantum version of the DS theorem and our main stability result. The detailed proof of the stability of the DS theorem is presented in Sec. III C.

**A. Quantum Darmois-Skitovich theorem**

We are interested in a quantum analog of the following theorem:

**Theorem 4** (Darmois-Skitovich). Let \( X_1, \ldots, X_n \) (\( n \geq 2 \)) be independent random variables and \( a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R} \setminus \{0\} \). If the two linear forms

$$Y_1 = \sum_i a_i X_i \quad \text{and} \quad Y_2 = \sum_i b_i X_i$$

are independent, \( X_1 \) is normally distributed.
Different proofs and the history of the classical DS theorem can be found in p. 78 in Ref. 10 and in Ref. 11. Our setup for the quantum version is the following. We consider two $n$-mode quantum states $\rho_1, \rho_2 \in \mathbb{B}(\mathcal{H})$ with respective canonical operators

$$
R_1 = (Q_1, P_1, \ldots, Q_n, P_n) \\
R_2 = (Q_{n+1}, P_{n+1}, \ldots, Q_{2n}, P_{2n})
$$

(13)

and write $R = (R_1, R_2)$. We assume that $\rho_1$ and $\rho_2$ are independent so that their state in $\mathbb{B}(\mathcal{H} \otimes \mathcal{H})$ is a product state $\rho_1 \otimes \rho_2$. We refer $\rho_1$ and $\rho_2$ as input states.

The action of producing linear forms of random variables can be mimicked by (Gaussian) unitary evolutions $U_\mathcal{S}$. These unitary evolutions are generated by Hamiltonians that are quadratic expressions in the canonical operators. Moreover, the unitary evolution $U_\mathcal{S} \in \mathbb{B}(\mathcal{H} \otimes \mathcal{H})$ is associated with a symplectic transformation $S \in \text{Sp}(4n, \mathbb{R})$. In other words, the linear transformation

$$
R \mapsto SR,
$$

(14)

corresponds to a unitary evolution $\rho \mapsto U_\mathcal{S}U_\mathcal{S}^*$. In order to obtain an analog of Eq. (12), we classify the set of unitaries $U_\mathcal{S}$ that produce a bipartite independent output, i.e., such that

$$
U_\mathcal{S}(\rho_1 \otimes \rho_2)U_\mathcal{S}^* = \rho_1 \otimes \rho_2,
$$

(15)

where $\rho_1, \rho_2 \in \mathbb{B}(\mathcal{H})$ are $n$-mode quantum states. This is equivalent to classifying the respective set of symplectic transformation for which Eq. (15) holds. If we are to expect that $U_\mathcal{S}$ preserves independence, the transformation $S$ should at least preserve uncorrelated inputs (a generally weaker condition than independence that only deals with the second moments). Furthermore, acting locally on each state and swapping them are trivial operations which preserve independence for arbitrary states. Hence, we need to consider other operations in order to obtain a meaningful statement for the quantum DS theorem. The following two lemmas show in fact that there is only one non-trivial symplectic transformation for our setup.

**Lemma 5.** Let $S \in \text{Sp}(4n, \mathbb{R})$ be such that

$$
S \begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix} S^T = \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix}
$$

for all CM $\Gamma_1, \Gamma_2 \in \mathbb{R}^{2n \times 2n}$.

Here, $*, \cdot$ denote any CM $\in \mathbb{R}^{2n \times 2n}$. Then, $S$ is either of the form $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ or $\begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}$ with $A, B, C, D \in \text{Sp}(2n, \mathbb{R})$.

If we consider identical inputs, we obtain:

**Lemma 6.** Let $S \in \text{Sp}(4n, \mathbb{R})$ be such that

$$
S \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma \end{pmatrix} S^T = \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix}
$$

for all CM $\Gamma \in \mathbb{R}^{2n \times 2n}$.

(16)

Here, $*, \cdot$ denote any CM $\in \mathbb{R}^{2n \times 2n}$. Then,

$$
S = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1 + \alpha^2} \end{pmatrix} \begin{pmatrix} 1_{2n} & \alpha 1_{2n} \\ -\alpha 1_{2n} & 1_{2n} \end{pmatrix} = \begin{pmatrix} 0 & \alpha X \\ \alpha Y & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1 + \gamma^2} \end{pmatrix} \begin{pmatrix} 1_{2n} & \gamma 1_{2n} \\ -\gamma 1_{2n} & 1_{2n} \end{pmatrix}
$$

(17)

where $X, Y \in \text{Sp}(2n, \mathbb{R})$ and $\alpha, \gamma \in \mathbb{R} \cup \{ \pm \infty \}$.

Thus, the only non-trivial linear transformation in Eq. (14) is of the form of Eq. (17). We discard the trivial operations and set $\alpha = \tan \theta$ in Eq. (17),

$$
S_\theta = \begin{pmatrix} \cos \theta 1_{2n} & -\sin \theta 1_{2n} \\ \sin \theta 1_{2n} & \cos \theta 1_{2n} \end{pmatrix},
$$

(18)

which is the symplectic transformation associated with a $(n$-mode) beam splitter operation. We refer to the latter operation as a *non-trivial beam splitter* transformation.
Although for every covariance matrix $\Gamma$, $S_b(\Gamma \otimes \Gamma)S_b^\dagger = (\Gamma \otimes \Gamma)$, it turns out from the DS theorem that there does not exist a non-Gaussian state $\rho$ such that $U_b(\rho \otimes \rho)U_b^\dagger = \rho \otimes \rho$. See Fig. 1.

**Theorem 7** (Quantum Darmois-Skitovich). Let $U_b$ be the unitary operation corresponding to a non-trivial beam splitter Eq. (18). Consider the state $\rho_{ab} = U_b(\rho_1 \otimes \rho_2)U_b^\dagger$ obtained after the unitary evolution of an arbitrary product state. If the output state is a product state $\rho_{ab} = \rho_a \otimes \rho_b$, then $\rho_1$ and $\rho_2$ are Gaussian bosonic states with the same CM but not necessarily same displacement vector.

Due to the 1–1 correspondence between quantum states and their characteristic function, we have the following consequence:

**Corollary 8.** Let $\chi_1$ and $\chi_2$ be the characteristic function of the quantum states $\rho_1$ and $\rho_2$, respectively, which have finite second moments. If for a fixed $\theta = m\pi/2$, $m = 0, 1, 2, \ldots$, the characteristic functions satisfy the functional equation

$$\chi_1(\cos \theta \xi + \sin \theta \xi)\chi_2(\cos \theta \xi - \sin \theta \xi) = \chi_1(\cos \theta \xi_1)\chi_1(\sin \theta \xi_2)\chi_2(\cos \theta \xi_2)\chi_2(- \sin \theta \xi_1),$$

(19)

for all $\xi_1, \xi_2 \in \mathbb{R}^n$, then $\rho_1$ and $\rho_2$ are Gaussian bosonic states with the same CM but not necessarily same displacement vector.

An immediate proof of the quantum DS theorem can be obtained from its corresponding classical result and Corollary 2. We recall that the latter corollary tells us that we always obtain a positive Wigner function (a classical probability distribution) from the quantum characteristic function whenever we move through a fixed direction in phase space.

**Proof of Theorem 7.** Using Eq. (8), the evolution of the input states can be expressed in terms of characteristic functions as Eq. (19). We fixed the direction $\xi_1 = \xi_2 = \xi$ and parametrize $\xi_1 = \Gamma \xi$, $\xi_2 = \xi$ with $t, s \in \mathbb{R}$. Moreover, we introduce the classical characteristic functions $\chi_u(u) := \chi(u \xi), j = 1, 2$ with $u \in \mathbb{R}$ so that Eq. (19) reads

$$\chi_1(\cos \theta t + \sin \theta t)\chi_2(\cos \theta t - \sin \theta t) = \chi_1(\cos \theta s)\chi_1(\sin \theta s)\chi_2(\cos \theta s)\chi_2(- \sin \theta s).$$

This last equation is the functional version of the classical DS theorem [cf. Eq. (8.7) of Sec. XV.9 in Ref. 10]. From this classical result, it follows that $\chi_1$ and $\chi_2$ are one dimensional Gaussian characteristic functions with the same variance. We can compute the moments by taking derivates of the characteristic function to obtain that

$$\chi_1(t \xi) = \exp \left[ -\frac{t^2}{4} (\xi \cdot \Gamma \xi) + it \cdot \xi \cdot d \right], \quad j = 1, 2.$$

The result follows with $t = 1$.

**B. Stability**

In Sec. II A, we presented an exact version of the DS theorem that brings naturally an operational characterization of Gaussian bosonic states. There, exact factorizability of the output state is assumed. In practice, it is impossible to assure such thesis since there are always errors in the measurements, and therefore, the validity of the result is not completely clear in real life. Moreover, any practical application can immediately be ruled out if the conclusion is not robust against small changes in the defined conditions.

We are interested in finding to which extent the results of Theorem 7 are affected if the main assumption is not exact but approximately satisfied. It is not always the case that characterizations of the normal distribution are generally stable. Cramer’s characterization of the normal distribution states that if the sum $X + Y$ of two random independent random variables $X$ and $Y$ has a normal distribution, then necessarily both $X$ and $Y$ are normal. It turns out that the classical theorem of Cramer is only stable in a weak sense and that it fails to be robust for stronger notions of distance such as the entropic distance or the total variation norm. We are interested in finding to which extent the results of Theorem 7 are affected if the main assumption is not exact but approximately satisfied.

Before specifying the conditions of the stability of the quantum DS theorem, we briefly comment on the respective classical stability problem. The stability of the classical DS theorem is due to Gabovich. In his work, the word “approximately” is quantified in terms of the closeness of the cumulative distribution functions of the random variables. It is shown (Theorem 3 in Ref. 14) that for approximate independency, the considered classical probability distributions are $c(\log \log(1/\varepsilon))_{12}$-close to a normal distribution in the Levy metric. In Gabovich’s estimate, there is no explicit value of the constant $c$, and therefore, it is not known how it depends on the coefficients of the linear forms, neither on the number of random variables involved. We address this in our stability proof as it contains relevant information of the physical problem such as where instability could arise.

At the level of density operators, it was proven that weak operator topology is equivalent to trace-norm topology. Therefore, the Gabovich result and Corollary 2 should imply at least in a qualitative manner that the quantum DS theorem is stable. However, we can obtain a better concrete estimate of the stability of the DS theorem by considering not just the marginals and the classical result but rather using the entire phase space. The intuition that this could work in the quantum world lies on the natural restrictions that the quantum characteristic
functions have. Namely, due to the sigma-positiveness of the quantum characteristic function, which is essentially a form of Heisenberg’s uncertainty principle, we expect the construction of functions that are close point-wise to Gaussians but not in $L^2$-distance, to be ruled out. Quantum characteristic functions are always square-integrable, and, in particular, there are no quantum characteristic functions with compact support. With no more preambles, we give a precise statement of our result.

We say that a quantum state $\rho_{ab}$ is an $\epsilon$-approximate product state if there is a product state $\rho_a \otimes \rho_b$ such that
\[
\|\rho_{ab} - \rho_a \otimes \rho_b\|_1 \leq \epsilon.
\] (20)

Suppose that two independent states evolve according to the action of a beam splitter, but this time the output state is an $\epsilon$-approximate product state (see Fig. 2). The following theorem describes the robustness of the quantum DS theorem:

**Theorem 9** (Stability DS). Let $U_S$ be the unitary operation corresponding to a non-trivial beam splitter characterized in Eq. (18) and $\rho_1$, $\rho_2$ density operators of two $n$-mode systems with finite moments of all orders and whose CMs are $\Gamma_1$ and $\Gamma_2$. Consider the output state $\rho_{ab} = U_S(\rho_1 \otimes \rho_2) U_S^\dagger$ and define by $\rho'_1$, $\rho'_2$ Gaussian bosonic states that have the same CMs and displacement vectors as $\rho_1$, $\rho_2$. If for sufficiently small $\epsilon \in (0, 1)$, the output state $\rho_{ab}$ is $\epsilon$-close to a product state in trace norm, then
\[
\|\rho_j - \rho'_j\|_2 \leq c_1 \epsilon^{1/3} + \frac{c_2}{\sqrt{\log(1/\epsilon)}}, \quad j = 1, 2,
\] (21)
\[
\|\Gamma_1 - \Gamma_2\|_2 \leq c_3 \epsilon^{1/2},
\] (22)
where the constants $c_1$, $c_2$ and $c_3$ (which are made explicit in the proof) depend on the transmittivity $\theta$ of the beam splitter, the number of modes $n$, and the second and fourth absolute moments of the output state $\rho_{ab}$.

Note that under the conditions of Theorem 9, the bound of Eq. (21) also applies to the distance between the Wigner functions of $\rho$ and $\rho'$.

**1. Schwartz operators**

We begin this section by reviewing some important facts about Schwartz operators that are the non-commutative analog of Schwartz functions. The latter are infinite differentiable functions whose derivatives decay faster than any polynomial at infinity. The introduction of Schwartz operators allows us to handle differentiability and boundness problems in an elegant manner, and it is for this reason that they play an important technical role in this paper. This class of operators was first introduced in Ref. 4, and the reader may find there a detailed exposition.
In our proof of the stability of the DS theorem, we deal with terms of the form $\text{Tr}[R_i R \rho R_j R]$ that are \textit{a priori} not necessarily well-defined on any dense domain of $\rho$; it may happen that $\rho$ maps outside the domain of $R_i$. This is a common issue among others when dealing with unbounded operators (see Sec. 17.2.1 in Ref. 17 for some misleading formal manipulations with unbounded operators). There is an immense advantage when working with Schwartz operators since many regular properties for bounded operators become available for unbounded ones (e.g., “cycling under the trace” property holds). Indeed, Theorem 13 below plays a decisive role in the calculation of the constants appearing in our proof of the stability of DS theorem.

**Definition 10** (Schwartz operators). An operator $T \in B(H)$ is called a Schwartz operator if

$$\left\|P^\alpha Q^\eta T P^{\alpha'} Q^{\eta'}\right\|_1 < \infty \quad \text{for all } \alpha, \alpha', \beta, \beta' \in I_n,$$

where $I_n := \{\alpha = (a_1, \ldots, a_n) | a_i \in \mathbb{N} \cup \{0\} \text{ for all } i = 1, \ldots, n\}$ is the set of multi-indices, and

$$Q^\alpha = Q_1^{\alpha_1} \cdots Q_n^{\alpha_n}, \quad P^\alpha = P_1^{\alpha_1} \cdots P_n^{\alpha_n}.$$  \hspace{1cm} (23)

The set of all Schwartz operators will be denoted by $\mathcal{S}(H)$.

Hence, for Schwartz-density operators, all the statistical moments in $Q$ and $P$ exist and are finite. We denote by

$$\mathcal{S}(H) := \{\rho \in \mathcal{S}(H) \mid \rho \text{ is a density operator}\}$$

the space of density operators that are also Schwartz operators. For $\rho \in \mathcal{S}(H)$, we have the following neat characterization

**Proposition 11.** Let $T$ be a Hilbert-Schmidt operator. Then, $T$ is a Schwartz operator if and only if the respective Weyl transform is a Schwartz function.

**Corollary 12.** A density operator $\rho$ is a Schwartz operator if and only if its characteristic function $\chi$, or Wigner function $W$, is a Schwartz function. Moreover, the partial trace of a Schwartz-density operator is a Schwartz operator.

The following theorem contains the basic properties of Schwartz operators that we use:

**Theorem 13.** Let $H = L^2(\mathbb{R}^n)$ and $T \in B(H)$. Then,

(i) Let $f$ be a polynomial function on the entries of the vector $R = (Q_1, P_1, \ldots, Q_n, P_n)$ and $[W_\xi]$ a Weyl system. If $T \in \mathcal{S}(H)$, then $\text{Tr}[f(R)T] = \text{Tr}[f(R)]$. Moreover, $f(R)T \in \mathcal{S}(H)$ and $\text{Tr}[Wf(R)T] = \text{Tr}[f(R)TW] = \text{Tr}[TWf(R)]$.

(ii) If $T \in \mathcal{S}(H)$, then $T$ is trace-class.

(iii) If $T \in \mathcal{S}(H)$, then $|T| \in \mathcal{S}(H)$.

(iv) If $0 < T \in \mathcal{S}(H)$, then $\sqrt{T} \in \mathcal{S}(H)$.

(v) If $T \in \mathcal{S}(H)$ then $T^* \in \mathcal{S}(H)$.

For Schwartz-density operators, we can write explicit formulas for the gradient and Hessian of the characteristic function in terms of a trace: Consider $\{\xi_1\}_{i=1}^N$, any basis in $\mathbb{R}^N$; then, the gradient of $\chi(\xi)$, denoted by $\nabla \chi(\xi)$, is defined by the entries

$$\frac{\partial \chi(\xi)}{\partial \xi_k} = \frac{d}{dt} \chi(\xi + t\xi_k) \bigg|_{t=0}.$$  \hspace{1cm} \text{(26)}

The following Lemma generalizes Lemmas 5.4.2 and 5.4.3 in Ref. 8:

**Lemma 14.** (Gradient of the Weyl Operator). Let $T$ be a Schwartz operator and $\nabla_\eta := \eta \cdot \nabla$. Then, the following identities hold

$$\chi_{\nabla_\eta} \tau(\xi) = \left\{ \frac{1}{2} \xi_k \cdot \sigma \xi - i \frac{\partial}{\partial \xi_k} \right\} \chi_T(\xi),$$  \hspace{1cm} (24)

$$\chi_{\nabla \eta_k} (\xi) = \left\{ -\frac{1}{2} \xi_k \cdot \sigma \xi - i \frac{\partial}{\partial \xi_k} \right\} \chi_T(\xi).$$  \hspace{1cm} (25)
\[
(\nabla_{\kappa}(\kappa)) = \frac{i}{2} \text{Tr}(W_\xi (R_\eta) T) = \frac{i}{2} \text{Tr}(W_\xi (R_\eta, T)),
\]
(26)

\[
(\xi_\kappa \cdot \partial \kappa)(\kappa) = \text{Tr}(W_\xi (R_\eta) T) = \text{Tr}(W_\xi (R_\eta, T)).
\]
(27)

**Proof of Lemma 14.** First note that Eqs. (26) and (27) follow from adding and subtracting Eqs. (24) and (25) together with Theorem 13(i). We show first that

\[
\frac{d}{dt}(\text{Tr}(W_\eta T))_{t=0} = i\text{Tr}(R_\eta T).
\]
(28)

Since \( T \) is trace-class [Theorem 13(ii)], we can decompose \( T = T_1 + iT_2 \) with \( T_1, T_2 \) self-adjoint trace-class operators. Moreover, we can write \( T_1, T_2 \) as a finite linear combination of positive, trace-class operators, and thus, from the linearity of the trace, we can assume without loss of generality that \( T \) is a positive, trace-class operator. From the spectral decomposition \( R_\eta = / x dE(x) \) and the functional calculus, we obtain

\[
\left| \text{Tr} \left( \frac{W_\eta - 1}{it} - R_\eta \right) T \right| = \left| \int \left( \frac{e^{itx} - 1}{it} - x \right) \text{Tr} dE(x) T \right|
\leq \int \left| \frac{e^{itx} - 1}{it} - x \right| \text{Tr} dE(x) T.
\]

Using \( \left| \frac{e^{itx} - 1}{it} - x \right| \leq 2|x|, \) the Cauchy-Schwarz inequality, and Theorem 13(iv),

\[
\int \left| \frac{e^{itx} - 1}{it} - x \right| \text{Tr} dE(x) T \leq 2 \text{Tr} |R_\eta| T,
\]

\[
\leq 2 \left\| R_\eta \right\| \left\| \sqrt{T} \right\|_2 < \infty.
\]

Hence, from the dominated convergence theorem, we proved what is required. Now, we proceed to prove Eq. (24). From Theorem 13(i and ii), we have that \( R_\eta T \) is trace-class and, therefore, the Weyl transform exists. Moreover, \( \chi_{R_\eta T}(\kappa) \) is Schwartz, hence continuous, as it is the Weyl transform of a Schwartz operator (Proposition 11). Hence, we just need to verify the relation of Eq. (24) as Eq. (25) is similar. This follows directly from Eqs. (28) and (3),

\[
\chi_{R_\eta T}(\kappa) = -i \frac{d}{dt} \text{Tr}(W_\xi W_\eta T)_{t=0}
= -i \frac{d}{dt} \left( e^{it\kappa} \chi(\kappa + t\kappa) \right)_{t=0}.
\]

\[ \square \]

We remark that since \( T \) is Schwartz, higher order derivatives of \( \chi_T(\xi) \) can be written explicitly by using Theorem 13(i) and Eq. (26). For instance, the Hessian of \( \chi_T(\xi) \), where \( \rho \in \mathcal{A}(\mathcal{H}) \) has entries given by

\[
\frac{\partial^2 \chi(\xi)}{\partial \xi \partial \xi} = -\frac{1}{4} \text{Tr} [W_\xi (R_\eta, \{R_\xi, \rho\})].
\]
(29)

In particular, if the state \( \rho \) has covariance matrix \( \Gamma \) and displacement vector \( d \),

\[
\nabla \chi(0) = i d \quad \text{and} \quad (\text{Hessian} \; \chi)(0) = -\sigma \left( \frac{\Gamma}{2} + dd^T \right) \sigma^T.
\]
(30)

**III. PROOFS**

**A. Proof of Lemma 5**

We write \( S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \) and express the condition of the vanishing off-diagonal terms

\[
A \Gamma_1 \Gamma_1^T + B \Gamma_2 \Gamma_2^T = 0 \quad \text{for all} \quad \Gamma_1, \Gamma_2 \geq i \sigma.
\]
(31)
If we fix $A, B, C,$ and $D$, all different from zero such that Eq. (31) is satisfied, we can always find $\Gamma_1$ and $\Gamma_2$ such that for these choice of submatrices $A \Gamma_1 C^T + B \Gamma_2 D^T \neq 0$. Then, each summand in Eq. (31) must be zero. For instance, if $A \Gamma_1 C^T = -B \Gamma_2 D^T$ take now $\Gamma_1 \rightarrow 2 \Gamma_1$, then $A(2 \Gamma_1) C^T + B \Gamma_2 D^T = A \Gamma_1 C^T = 0$.

Without loss of generality, let us consider $A \Gamma_1 C^T = 0$, the other cases will be analogous. We take the singular value decomposition of $A = U D V$ and $C = W \Sigma Z$, where $U$, $V$, $W$, and $Z$ are unitaries. From here, we choose $\Gamma_1 = V^{-1} P Z^{-T}$, where $P$ is a positive definite matrix (recall that the sigma positive condition $\Gamma_1 \geq \sigma$ can be always obtained from rescaling a positive matrix). Then, $A \Gamma_1 C^T = U D P \Sigma W^T \neq 0$, every time we choose the proper $\Gamma_1$. Consequently, $A = 0$, $C = 0$, or both and likewise for $B$ and $D$.

The only matrices that fulfill the symplectic conditions are

$$
\begin{pmatrix}
A & 0 & 0 & B \\
0 & D & C & 0
\end{pmatrix},
$$

(32)

provided $A, B, C$, and $D$ are symplectic.

B. Proof of Lemma 6

First, it should be noted that the given assumptions immediately imply $S \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma \end{pmatrix} S^T = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}$ for all $\Gamma > 0$ and by continuity for all $\Gamma \geq 0$ and consequently for all $\Gamma = \Gamma^T$. The latter is due to the fact that every symmetric matrix can be decomposed in a semi-definite positive and negative part. The following Lemma shows how the block structure of $S$ is. We postpone its proof until the end.

**Lemma 15.** Let $S \in GL(4n, \mathbb{R})$ such that

$$
S \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma \end{pmatrix} S^T = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}
$$

is $2n \times 2n$ block diagonal for all symmetric $\Gamma \in \mathbb{R}^{2n \times 2n}$.

Here, $*, \cdot$ denote any $CM \in \mathbb{R}^{2n \times 2n}$. Then, it follows that:

(i) $S$ is either of the form

(a) $S = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ or $S = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}$.

(b) $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$,

with $A, B, C$ and $D \in \mathbb{R}^{2n \times 2n}$ invertible.

(ii) $S \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} S^T = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}$ is $2n \times 2n$ block diagonal for all $X \in \mathbb{R}^{2n \times 2n}$.

(34)

We consider the case where $A, B, C$, and $D$ are invertible, and the other cases will be contained here as we will see. Using part (ii) of Lemma 15, we have

$$
A \otimes C = -B \otimes D.
$$

(35)

We multiply Eq. (35) by $(A^{-1} \otimes \mathbb{1})$ from the left and take the trace in the first component, and likewise, we multiply Eq. (35) by $(1 \otimes D^{-1})$ and take the trace in the second component to obtain

$$
C = \alpha D \quad \text{where} \quad \alpha := \frac{-\text{Tr} A^{-1} B}{2n} \in \mathbb{R}
$$

(36)

and

$$
B = \beta A \quad \text{where} \quad \beta := \frac{-\text{Tr} D^{-1} C}{2n} \in \mathbb{R}.
$$

(37)
From equation Eq. (35), we obtain that \( \alpha = -\beta \) and that
\[
S = \begin{pmatrix}
A & aA \\
-aD & D
\end{pmatrix}.
\]
Moreover, the symplectic constraints on \( S \) give us that \( A, D \in \frac{1}{\sqrt{1+\alpha^2}} \text{Sp}(2n, \mathbb{R}) \). We write \( A = \frac{1}{\sqrt{1+\alpha^2}} X \) and \( D = \frac{1}{\sqrt{1+\alpha^2}} Y \) with \( X, Y \in \text{Sp}(2n, \mathbb{R}) \) to obtain Eq. (17). Finally, for \( \alpha \neq 0 \), we define \( y = \frac{1}{\alpha} \) and obtain the remaining equation. The case \( y = 0 \) is covered by \( \alpha \to \pm \infty \), and it gives the swap operation. Clearly, \( \alpha \neq 0 \) gives the local transformation.

**Proof of Lemma 15.** We decompose \( S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \) into blocks \( A, B, C, \) and \( D \in \mathbb{R}^{n \times n} \) and observe that Eq. (33) is equivalent to
\[
AC^T + BD^T = 0 \quad \text{for all } \Gamma = \Gamma^T.
\] (38)
Using tensor notation, this equation can be written\(^7\) as
\[
(A \otimes C + B \otimes D) \Gamma = 0 \quad \text{for all } \Gamma \in \mathbb{R}^{4n} \otimes \mathbb{R}^{4n} \text{ symmetric}
\]
\[
\iff (A \otimes C + B \otimes D) P_\alpha |X\rangle = 0 \quad \text{for all } |X\rangle \in \mathbb{R}^{4n} \otimes \mathbb{R}^{4n} \\
\iff (A \otimes C + B \otimes D) P_\alpha = 0
\]
\[
\iff P_\alpha (A \otimes C + B \otimes D)^T = 0,
\] (39)
where \( P_\alpha \) denotes the projector onto the symmetric subspace of \( \mathbb{R}^{4n} \otimes \mathbb{R}^{4n} \). We make the following remark that will be used frequently during this proof: the symmetrization or symmetric component of a non-zero product state does not vanish. Suppose it does, then \( P_\alpha ([|v\rangle \otimes |w\rangle] = 0 \) and \( |v\rangle \otimes |v\rangle + |w\rangle \otimes |w\rangle = 0 \). Performing the scalar product with \( |v\rangle \otimes |w\rangle \) leads to \( ||v|^2 + |w|^2|^2 = 0 \) which is only zero if and only if \( |v\rangle = |w\rangle = 0 \).

**On (i):** We now prove the first part of the Lemma by considering the two cases:

(a) One of the submatrices \( A, B, C, \) or \( D \) is zero; we only treat the case \( A = 0 \), the others are similar. Then, \( A^T \otimes C^T = 0 \) and \( B^T \) is invertible (otherwise \( S \) would not have full rank). Since any non-zero product \( B^T v \otimes D^T w \) (for some \( v, w \in \mathbb{R}^{2n} \)) would contain a non-vanishing symmetric component, Eq. (39) implies that \( D^T = 0 \).

(b) We prove by contradiction that in this case, the four submatrices are invertible. For instance, assume that \( A \) is not invertible. Then, there exists a vector \( 0 \neq |a\rangle \in \text{Ker} A^T \). We choose \( |e\rangle \notin \text{Ker} D^T \) (recall that \( D \neq 0 \)), and with (39), we then find
\[
0 = P_\alpha (A^T \otimes C^T + B^T \otimes D^T) |a\rangle \otimes |e\rangle = P_\alpha (B^T |a\rangle \otimes D^T |e\rangle).
\] (40)
By the same argument as in (a), we now conclude \( |a\rangle \in \text{Ker} B^T \). Moreover,
\[
S^T |a\rangle = \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} |a\rangle = 0.
\]
The latter is in contradiction to the invertibility of \( S \) (if \( S \) is an invertible matrix then the kernel is trivial), hence \( A \) — and due to analogous reasoning — \( B, C, \) and \( D \) are invertible.

**On (ii):** We now prove the second part of the Lemma, Eq. (34), by showing that the equivalent expression \( AXC^T + BXD^T = 0 \) for all \( X \in \mathbb{R}^{2n \times 2n} \) is true.

This is trivially satisfied if \( S \) is given in form of case \( [i(a)] \). Therefore, we are left with the case \( [i(b)] \) where, in particular, \( B \) and \( C \) are invertible. Without loss of generality, we choose \( B = C = I \) [this can be done by redefining \( A \to A^{-1}B \) and \( D \to D^{-1}C \) in (38)], and Eq. (39) reads
\[
P_\alpha (A^T \otimes 1 + 1 \otimes D^T) = 0.
\] (41)
We now show in three steps that \( A^T \otimes 1 + 1 \otimes D^T = 0 \), which then concludes the proof. First, we show that there exists \( \lambda \in \mathbb{C} \) such that \( \text{Spec}(A^T) = \{ \lambda \} \) and \( \text{Spec}(D^T) = \{ -\lambda \} \). To this purpose, we choose eigenvectors \( |e\rangle \) of \( A^T \) and \( |f\rangle \) of \( D^T \) with eigenvalues \( \lambda \) and \( \omega \) respectively. Then,
Again, since the symmetrization of a non-zero product state is different from zero, we find $\lambda = -\omega$. Note that this holds for arbitrary eigenvalues $\lambda$ of $A^T$ and $\omega$ of $D^2$.

Second, using the Jordan normal form decomposition, we decompose $A^T$ (and $D^2$) into a diagonalizable $\lambda I$ and nilpotent part $N_A(N_D)$ and observe that $(A^T \otimes 1 + 1 \otimes D^2) = (\lambda I \otimes 1 + N_A \otimes 1) + 1 \otimes (-\lambda I + N_D) = (N_A \otimes 1 + 1 \otimes N_D)$.

Finally, Eq. (41) reads

$$P_t(N_A \otimes 1 + 1 \otimes N_D) = 0,$$

and we can conclude the proof by deriving that this implies $(N_A \otimes 1 + 1 \otimes N_D) = 0$. This is the third step.

Assume $N_A \otimes 1 + 1 \otimes N_D \neq 0$. Using the symmetry argument about non-zero product states, we find $N_A \neq 0$ and $N_D \neq 0$. Let $s$ be such that $N_A^{-1} = 0$ and $N_A^{-1} = 0$. Then, we multiply (42) from the right by $N_A^{-1} \otimes 1$ to get $P_t(N_A^{-1} \otimes N_D) = 0$. But this, in turn, implies $N_A^{-1} \otimes N_D = 0$ and leads to a contradiction. Therefore,

$$(N_A \otimes 1 + 1 \otimes N_D) = 0.$$  

\[\square\]

C. Proof of the stability of DS Theorem 9

The proof involves a series of steps. We first use Parseval’s theorem to express the distance of the quantum states in terms of the $L_2$-distance of their respective characteristic functions. Next, we show that there is a ball $B_\varepsilon$ around the origin of phase space where the characteristic functions do not vanish. The radius of this ball scales inversely proportional to the largest variance of the input states and proportional to $\log(1/\varepsilon)$. Thus, the smaller the error parameter $\varepsilon$, the bigger this region is. We then proceed to bound the distance separately on $B_{\varepsilon/2}$ and its complement $B_{\varepsilon/2}^c$. For the latter region, we exploit the relation between the tails of a distribution and the finiteness of its moments. Inside $B_{\varepsilon/2}$, the problem is equivalent to the stability of the Gaussian functional Eq. (19) with the restricted convex domain. For that matter, the stability of the Gaussian functional equation is reduced to the stability of the fundamental functional equation, namely, the Cauchy functional equation. In the course of the proof, we have the particular goal of achieving bounds in terms of integrals of the type $\|\xi\|_2 \int |\hat{g}(\xi)|^2 d\xi$, $n = 2, 4$, because these can further be bounded in terms of traces of operator moments of the state, which exist due to the crucial assumption of the state being a Schwartz operator. Such a bound can easily be achieved with $\xi$ bounded away from the origin, which is why the integral is split into the two parts.

We may assume without loss of generality that the input states (and therefore the output states) are centered. This is justified by the fact that the trace and HS norms are unitarily invariant and the operation of “Gaussification” (the operation on bosonic quantum states which produces Gaussian states with the same first and second moments of the input state) commutes with the displacement operation $\rho \mapsto W_\rho D W_\rho^*$. We denote by $\chi_{a,b}, \chi_a$ and $\chi_b$ the characteristic functions of $\rho_{a,b}, \rho_a$ and $\rho_b$, respectively. We define

$$g := \rho_{ab} - \rho_a \otimes \rho_b.$$

From Lemma 17, we have $\|g\|_1 \leq 3\varepsilon$. Let $\eta_1, \eta_2 \in \mathbb{R}^{2n}$ and $R_1, R_2$ be vectors of canonical operators as in Eq. (13). We write

$$G(\eta_1, \eta_2) := \text{Tr}[\omega^{\eta_1 \cdot R_1} \otimes \omega^{\eta_2 \cdot R_2} g] = \chi_{a,b}(\eta_1, \eta_2) - \chi_a(\eta_1) \chi_b(\eta_2).$$

(44)

Now, using the covariance property of Gaussian unitary operations Eq. (8),

$$\chi_{a,b}(\xi) = \chi_{a,b}(S^T \xi),$$

and the fact that $G(\eta_1, 0) = G(0, \eta_2) = 0$, we write for all $\eta_1, \eta_2 \in \mathbb{R}^{2n}$, the dynamical process in terms of characteristic functions as

$$\chi_1(\cos \theta \eta_1 + \sin \theta \eta_2) \chi_2(\cos \theta \eta_2 - \sin \theta \eta_1) = \chi_1(\cos \theta \eta_1) \chi_1(\sin \theta \eta_2) \chi_2(\cos \theta \eta_2) \chi_2(-\sin \theta \eta_1) + G(\eta_1, \eta_2).$$

(45)

This last equation resembles the ideal functional equation Eq. (19) plus a new remainder term $G(\eta_1, \eta_2)$. From Hölder’s inequality and the definition of $G$, we note that $\|G\| \leq 3\varepsilon$. The following lemma establishes the ball where the characteristic functions do not vanish.

**Lemma 16.** Let $\chi_1$ and $\chi_2$ be two characteristic functions with respective density operators $\rho_1, \rho_2$ and covariance matrices $\Gamma_1, \Gamma_2$. Define for any $\theta \in \mathbb{R}/\{\pm 1\}$,
\[ G(\eta_1, \eta_2) := \chi_1(\cos \theta_{\eta_1})\chi_1(\sin \theta_{\eta_2})\chi_2(\cos \theta_{\eta_2})\chi_2(-\sin \theta_{\eta_1}) - \chi_1(\cos \theta_{\eta_1} + \sin \theta_{\eta_2})\chi_2(\cos \theta_{\eta_2} - \sin \theta_{\eta_1}). \]

Assume \(|G(\eta_1, \eta_2)| \leq 3\epsilon\) for all \(\eta_1, \eta_2 \in \mathbb{R}^{2n}\), let \(\lambda := \frac{1}{2} \max\{\|\Gamma_1\|, \|\Gamma_2\|\}\) be the largest variance of the states \(\rho_1\) and \(\rho_2\) and define

\[ r := \sqrt{\frac{1}{\lambda} \log_2 \frac{1}{\epsilon^{1/12}}}. \tag{46} \]

Then, for \(\eta \in \mathcal{B} := \{\xi \| \|\xi\|_2 \leq r\}\),

\[ |\chi_i(\eta)| \geq 12\epsilon^{1/12} \quad i=1, 2. \tag{47} \]

In order not to distract the reader from the key points of the main proof, we relegate the Proof of Lemma 16 to Sec. III D. The choice of exponent 1/12 for \(\epsilon\) in Eq. (46) will become clear in the estimates done in Sec. V.

We divide phase space in two separating regions. One where the characteristic functions do not vanish, namely, inside the ball \(\mathcal{B}_r := \{\xi \| \|\xi\|_2 \leq r\}\), and its complement which we denote by \(\mathcal{B}_r^c\). So that with the help of Parseval’s theorem, we express the distance between the two states as

\[ (2\pi)^n \|\rho_1 - \rho_2\|_2^2 = \|\chi_1 - \Phi\|_2^2 \]

\[ = \int_{\xi \in \mathcal{B}_r^c} |\chi_1(\xi) - \Phi(\xi)|^2 d\xi + \int_{\xi \in \mathcal{B}_r^c} |\chi_2(\xi) - \Phi(\xi)|^2 d\xi. \tag{48} \]

We compute the bound for \(\mathcal{B}_{r/2}\) and \(\mathcal{B}_{r/2}^c\), separately.

1. **Bound on the region where the characteristic function might vanish**

We use \(|z_1 - z_2|^2 \leq (|z_1|^2 + |z_2|^2)/2\) for \(z_1, z_2 \in \mathbb{C}\) to express the bound as

\[ \int_{\xi \in \mathcal{B}_{r/2}^c} |\chi_1(\xi) - \Phi(\xi)|^2 d\xi \leq \frac{1}{2} \int_{\xi \in \mathcal{B}_{r/2}^c} |\chi_1(\xi)|^2 d\xi + \frac{1}{2} \int_{\xi \in \mathcal{B}_{r/2}^c} |\Phi(\xi)|^2 d\xi. \tag{49} \]

For the first term of the RHS of Eq. (49), we use that \(1 < 2\|\xi\|_2/r\) for \(\xi \in \mathcal{B}_{r/2}^c\) so that

\[ \int_{\xi \in \mathcal{B}_{r/2}^c} |\chi_1(\xi)|^2 d\xi \leq \frac{4}{r^2} \int_{\xi \in \mathbb{C}^n} \|\xi\|^2_2 |\chi_1(\xi)|^2 d\xi. \tag{50} \]

Let us denote by \(\mathcal{W}(\eta)\) the Wigner function of \(|\chi_1(\xi)|^2\) (the product of two characteristic functions is a characteristic function),

\[ \mathcal{W}(\eta) = \frac{1}{(2\pi)^{n/2}} \int_{\xi \in \mathbb{C}^n} e^{i\eta \cdot \xi} |\chi_1(\xi)|^2 d\xi. \tag{51} \]

It can be easily verified by direct computation that the characteristic function \(|\chi_1(\xi)|^2\) is centered and has CM \(2\Gamma_1\). Moreover, we have

\[ \int_{\xi \in \mathbb{C}^n} \|\xi\|^2_2 |\chi_1(\xi)|^2 d\xi = -(2\pi)^n \sum_{k=1}^{2n} \frac{\partial^2 \mathcal{W}(0)}{\partial \eta_k^2}, \]

where \(\eta_k\) are the components of the vector \(\eta\) in an arbitrary but fixed basis.

We use now the representation of the Wigner function in terms of the expectation values of the parity operator \(P\)

\[ \mathcal{W}(\eta) = \frac{1}{n^2} \text{Tr}[\rho W_\eta PW_\eta], \tag{52} \]

where \(\rho\) is the density operator corresponding to the characteristic function \(|\chi_1(\xi)|^2\). The operator \(\rho\) is clearly Schwartz as its characteristic function is a Schwartz function (Corollary 12). The parity operator \(P\) is the n-fold tensor product of the parity operators for a single degree of freedom and is the unitary operator that satisfies
\[ PW_{\xi}P^* = W_{-\xi}, \]
\[ PR_{k}P^* = -R_{k}, \]
\[ P = P^* = P^{-1}. \]

Using Eqs. (53) and (26), we compute
\[ \frac{\partial^2 W(0)}{\partial \eta_k \partial \eta_l} = -\frac{2}{\pi^2} \text{Tr} [PP(R_{\eta_k}R_{\eta_l})]. \] (54)

Thus, from Eqs. (50), (52), and (54), we find that
\[ \int_{\xi \in \mathbb{R}^{2n}} |\chi_{1}(\xi)|^2 d\xi \leq \frac{2^{n+4}}{\pi^2} \sum_{k=1}^{2n} \text{Tr} [PPR_k^2]. \] (55)

Hence, we just need to bound the terms Tr[PPR_k^2]. In order to do this, we notice that
\[ PP = \sqrt{\rho} \sqrt{\rho} = \rho P. \]

Moreover, from the spectral decomposition of \( \rho \), we have \( \rho P = \rho \sqrt{\rho} \sqrt{\rho} \). Accordingly,
\[ \text{Tr}[PPR_k^2] = \text{Tr}[\sqrt{\rho} P \sqrt{\rho}] \leq \| \sqrt{\rho} \sqrt{\rho} \|_1 \]
\[ = \text{Tr}[\rho R_k^2]. \]

Here, we have used the Cauchy-Schwarz inequality and the cyclicity of the trace that comes from the properties of the Schwartz operator \( \rho \) [see Theorem 13 (i), (ii), and (iv)].

In summary, we have the following bound for the tails of our characteristic function:
\[ \int_{\xi \in \mathbb{R}^{2n}} |\chi_{1}(\xi)|^2 d\xi \leq \frac{2^{n+4}}{\pi^2} \text{Tr} \Gamma_1 \]
\[ = \left( 2^{n+4}12\lambda \text{Tr} \Gamma_1 \right) \frac{1}{\log \frac{1}{\tau}}. \]

Since \( \Phi(\xi) \) has the same CM as \( \chi(\xi) \), we can use the same bound to obtain
\[ \frac{1}{(2\pi)^n} \int_{\xi \in \mathbb{R}^{2n}} |\chi_{1}(\xi) - \Phi(\xi)|^2 d\xi \leq \left( \frac{192\lambda \text{Tr} \Gamma_1}{\pi^n} \right) \frac{1}{\log \frac{1}{\tau}} \]
\[ \leq \frac{c_2}{\log \frac{1}{\tau}}, \] (56)
where
\[ c_2 := 8 \sqrt{3\lambda \text{Tr} \Gamma_{ab} \pi^n}, \] (57)

and \( \Gamma_{ab} \) is the CM of the output state \( \rho_{ab} \). Note that since the BS is a passive transformation, the trace of the input CM, \( \Gamma_1 \oplus \Gamma_2 \), is the same as the trace of the output CM \( \Gamma_{ab} \). Thus, it is clear that \( \text{Tr} \Gamma_1 \leq \text{Tr} \Gamma_{ab} \).
2. Bound inside the region where $\chi$ does not vanish

We proceed to compute the bound for the first term of the RHS of Eq. (48). Our strategy here is to reduce the perturbed functional Eq. (45) to the perturbed functional equation of a linear map in a restricted domain (cf. Refs. 14, 20, and 21). This will imply that inside the ball $B_r$, the characteristic functions are a product of a Gaussian and a non-Gaussian part whose moments can be controlled by the Schwartz properties of $\rho_1$ and $\rho_2$. For that matter, let $\eta_1, \eta_2 \in \mathbb{R}^{2n}$ with $|\eta_j|^2 < r/2$ so that $\cos \theta_{\eta_1} + \sin \theta_{\eta_2} - \sin \theta_{\eta_1} \in B_r$. Let us take the logarithm (principal branch) on both sides of Eq. (45). We write

$$\Psi_1(\cos \theta_{\eta_1} + \sin \theta_{\eta_2}) + \Psi_2(\cos \theta_{\eta_2} - \sin \theta_{\eta_1}) = \Psi_1(\cos \theta_{\eta_1}) + \Psi_2(\cos \theta_{\eta_2}) + \Psi_2(-\sin \theta_{\eta_1}) + Q(\eta_1, \eta_2),$$

(58)

where $\Psi_j(\eta) := -\log x_j(\eta)$ for $j = 1, 2$ and

$$Q(\eta_1, \eta_2) := -\log \left( 1 + \frac{Q(\eta_1, \eta_2)}{x_1(\cos \theta_{\eta_1})x_2(\cos \theta_{\eta_1})x_2(-\sin \theta_{\eta_1})} \right).$$

(59)

Since $\rho_1, \rho_2$ is Schwartz, we can define continuous vector-valued functions $\phi_j(\xi) : B_{r/2} \rightarrow \mathbb{C}^{2n}$ by

$$\phi_j(\xi) := \nabla \Psi_j(\xi).$$

Note that $\phi_j(\xi), j = 1, 2$ are in fact conservative vector fields and that $\phi_j(0) = 0$ as $\rho_j$ is centered. The gradient of $\phi_j$ is the Hessian of $x_j$ (see Lemma 14) and $2\nabla \phi_j(0) = \sigma_j^2 \sigma_i^T$ for $j = 1, 2$.

3. Inhomogeneous Cauchy functional equation

Next, we want to obtain a functional equation only depending on $x_1$ or $x_2$. In order to do so, we differentiate Eq. (58) in the direction of $\eta_1$ to find

$$\cos \theta_1(\cos \theta_{\eta_1} + \sin \theta_{\eta_2}) - \sin \theta_2(\cos \theta_{\eta_2} - \sin \theta_{\eta_1}) = \cos \theta_1(\cos \theta_{\eta_1}) - \sin \theta_2(\cos \theta_{\eta_2}) + Q_1(\eta_1, \eta_2).$$

(60)

Here, $Q_1(\eta_1, \eta_2)$ is the gradient of $Q(\eta_1, \eta_2)$ with respect to $\eta_1$, whose entries in the coordinate system $(\xi_1)_{k=1}^{2n}$ are $\frac{\partial Q_1(\xi_1)}{\partial \xi_k}$. We evaluate in Eq. (60) $\eta_1 = 0$ to get

$$\cos \theta_1(\sin \theta_{\eta_2}) - \sin \theta_2(\cos \theta_{\eta_2}) = Q_1(0, \eta_2).$$

(61)

In a similar fashion, we differentiate Eq. (58) in the direction of $\eta_2$ and set to zero to obtain

$$\sin \theta_1(\cos \theta_{\eta_1} + \sin \theta_{\eta_2}) + \cos \theta_2(\cos \theta_{\eta_2} - \sin \theta_{\eta_1}) = \sin \theta_1(\sin \theta_{\eta_1} + \cos \theta_{\eta_2}) + \cos \theta_2(\cos \theta_{\eta_2}) + Q_2(\eta_1, \eta_2),$$

(62)

$$\sin \theta_1(\cos \theta_{\eta_1}) + \cos \theta_2(-\sin \theta_{\eta_1}) = Q_2(\eta_1, 0),$$

(63)

where $Q_2(\eta_1, \eta_2)$ is the gradient of $Q(\eta_1, \eta_2)$ with respect to $\eta_2$. The entries of $Q_2(\eta_1, \eta_2)$ are $\frac{\partial Q_2(\eta_1, \eta_2)}{\partial \eta_k}$.

Now, we will be able to decouple $\phi_1$ and $\phi_2$. First, we substract Eq. (61) from Eqs. (60) and (63) from Eq. (62) to obtain

$$\left[ \phi_1(\cos \theta_{\eta_1} + \sin \theta_{\eta_2}) - \phi_1(\cos \theta_{\eta_1}) - \phi_2(\sin \theta_{\eta_2}) \right] = \tan \theta \left[ \phi_2(\cos \theta_{\eta_2} - \sin \theta_{\eta_1}) - \phi_2(\cos \theta_{\eta_2}) - \phi_2(-\sin \theta_{\eta_1}) \right] + \frac{Q_1(\eta_1, \eta_2) - Q_1(0, \eta_2)}{\cos \theta},$$

(64)

$$\left[ \phi_2(\cos \theta_{\eta_2} - \sin \theta_{\eta_1}) - \phi_2(\cos \theta_{\eta_2}) - \phi_2(-\sin \theta_{\eta_1}) \right] = -\tan \theta \left[ \phi_1(\cos \theta_{\eta_1} + \sin \theta_{\eta_2}) - \phi_1(\cos \theta_{\eta_1}) - \phi_1(\sin \theta_{\eta_2}) \right] + \frac{Q_2(\eta_1, \eta_2) - Q_2(\eta_1, 0)}{\cos \theta}. $$

(65)

Thus, from Eqs. (64) and (65), we find the following inhomogeneous Cauchy equations
\[
\begin{align*}
\phi_1(\cos \theta \eta_1 + \sin \theta \eta_2) - \phi_1(\cos \theta \eta_1) - \phi_1(\sin \theta \eta_2) & = \sin \theta\left(Q_2(\eta_1, \eta_2) - Q_2(\eta_1, 0)\right) \\
& \quad + \cos \theta\left[Q_2(\eta_1, \eta_2) - Q_2(0, \eta_2)\right], \tag{66}
\end{align*}
\]

\[
\begin{align*}
\phi_2(\cos \theta \eta_2 - \sin \theta \eta_1) - \phi_2(\cos \theta \eta_2) - \phi_2(-\sin \theta \eta_1) & = \sin \theta\left[Q_2(\eta_1, \eta_2) - Q_2(0, \eta_2)\right] \\
& \quad + \cos \theta\left[Q_2(\eta_1, \eta_2) - Q_2(\eta_1, 0)\right]. \tag{67}
\end{align*}
\]

4. Bound on $\Omega_{r/2}$

Now that $\phi_1$ and $\phi_2$ are decoupled, we continue only with $\phi_1$ as with $\phi_2$ is analogous and the same upper bound is obtained. We recall that the derivative of a vector with respect to a vector can be represented as a matrix. Thus, when we differentiate Eq. (66) in the direction of $\eta_2$ and evaluate at $\eta_2 = 0$, we obtain the following matrix-valued equation:

\[
\nabla \phi_1(\cos \theta \eta_1) = \frac{\alpha F}{2} + \frac{Q_{22}(0, 0)}{\tan \theta} = Q_{22}(\eta_1) + \frac{Q_{12}(\eta_1)}{\tan \theta}. \tag{68}
\]

Here, $Q_{12}(\eta_1, 0) : \mathbb{R}^{2n} \to \mathbb{C}^{2n \times 2n}$ and $Q_{22}(\eta_1, 0) : \mathbb{R}^{2n} \to \mathbb{C}^{2n \times 2n}$ are matrix-valued functions defined to have the following $(k, l)$-entries:

\[
\left. \frac{\partial^2 Q(\eta_1, s\xi_2 + t\xi_2)}{\partial s \partial t} \right|_{s=0} \quad \text{and} \quad \left. \frac{\partial^2 Q(\eta_1, s\xi_2 + t\xi_2)}{\partial s \partial t} \right|_{t=0}, \tag{69}
\]

respectively.

Accordingly, we integrate the previous equation twice from zero to $\eta$, $\|\eta\|_2 \leq r/2$. We obtain for $\xi \in \Omega_{r/2}$,

\[
\chi_1(\xi) = \exp\left(-\xi \cdot \left(\frac{\alpha F}{2} - \frac{V}{2}\right)\xi - F\left(\frac{\xi}{\cos \theta}\right)\right), \tag{70}
\]

with

\[
V := -\frac{Q_{12}(0, 0)}{\tan \theta} = \frac{\alpha (\text{Tr} R_1 R_2^T) \sigma^T}{\tan \theta},
\]

\[
F(\xi) := \cos^2 \theta \int_{C(\xi)} \left(\int_{C(\eta)} \left(\frac{Q_{22}(\eta_1)}{\tan \theta} + \frac{Q_{12}(\eta_1)}{\tan \theta}\right) \cdot d\eta_1\right) \cdot d\eta,
\]

where $C(\xi), C(\eta)$ are curves in phase space connecting the origin with the vectors $\xi$ and $\eta$. Moreover, these last terms can be upper bounded by (see Appendixes A and B)

\[
\|V\|_2 \leq \left(\frac{\sqrt{24 n R \kappa}}{|\tan \theta|}\right)^{\sqrt{C}}, \tag{71}
\]

\[
\left|F\left(\frac{\xi}{\cos \theta}\right)\right| \leq \left(\frac{n^2 \kappa}{2 \tan^2 \theta}\right)^{\xi^2/3}, \tag{72}
\]

where the largest absolute fourth moment of $\rho_{ab}$ is defined as

\[
\kappa := \max\{\|\rho_{ab} R_2^T\|_1 \quad | \quad \|\xi\|_2 = \|\eta\|_2 = 1\}. \tag{73}
\]

At this point, we can show that the CM of $\rho_1$ and $\rho_2$ are $\varepsilon$–close. From differentiating Eq. (61), with respect to $\eta_2$ and evaluating at zero, we get the relation between the CMs of $\rho_1$ and $\rho_2$. 
\[ \Gamma_1 - \Gamma_2 = \frac{2}{\cos^2 \theta} V. \]

Hence, from Eq. (71),
\[ |\Gamma_1 - \Gamma_2| \leq \left( \frac{\sqrt{384n^2 \kappa}}{\sin 2\theta} \right) \sqrt{\epsilon}. \] (74)

Now, we can proceed to show that for \( \xi \in \mathbb{B}_{r/2} \), the characteristic function of the state \( \rho_1 \) is \( \epsilon \)-close to the Gaussian characteristic function \( \Phi = \exp[\xi' \cdot i \xi/4] \).

We use Eq. (70), \(|e^i - 1| \leq |x| \max \{1, e^{2\Re(x)}\}\) and \(|\chi_1(\xi)|^2 = |\Phi(\xi)|^2 \epsilon | \xi V - 2\Re[\Phi(\xi) / \cos \theta] | \) to write the bound as
\[
\begin{align*}
\int_{\xi \in \mathbb{B}_{r/2}} |\chi_1(\xi) - \Phi(\xi)|^2 d\xi &= \int_{\xi \in \mathbb{B}_{r/2}} |\Phi(\xi)|^2 |\exp[\xi \cdot V / 2 - F(\xi / \cos \theta)] - 1|^2 d\xi \\
&\leq \int_{\xi \in \mathbb{B}_{r/2}} |\chi_1(\xi)|^2 \left| \frac{\xi \cdot V \xi}{2} - F(\xi / \cos \theta) \right|^2 d\xi \\
&\leq \frac{1}{2} \int_{\xi \in \mathbb{B}_{r/2}} |\chi_1(\xi)|^2 \left( \left| \frac{\xi \cdot V \xi}{2} \right|^2 + |F(\xi / \cos \theta)|^2 \right) d\xi \\
&\leq \left( \frac{4n^2 \kappa \epsilon^2 / 3}{\tan^2 \theta} \right) \int_{\xi \in \mathbb{B}_{r/2}} |\chi_1(\xi)|^2 |\xi|^2 d\xi. \tag{75}
\end{align*}
\]

Here, in the second inequality, we have used again \(|z_1 - z_2|^2 \leq (|z_1|^2 + |z_2|^2) / 2\) for \( z_1, z_2 \in \mathbb{C} \) and in the last inequality Eqs. (71) and (72). We show now that
\[
\frac{1}{(2\pi)^n} \int_{\xi \in \mathbb{B}_{r/2}} |\chi_1(\xi)|^2 |\xi|^2 d\xi \leq \frac{512n^2 \kappa}{\pi^n} \left( 1 + \frac{3 \sin 2\theta}{\cos^2 \theta} \right). \tag{76}
\]

We used again the Wigner representation, Eq. (51), of the characteristic function \(|\chi_1(\xi)|^2\) in order to compute the bound of Eq. (76),
\[
\begin{align*}
\frac{1}{(2\pi)^n} \int_{\xi \in \mathbb{B}_{r/2}} |\chi_1(\xi)|^2 |\xi|^2 d\xi &\leq \frac{1}{(2\pi)^n} \int_{\xi \in \mathbb{R}^n} |\chi_1(\xi)|^2 |\xi|^2 d\xi \\
&= \frac{1}{(2\pi)^n} \sum_{k \in \mathbb{Z}^n} \int_{\xi \in \mathbb{R}^n} |\chi_1(\xi)|^2 |\xi|^2 d\xi \\
&= \sum_{k \in \mathbb{Z}^n} \left. \frac{\partial^2 \mathcal{W}(0)}{\partial \eta_k^2 \partial \eta_l} \right|_{\xi = 0}. \\
\end{align*}
\]

Using Lemma 14 with the help of the CCR gives
\[
\sum_{k, l = 1}^{2n} \left. \frac{\partial^2 \mathcal{W}(0)}{\partial \eta_k^2 \partial \eta_l} \right|_{\xi = 0} = \frac{8}{\pi^n} \sum_{k, l = 1}^{2n} \text{Tr}[\rho^2 \rho (R_k R_l^\dagger)].
\]

Here, \( \rho \) is again the density operator corresponding to the characteristic function \(|\chi(\xi)|^2\). We use again that \( \mathcal{P} = \rho \mathcal{P} \) (see Subsection III C) and Hölder’s inequality to bound
\[
\frac{8}{\pi^n} \sum_{k, l = 1}^{2n} \text{Tr}[\rho^2 \rho (R_k R_l^\dagger)] \leq \frac{8}{\pi^n} \sum_{k, l = 1}^{2n} \left( \| \sqrt{\rho} R_k^\dagger R_l^{1/2} \|^1 + \| \sqrt{\rho} R_k^\dagger R_l^{1/2} \|_1 \right).
\]

Now using Cauchy-Schwarz inequality and the cyclicity properties for Schwartz operators, we find
Since we want to specify all the constants in terms of the moments of the output state \( \rho_{ab} \), we need to do the following computations. Using the explicit form of the BS transformation Eq. (18) and the derivatives of \( \chi(x) \), we obtain after a tedious but straightforward calculation

\[
\cos^4 \theta \sum_{k=1}^{2n} \text{Tr} \rho R^4_k + \sin^4 \theta \sum_{k=1}^{2n} \text{Tr} \rho R^4_{2k} - 6 \sin^2 \theta \cos^2 \theta \sum_{k=1}^{2n} \text{Tr} \rho \rho R^2_{1k} \text{Tr} \rho R^2_{2k} = \sum_{k=1}^{2n} \text{Tr} \rho_{ab} R^4_k,
\]

which together with the positivity of the fourth moments implies

\[
\sum_{k=1}^{2n} \text{Tr} \rho R^4_k \leq 6 \tan^2 \theta \sum_{k=1}^{2n} \text{Tr} \rho \rho R^2_{1k} \text{Tr} \rho R^2_{2k} + \frac{1}{\cos^4 \theta} \sum_{k=1}^{2n} \text{Tr} \rho_{ab} R^4_k. \tag{78}
\]

Moreover, the relation between the fourth moments of the symmetrized state \( \rho \) and \( \rho_1 \) is given by

\[
\text{Tr} \rho R^4_k = 2 \text{Tr} \rho_1 R^4_k + 6 (\text{Tr} \rho_1 R^2_k)^2, \quad k = 1, \ldots, 2n. \tag{79}
\]

Combining Eqs. (78) and (79) and using \( (\text{Tr} \rho_1 R^2_k)^2 \leq \text{Tr} \rho_1 R^4_k \leq \kappa \), we obtain

\[
\sum_{k=1}^{2n} \text{Tr} \rho R^4_k \leq 16n \kappa \left( 1 + 3 \sin 2\theta \right),
\]

and Eq. (77) gives the claimed bound in Eq. (76). Hence, inserting Eq. (76) in Eq. (75), we obtain the bound for the non-vanishing region \( \mathbb{B}_{1/2} \),

\[
\frac{1}{(2\pi)^n} \int_{x \in \mathbb{B}_{1/2}} |\chi(x) - \Phi(x)|^2 dx \leq c_1 2^{1/3}, \tag{80}
\]

where

\[
c_1 := 32 \sqrt{\frac{2}{\pi^4} \left( 1 + 3 \sin 2\theta \right) \pi^2 \kappa}. \tag{81}
\]

The result of the theorem follows from Eqs. (56) and (80) and the elementary inequality for non-negative scalars \( x_1, x_2, \sqrt{x_1^2 + x_2^2} \leq \sqrt{x_1^2} + \sqrt{x_2^2} \).

**D. Auxiliary Lemmas**

**Lemma 17.** Let \( \rho_x \) be a density operator of a bipartite system with reduced states \( \rho_1 \) and \( \rho_2 \). If \( \rho_1 \otimes \rho_2 \) describe an arbitrary product state and \( \| \rho_12 - \rho_1 \otimes \rho_2 \|_1 \leq \epsilon \), then \( \| \rho_12 - \rho_1 \otimes \rho_2 \|_1 \leq 3 \epsilon \).

**Proof.** Using the triangle inequality twice, we find

\[
\| \rho_12 - \rho_1 \otimes \rho_2 \|_1 \leq \| \rho_12 - \rho_1 \otimes \rho_2 \|_1 + \| \rho_1 \otimes \rho_2 - \rho_1 \otimes \rho_2 \|_1 \leq \epsilon + \| \rho_1 \otimes \rho_2 - \rho_1 \otimes \rho_2 \|_1 + \| \rho_1 \otimes \rho_2 - \rho_1 \otimes \rho_2 \|_1 = \epsilon + \| \rho_2 - \rho_2 \|_1 + \| \rho_1 - \rho_1 \|_1.
\]

Exploiting that \( \| X \|_1 = \sup_{Y \| Y \|_1 \leq 1} \text{Tr} \[ YX \] \), we can bound
\[
\|\hat{\rho}_1 - \rho_1\|_1 = \sup_{Y \in \mathbb{Y}} \text{Tr}(Y \otimes I)(\rho_1 \otimes \hat{\rho}_2 - \rho_1)
\]
\[
\leq \sup_{\hat{Y} \in \mathbb{Y}} \text{Tr}\{\hat{Y}(\rho_1 \otimes \hat{\rho}_2 - \rho_1)\},
\]
\[
= \|\rho_1 \otimes \rho_2 - \rho_1\|_1 \leq \varepsilon
\]

and similar for the other term.

**Proof of Lemma (16).** We follow the proof idea of Lemma 1 from Ref. 14. Without loss of generality, assume \(0 < \theta \leq \pi/4\) and set \(\eta_2 = \tan \theta \eta_1, \eta_1 = \xi\) in Eq. (45). In case \(\theta > \pi/4\), set \(\eta_1 = \eta_2 / \tan \theta\) in Eq. (45) and proceed likewise. Then, Eq. (45) becomes

\[
\chi_1\left(1 + \tan^2 \theta \right) = \chi_1(\cos \theta \xi)\chi_2(\cos \theta \tan \theta \xi)(-\sin \theta \xi) + G(\xi, \tan \theta \xi),
\]

for all \(\xi \in \mathbb{R}^{2n}\). Replace \(\xi \mapsto (\xi / \cos \theta)\) in the previous equation to obtain

\[
\chi_1\left(1 + \tan^2 \theta \right)\frac{\xi}{\cos \theta} = \chi_1(\xi)\chi_2(\tan \theta \xi)^2 + G\left(\frac{\xi}{\cos \theta}, \frac{\tan \theta \xi}{\cos \theta}\right).
\]

Since \(\|G\| \leq 3\varepsilon\), we have for all \(\xi \in \mathbb{R}^{2n}\),

\[
\left|\chi_1\left(1 + \tan^2 \theta \right)\frac{\xi}{\cos \theta}\right| \geq \left|\chi_1(\xi)\chi_2(\tan \theta \xi)^2\right| - 3\varepsilon.
\]

Similarly, for \(0 < \theta \leq \pi/4\) and \(\eta_1 = -\tan \theta \eta_2, \eta_2 = \xi\), we arrive to

\[
\left|\chi_2\left(1 + \tan^2 \theta \right)\frac{\xi}{\cos \theta}\right| \geq \left|\chi_2(\xi)\chi_2(\tan \theta \xi)^2\right| - 3\varepsilon.
\]

With \(\gamma(\xi) = \min_{\eta \in [0, \xi]} |\chi_0(\eta)|\), we obtain \(\gamma\left(1 + \tan^2 \theta \right)\xi \geq \gamma(\xi) - 3\varepsilon\). Replacing \(\xi\) by \(\left(1 + \tan^2 \theta \right)^k \xi\) with \(k \in \mathbb{N}\) in the previous equation gives

\[
\gamma\left(1 + \tan^2 \theta \right)^{k+1} \xi \geq \gamma(\xi) - 3\varepsilon \quad \text{for all } \xi \in \mathbb{R}^{2n}, k \in \mathbb{N}.
\]

(82)

It is a fact that for any classical characteristic function \(\phi(t)\) with variance \(\lambda\), the following inequality holds (see for instance Ref. 22, p. 89):

\[
|\phi(t)| \geq 1 - \frac{1}{2} \lambda t^2.
\]

(83)

Let us fix \(\xi\) in the direction of phase space in which we obtain the largest variance \(\lambda\) of \(\rho_1\) and \(\rho_2\) and consider the region where \(\|\xi\|_2 \leq \sqrt{\frac{1}{\lambda}}\). Then, from Eq. (83), for \(\|\xi\|_2 \leq \sqrt{\frac{1}{\lambda}}\), we have \(|\chi(\xi)| \geq 1/2\).

Using Eq. (82) and the inequality \((1 + a)^n \geq 1 + na, \forall n \in \mathbb{N}, \forall a \in [0, 1]\), we can show by induction that

\[
\gamma\left(1 + \tan^2 \theta \right)^{k+1} \xi \geq \left(\frac{1}{2}\right)^{k+1} - 4\varepsilon \quad \text{for } k \in \mathbb{N}, \|\xi\|_2 \leq \sqrt{\frac{1}{\lambda}}.
\]

Moreover, for \(\varepsilon < 1\), we clearly have

\[
\gamma\left(1 + \tan^2 \theta \right)^{k+1} \xi \geq \left(\frac{1}{2}\right)^{k} - 4\varepsilon \quad \text{for } k \in \mathbb{N}, \|\xi\|_2 \leq \sqrt{\frac{1}{\lambda}}.
\]

(84)

Finally, we take \(k_0\) such that \(2^{k_0} = \sqrt{\frac{1}{2\pi^2}}\), to obtain with the help of Eq. (84),
\[ y \left( 1 + \tan^2 \theta \right)^{\gamma_0} \xi \geq \left( \frac{1}{2} \right)^{d_{\gamma_0} - 1} - 4\epsilon^{1/12} = 12\epsilon^{1/12}. \]

We thus have \( y \left( 1 + \tan^2 \theta \right)^{\gamma_0} \xi > \epsilon^{1/12} \) and \( 1 + \tan^2 \theta \right)^{\gamma_0} \xi \in B_\gamma \) as claimed.

\section*{IV. DISCUSSION}

The DS theorem can be understood as the statement that Gaussian bosonic states with same covariance matrix are the only fixed point states of a non-trivial beam splitter transformation. In contrast with other characterizations of Gaussian states such as the one from Hudson,\textsuperscript{11} the DS theorem does not require any constraint on the purity of the state. The stability result of Theorem 9 provides an explicit estimate of the robustness of a characterization of Gaussian states through linear independence. In particular, we have obtained an estimate of the constants which reflects the fact that the quantum DS theorem is unstable when the beam splitter is close to being transparent \((\theta = 0)\) or a mirror \((\theta = \pi/2)\). Throughout this work, we have made an effort to present explicit constants as well as to improve the order of the error parameter; however, this does not mean that they are anywhere close to optimal. In fact, it is not known to us if the optimal constant must necessarily depend on the number of modes \(n\) or whether the \(\log(1/\epsilon)^{1/2}\) dependence can be lifted to a polynomial dependence.

The Darmois-Skitovich theorem is not only interesting as a neat characterization problem but also because of its practical applications: it is the main theoretical concept behind the signal reconstruction method known as \textit{blind source separation}\textsuperscript{24} and \textit{independent component analysis},\textsuperscript{25} which are actively studied in the field of communication and signal processing. These classical applications only work because the result is sufficiently robust. It is conceivable that there will be analogous quantum applications where robustness plays a similar role. We hope with this study of the stability of the quantum DS theorem to stimulate a further investigation of this theorem and its extensions in the quantum information community.

Finally, since the stability bound only depends on the moments up to 4 order, it seems reasonable to expect that the result would hold for all density operators with finite 4th moments. In order to extend the stability result to the class of states with finite 4th moments, it would be enough to prove that Schwartz density operators are dense in this class. That is to say, given a density operator with finite 4th moments, we need to show that there always exist a Schwartz density operator that approximates the original density operator arbitrarily well in trace norm \textit{and} whose moments up to 4 also approximate arbitrarily well the moments of the original density operator. With such results and the triangle inequality, one could show that the stability of the DS theorem for quantum states with finite 4th moments can be reduced to Theorem 9. It was proven in Ref. 4 that Schwartz operators are dense in the set of quantum states in the trace norm, however, without any information about how the moments are approximated. The required result for extending the stability result seems to be a quantum version of the Meyers-Serrin inequality, one could show that the stability of the DS theorem for quantum states with finite 4th moments can be reduced to Theorem 9. It was proven in Ref. 4 that Schwartz operators are dense in the set of quantum states in the trace norm, however, without any information about how the moments are approximated. The required result for extending the stability result seems to be a quantum version of the Meyers-Serrin theorem in Sobolev spaces (see Sec. 3.4 in Ref. 26 for a quantum version of the Meyers-Serrin theorem for symmetric moments). A general version of the quantum Meyers-Serrin theorem would be an important technical result in its own right. It would make the set of Schwartz density operators available for proving results in larger classes of quantum states whose direct use of restrictive technical properties had made the proof very difficult to achieve at first place.

\section*{ACKNOWLEDGMENTS}

I would like to thank Michael M. Wolf for his continuous interest and many insightful discussions throughout this work. In particular, I thank him for suggesting the problem considered in this paper. In addition, I would like to thank A. Michelangeli for an insightful discussion of Lemma 1 from Ref. 14, and M. Keyl for clarifications concerning Schwartz operators. Furthermore, I wish to thank M. Christandl and J. P. Solovej for their hospitality during my visit at the QMATH Center as well as the financial support from the Danish Council for Independent Research (Sapere Aude). This work was partly written during this visit.

\section*{APPENDIX A: UPPER BOUND OF Eq. (71)}

Let us write \( R_{1k}, k = 1, \ldots, 2n, \) and \( R_{2l}, l = 1, \ldots, 2n, \) for the entries of the vector \( R_1 = (Q_1, P_1, \ldots, Q_n, P_n) \) and \( R_2 = (Q_{n+1}, P_{n+1}, \ldots, Q_{2n}, P_{2n}) \), respectively. First, from the orthogonality of the symplectic matrix \( \sigma \),

\[ \| V \|_2 = \frac{\| \sigma(Tr gR_{1k} \sigma^T) \|_2}{\| \tan \theta \|} = \frac{\| \sigma(Tr gR_{1k} R_{2l}^T) \|_2}{\| \tan \theta \|} = \left( \frac{\| \sum \| \sigma(Tr gR_{1k} R_{2l}^T) \|_2}{\| \tan \theta \|} \right)^{1/2}. \]

Our plan is to bound each entry \( \| \sigma(Tr gR_{1k} R_{2l}) \|_2^2 \). Since the operator \( g \) is the difference of two Schwartz operators, it is also Schwartz. Thus, we have from Theorem 13 (iii) and (iv) that \( g^{1/2} \) is a Schwartz operator. If \( g = \sum \lambda_i |i\rangle\langle i| \rangle \) is the spectral decomposition of \( g \), we consider the following factorization:
Clearly, $Q = Q^{-1}$ commutes with $|g|$ and $|g|^{1/2}$. Moreover, from Theorem 13 (i), we will be able to use the trace cyclicity for our following computation. Using twice the Cauchy-Schwarz inequality and $\|g\|_1 \leq 3\varepsilon$, we find

$$|\Tr g R_{1k} R_{2j}|^2 = |\Tr |g|^{1/2} g R_{1k} R_{2j}|^2 \leq (\Tr |g|)(\Tr |g| R_{1k}^2 R_{2j}^2) \leq 3\varepsilon \sqrt{\Tr |g| R_{1k}^2 \Tr |g| R_{2j}^2} \leq (3\varepsilon) \max\{\Tr |g| R_{1k}^2, \Tr |g| R_{2j}^2\}.$$

Using again the decomposition of $g$, we obtain for $j = 1, 2$ that

$$|\Tr g R_{1k}^i| = |\Tr Q g R_{1k}^i | \leq \|Q\| \|g R_{1k}^i\|_1 = \|\rho_{ab} - \rho_a \otimes \rho_b\|_1 \leq 2\|\rho_{ab}\|_1,$$

since $R_{1k}^i$ is a local operator on one part of the output. Consequently, $|\Tr g R_{1k} R_{2j}|^2 \leq 6\varepsilon \max\{\Tr \rho_{ab} R_{1k}^i, \Tr \rho_{ab} R_{2j}^i\} \leq 6\varepsilon\kappa$, where $\kappa := \max\{\|\rho_{ab} R_{1k}^i\|_1, \|\xi\|_2 = \|\eta\|_2 = 1\}$ is the largest generalized fourth moment of $\rho_{ab}$. Note that since $\rho_{ab}$ is a Schwartz operator, $\kappa < \infty$. Thus,

$$|V|_2 \leq \frac{\sqrt{24n^2\varepsilon\kappa}}{\tan \theta}.$$

**APPENDIX B: UPPER BOUND OF Eq. (72)**

The line integral of a matrix $A \in \mathbb{C}^{2n \times 2n}$ is defined in terms of the line integrals of the rows of $A$. Namely, if $A_k$, $k = 1, \ldots, 2n$ are the rows of $A$, then

$$\int A \cdot d\eta := \left( \int A_1 \cdot d\eta \right) : \left( \int A_{2n} \cdot d\eta \right).$$

Thus, the line integral of a matrix-valued function is a vector, and the line integral of a vector field is a scalar. Let us denote by $M : \mathbb{R}^{2n} \to \mathbb{C}^{2n \times 2n}$ a matrix-valued function. The upper bound for Eq. (72) is equivalent to bound

$$F(\xi) := \int_{C(\xi)} \left( \int_{C(\eta)} M(z) \cdot dz \right) \cdot d\eta,$$

with $\xi, \eta \in \mathbb{B}_{1/2}$ and

$$M(z) = \cos^2 \theta \left( Q_{22}(z) + \frac{Q_{12}(z)}{\tan \theta} \right).$$

As the rows of $M(z)$ will not be conservative vector fields, we choose the curves $C(\xi)$ and $C(\eta)$ to be straight lines. Let us parametrize the curve $C(\xi)$ via $\theta : [0, 1] \to C$, $t \mapsto t\xi$ and write $M_{ik}(z)$ for the entries of the matrix $M(z) \in \mathbb{C}^{2n \times 2n}$. Define the matrix-valued function $\mathbb{R}^{2n} \times z \mapsto Y(z) \in \mathbb{C}^{2n \times 2n}$ to have the entries $Y_{ik}(z) := \int_0^1 M_{ik}(sz)\, ds$. From the explicit parametrization of the line integrals and the Cauchy-Schwarz inequality, we obtain that

$$|F(\xi)| \leq \max_{t \in [0, 1]} |t\xi \cdot Y(t\xi)\xi| \leq \|\xi\|^2 \max_{t \in [0, 1]} \left( \sum_{k=1}^{2n} |Y(t\xi)_{ik}|^2 \right)^{1/2}. \tag{B1}$$

In order to bound $|Y(t\xi)_{ik}|$, we differentiate Eq. (59) with the help of Lemma 14 to find...
\[ Q_{22}(z) = \frac{G_{2}(z) G_{2}^T(z)}{\chi_1^2(\cos \theta \beta \mathbb{I}) \chi_2^2(- \sin \theta \beta \mathbb{I})} - \frac{G_{22}(z)}{\chi_1^2(\cos \theta \beta \mathbb{I}) \chi_2^2(- \sin \theta \beta \mathbb{I})} \in \mathbb{C}^{2n \times 2n}, \]
\[ Q_{12}(z) = -\frac{G_{12}(z)}{\chi_1^2(\cos \theta \beta \mathbb{I}) \chi_2^2(- \sin \theta \beta \mathbb{I})} - \frac{\nabla \rho_{\beta}(z) G_{2}^T(z)}{\chi_1^2(\cos \theta \beta \mathbb{I}) \chi_2^2(- \sin \theta \beta \mathbb{I})} \in \mathbb{C}^{2n \times 2n}, \]

for \( z \in \mathbb{R}^{2n}, \| z \|_2 \leq r/2 \), where
\[ G_{2}(z) = \frac{i}{2} \sigma \text{Tr}[e^{iz R_1} \{ R_2, g \}] \in \mathbb{C}^{2n}, \]
\[ G_{22}(z) = -\frac{i}{4} \sigma \text{Tr}[e^{iz R_1} \{ \{ R_2, g \}, R_2^T \}] \in \mathbb{C}^{2n \times 2n}, \]
\[ G_{12}(z) = -\frac{i}{4} \sigma \text{Tr}[e^{iz R_1} \{ \{ R_1, g \}, R_2^T \}] \in \mathbb{C}^{2n \times 2n}, \]
\[ \nabla \rho_{\beta}(z) = \frac{i}{2} \text{Tr}[e^{iz R_1} \{ \{ R_1, \rho_{\beta} \}, R_2^T \}] \in \mathbb{C}^{2n}. \]

Let us write \( R_{1k}, k = 1, \ldots, 2n, \) and \( R_{2l}, l = 1, \ldots, 2n, \) for the entries of the vector \( R_1 = (Q_1, P_1, \ldots, Q_n, P_n) \) and \( R_2 = (Q_{n+1}, P_{n+1}, \ldots, Q_{2n}, P_{2n}) \), respectively. From Lemma 16, we know that for \( z \in \mathbb{B}_n, \chi(z) = 12 \epsilon^{1+n} \), and therefore, we can upper bound each entry of \( Y(\xi)_{kl} \) by
\[ \| Y(\xi)_{kl} \| \leq \cos^2 \theta \left( \frac{1}{4(12 \epsilon^{1/12})^2} + \frac{1}{4(12 \epsilon^{1/12})^2} + \frac{1}{4(12 \epsilon^{1/12})^2} + \frac{1}{4(12 \epsilon^{1/12})^2} \right). \]

Following a similar procedure as for the bound of \( \| V \|_2 \) (see Appendix A), we obtain
\[ \| Y(\xi)_{kl} \|^2 \leq \left( \frac{\kappa \cos^2 \theta}{8 \tan^2 \theta} \right)^{2/3}, \]
so from Eq. (B1)
\[ \left| \frac{i}{2} \xi \right|^2 \leq \left( \frac{n^2 \kappa | \xi |_{T}^2}{2 \tan^2 \theta} \right)^{2/3}. \]

REFERENCES

22 N. G. Ushakov, *Selected Topics in Characteristic Functions* (De Gruyter, 1999).