An operational measure for squeezing

Martin Idel, Daniel Lercher and Michael M Wolf

Zentrum Mathematik, Technische Universität München, Germany
E-mail: martin.idel@tum.de

Received 5 July 2016, revised 26 August 2016
Accepted for publication 14 September 2016
Published 13 October 2016

Abstract
We propose and analyse a mathematical measure for the amount of squeezing contained in a continuous variable quantum state. We show that the proposed measure operationally quantifies the minimal amount of squeezing needed to prepare a given quantum state and that it can be regarded as a squeezing analogue of the ‘entanglement of formation’. We prove that the measure is convex and subadditive and we provide analytic bounds as well as a numerical convex optimisation algorithm for its computation. By example, we then show that the amount of squeezing needed for the preparation of certain multi-mode quantum states can be significantly lower than naive state preparation suggests.

Keywords: squeezing, continuous variable quantum information, operational measure, Euler decomposition, bosonic systems

(Some figures may appear in colour only in the online journal)

1. Introduction
The interplay between quantum optics and the field of quantum information processing, in particular via the subfield of continuous variable quantum information, has been developing for several decades and is interesting also due to its experimental success (see [KL10] for a thorough introduction).

Coherent bosonic states and the broader class of Gaussian bosonic states, quantum states whose Wigner function is characterised by its first and second moments, are of particular interest in the theory of continuous variable quantum information. Their interest is also due to the fact that modes of light in optical experiments behave like Gaussian coherent states.

For any bosonic state, its matrix of second moments, the so called covariance matrix, must fulfil Heisenberg’s uncertainty principle in all modes. If the state possesses a mode, where despite this inequality $\Delta x \Delta p \geq \hbar / 2$ either $\Delta x$ or $\Delta p$ is strictly smaller than $\sqrt{\hbar / 2}$, it is called squeezed. The production of squeezed states is experimentally possible, but it
requires the use of nonlinear optical elements [Bra05], which are more difficult to produce and handle than the usual linear optics (i.e. beam splitters and phase shifters). Nevertheless, squeezed states play a crucial role in many experiments in quantum information processing and beyond. Therefore, it is natural both theoretically and practically to investigate the amount of squeezing which is necessary to create an arbitrary quantum state.

As a qualitative answer, squeezing is known to be an irreducible resource with respect to linear quantum optics [Bra05]. In the Gaussian case, it is also known to be closely related to entanglement of states [WEP03] and the non-additivity of quantum channel capacities [LGW13]. In addition, quantitative measures of squeezing have been provided on multiple occasions [Kra+03, Lee88], yet none of these measures are operational for more than a single mode in the sense that they do not measure the minimal amount of squeezing necessary to prepare a given state.

The goal of this paper is therefore twofold: first, we define and study operational squeezing measures, especially measures quantifying the amount of squeezing needed to prepare a given state. Second, we reinvestigate in how far squeezing is a resource in a mathematically rigorous manner and study the resulting resource theory by defining preparation measures.

In order to give a brief overview of the results, we assume the reader is familiar with standard notation of the field, which is also gathered in section 2. In particular, let γ denote covariance matrices. A squeezed state is a state where at least one of the eigenvalues of γ is smaller than one.

To obtain operational squeezing measures, we first study operational squeezing in section 3: suppose we want to implement an operation on our quantum state corresponding to some unitary \( U \). Any such unitary can be implemented as the time-evolution of Hamiltonians. Recall that any quantum-optical Hamiltonian can be split into ‘passive’ and ‘active’ parts, where the passive parts are implementable by linear optics and the active parts require nonlinear media. We assume that the active transformations available are single-mode squeezers with Hamiltonian

\[
H_{\text{squeeze},j} = \frac{i}{2} (a_j^2 - a_j^{\dagger 2}),
\]

where the \( j \) denotes squeezing in the \( j \)th mode. We therefore consider any Hamiltonian of the form

\[
H = H_{\text{passive}}(t) + \sum_k c_k(t) H_{\text{squeeze},j},
\]

where \( c_k \) are complex coefficients, which can be seen as the interaction strength of the medium and \( H_{\text{passive}} \) is an arbitrary passive Hamiltonian. Then, a natural measure of the squeezing costs to implement this Hamiltonian would be given by

\[
f_{\text{squeeze}}(H) = \int \sum_k |c_k(t)| \, dt.
\]

Our squeezing measure for the operation \( U \) is then defined as the minimum of \( f_{\text{squeeze}}(H) \) for all Hamiltonians implementing the operation \( U \) of the form (1). With this definition, we have an operational measure answering the question: given an operation \( U \), what is the most efficient way (in terms of squeezing) to implement it using passive operations and single-mode squeezers?

Instead of working with the generators, which are unbounded operators and therefore introduce a lot of analytic problems, we will work on the level of Wigner functions and therefore with the symplectic group. The unitary \( U \) then corresponds to a symplectic matrix \( S \) and we prove that the most efficient way to implement it is by using the Euler decomposition, also known as Bloch–Messiah decomposition. We show this result first in the case where the
functions $c_i$ are step functions and later on in the more general case of measurable $c$ (section 3.2). In particular, the result implies that the minimum amount of squeezing to implement the symplectic matrix $S \in \mathbb{R}^{2n \times 2n}$ is given by

$$F(S) := \sum_{i=1}^{n} \log s_i(S),$$

where $s_i$ denotes the $i$th singular value of $S$ ordered decreasingly.

With this in mind, we define a squeezing measure for preparation procedures where one starts out with a covariance matrix of an unsqueezed state and then performs symplectic (and possibly other) operations to obtain the state. More precisely, we define

$$G(\gamma) := \inf \left\{ \sum_{i=1}^{n} \log s_i(S) \left| \begin{array}{l}
\gamma \geq S^T S, \ S \in Sp(2n) \end{array} \right. \right\}.$$ 

One of the main results of this paper, which will be proven in section 5, is that this measure is indeed operational in that it quantifies the minimal amount of single-mode squeezing necessary to prepare a state with covariance matrix $\gamma$, using linear optics with single-mode squeezers, ancillas, measurements, convex combinations and addition of classical noise.

We also define a second squeezing measure, which is a squeezing-analogue of the entanglement of formation, the ‘squeezing of formation’, i.e. the amount of single-mode squeezed resource states needed to prepare a given state using only passive operations and adding of noise. This is done in section 5.3, where we also prove that this measure is equal to $G$.

In addition, we prove several structural facts about $G$ in section 4. In particular, $G$ is convex, lower semicontinuous everywhere, continuous on the interior and subadditive. Moreover, we show

$$\frac{1}{2} \sum_{\lambda_i \leq 1} n \log(\lambda_i(\gamma)) \leq G(\gamma)$$

with the eigenvalues $\lambda_i$ of $\gamma$. Equality in this lower bound is usually not achievable, albeit numerical tests have shown that the bound is often very good.

The measure would lose a lot of its appeal, if it could not be computed. Although we cannot give an efficient analytical formula for more than one mode, we provide a numerical algorithm to obtain $G$ for any state. To demonstrate that this works in principle, we calculate $G$ approximately for a state studied in [MK08] (section 6). The calculations also demonstrate that the preparation procedure obtained from minimising $G$ can greatly lower the squeezing costs when compared to naive preparation procedures. Finally, we critically discuss the flexibility and applicability of our measures in section 7. We believe that while we managed to give reasonable measures and interesting tools to study the resource theory of squeezing from a theoretical perspective, $G$ might not reflect the experimental reality in all parts. In particular, it becomes extraordinarily difficult to achieve high squeezing in a single mode [And+15], which is not reflected by taking the logarithm of the squeezing parameter. We show that this shortcoming can be easily corrected for a broad class of cost functions. In addition, the form of the active part of the Hamiltonian (1) might not reflect the form of the Hamiltonian in the lab. This cannot be corrected as easily but in any case, our measure will give a lower bound.
2. Preliminaries

In this section, we collect basic notions from continuous variable quantum information and symplectic linear algebra that we need later on. For a broader overview, we refer to [ARL14, BL05].

2.1. Phase space in quantum physics

Consider a bosonic system with $n$-modes, each of which is characterised by a pair of canonical variables $(Q_k, P_k)$. Setting $R = (Q_1, Q_2, \ldots, Q_n)$, the canonical commutation relations (CCRs) take on the form $[R_k, R_l] = i\sigma_{kl}$ with the standard symplectic form

$$\sigma = \bigoplus_{i=1}^{n} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Since it will sometimes be convenient, we also introduce another basis of the canonical variables: let $\tilde{R} = (Q_1, Q_2, \ldots, Q_n, P_1, P_2, \ldots, P_n)\}$, then the symplectic CCRs take on the form $[\tilde{R}_k, \tilde{R}_l] = i\delta_{kl}$ with the symplectic form

$$J = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}.$$

Clearly, $J$ and $\sigma$ differ only by a permutation, since $R$ and $\tilde{R}$ differ only by a permutation. From functional analysis, it is well-known that the operators $Q_k$ and $P_k$ cannot be represented by bounded operators on a Hilbert space. In order to avoid complications associated to unbounded operators, it is usually easier to work with a representation of the CCR-relations on some Hilbert space $\mathcal{H}$, instead. The standard representation is known as the Schrödinger representation and defines the Weyl system, a family of unitaries $W_\xi$ with $\xi \in \mathbb{R}^{2n}$ and

$$W_\xi = \exp(i\xi R), \quad \xi \in \mathbb{R}^{2n}$$

fulfilling the Weyl relations $W_\xi W_\eta = \exp^{-1/2\imath\eta} W_{\xi+\eta}$ for all $\xi, \eta$. Such a system is unique up to isomorphism under further assumptions of continuity and irreducibility as obtained by the Stone–von Neumann theorem. Given $W_\xi$ it is important to note that

$$W_\xi R_k W_\xi^* = R_k + \xi_k 1_n \quad \forall \xi \in \mathbb{R}^{2n}. \tag{4}$$

In this paper, we will not use many properties of the Weyl system, since instead, we can work with the much simpler moments of the state: given a quantum state $\rho \in S(L^2(\mathbb{R}^{2n}))$ (trace-class operators on $L^2$), its first and second centred moments are given by

$$d_k = \text{tr}(\rho R_k), \tag{5}$$

$$\gamma_{kl} = \text{tr}(\rho [R_k - d_k 1, R_l - d_l 1]), \tag{6}$$

with $\{\cdot, \cdot\}$ the regular anticommutator. We will write $\Gamma$ instead of $\gamma$ for the covariance matrix, if we work with $\tilde{R}$ instead of $R$. Again, a simple permutation relates the two.

An important question one can ask is when a matrix $\gamma$ can occur as a covariance matrix of a quantum state. The answer is given by Heisenberg’s principle, which here takes the form of a matrix inequality:

**Proposition 2.1.** Let $\gamma \in \mathbb{R}^{2n \times 2n}$, then there exists a quantum state $\rho$ with covariance matrix $\gamma$ if and only if

$$\gamma \succeq i\sigma,$$

with $\succeq$ the symmetric ordering relation.
where $\geq$ denotes the standard partial order on matrices (i.e. $\gamma \geq \sigma$ if $\gamma - \sigma$ is positive semidefinite). Note that we leave out the usual factor of $\hbar/2$ to simplify notation.

Another question one might ask is when a covariance matrix belongs to a pure quantum state. This question cannot be answered without more information about the higher order terms. If we however require the state to be uniquely determined by its first and second moments, i.e. if we consider the so called Gaussian states, we have an answer (see [ASI04]):

**Proposition 2.2.** Let $\rho$ be an $n$-mode Gaussian state (i.e. completely determined by its first and second moments), then $\rho$ is pure if and only if $\det(\gamma_\rho) = 1$.

### 2.2. The linear symplectic group and squeezing

A very important set of operations on a quantum system are those, that leave the CCRs invariant, i.e. linear transformations $S$ such that $[SR, SR] = i\sigma_\mathbb{R}$. Such transformations are called symplectic transformations.

**Definition 2.3.** Given a symplectic form $\sigma$ on $\mathbb{R}^{2n \times 2n}$, the set of matrices $S \subset \mathbb{R}^{2n \times 2n}$ such that $S^T\sigma S = \sigma$ is called the linear symplectic group and is denoted by $Sp(2n, \mathbb{R}, \sigma)$.

We will usually drop both $\sigma$ and $\mathbb{R}$ in the description of the symplectic group since this will be clear from the context. The linear symplectic group is a Lie group and as such contains a lot of structure. For more information on the linear symplectic group and its connection to physics, we refer the reader to [Gos06, MS98] chapter 2. An overview for physicists is also found in [Arv+95a]. All of the following can be found in that paper:

**Definition 2.4.** Let $O(2n, \mathbb{R})$ be the real orthogonal group, Then we define the following three subsets of $Sp(2n)$:

- $K(n) = Sp(2n, \mathbb{R}) \cap O(2n, \mathbb{R})$,
- $Z(n) = \{I_{2(j-1)} \oplus \text{diag}(s_j, s_j^{-1}) \oplus I_{2(n-(j+1))} | s_j \geq 0, j = 1, \ldots, n\}$,
- $\Pi(n) = \{S \in Sp(2n, \mathbb{R}) | S \geq 0\}$.

The first subset is the maximally compact subgroup of $Sp(2n)$, the second subset is the subset of single-mode-squeezers. It generates the multiplicative subgroup $\mathcal{A}(2n)$, a maximally abelian subgroup of $Sp(2n)$. The third set is the set of positive definite symplectic matrices.

In addition, since $Sp(2n)$ is a Lie group, it possesses a Lie algebra. Let us collect a number of relevant facts about the Lie algebra and some subsets:

**Proposition 2.5.** The Lie algebra $sp(2n)$ of $Sp(2n)$ is given by

$$sp(2n) = \{T \in \mathbb{R}^{2n \times 2n} | \sigma T + T \sigma = 0\}$$

together with the commutator as Lie bracket. Certain other Lie algebras or subsets of Lie algebras are of relevance to us:

1. $\mathfrak{so}(2n) = \{A \in \mathbb{R}^{2n \times 2n} | A + A^T = 0\}$ the Lie algebra of $SO(2n)$.
2. $\mathfrak{t}(n) = \{A \in \mathbb{R}^{2n \times 2n} | A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, a = -a^T, b = b^T\}$ the Lie algebra of $K(n)$.
(3) \( \pi(n) := \{ A \in \mathbb{R}^{2n \times 2n} | A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, a = a^T, b = b^T \} \) the subspace of the Lie algebra \( \mathfrak{sp}(2n) \) corresponding to \( \Pi(n) \).

Since the Lie algebra is a vector space, it is spanned by a set of vectors, the generators. A standard decomposition is given by taking the generators of \( \mathfrak{t}(n) \), the so called passive transformations as one part and the generators of \( \pi(n) \), the so called active transformations as the other part. That these two sets together determine the Lie algebra completely can be seen with the polar decomposition:

**Proposition 2.6** (Polar decomposition [Arv+95a]). For every symplectic matrix \( S \in \text{Sp}(2n) \) there exists a unique \( U \in K(n) \) and a unique \( P \in \Pi(n) \) such that \( S = UP \).

A basis for the Lie algebras \( \mathfrak{t}(n) \) and \( \pi(n) \) therefore characterises the complete Lie algebra \( \mathfrak{sp}(2n) \). Elements of the Lie algebras are also called generators and a basis of generators therefore fixes the Lie algebra. Via the polar decomposition, this implies that they also generate the whole Lie group. We will need a set of generators \( g_{ij}^{(p)} \in \mathfrak{t}(n) \) and \( g_{ij}^{(a)} \in \pi(n) \) later on, which we will fix via the metaplectic representation:

**Proposition 2.7** (Metaplectic representation [Arv+95a]). Let \( W_\xi \) be the continuous irreducible Weyl system defined above and let \( S \in \text{Sp}(2n) \). Then there exists an up to a phase unique unitary \( U_\xi \) with

\[
\forall \xi : \quad U_\xi W_\xi U_\xi^\dagger = W_\xi^C.
\]

Since we have the liberty of a phase, this is not really a representation of the symplectic group, but of its two-fold cover, the metaplectic group. We can also study the generators of this representation, which are given by \( 1/2 [R_\xi, R_\eta]_+ \).

For the reader familiar with annihilation and creation operators, if we denote by \( a_i, a_i^\dagger \) the annihilation and creation operators of the \( n \) bosonic modes, the generators of the metaplectic representation are given by

\[
G_j^{p(1)} := i(a_j^\dagger a_i - a_i^\dagger a_j), \quad G_j^{p(2)} := a_j^\dagger a_j + a_j a_j^\dagger, \quad (7)
\]

\[
G_j^{a(3)} := i(a_j^\dagger a_i^\dagger - a_i^\dagger a_j), \quad G_j^{a(4)} := a_j a_j^\dagger + a_j^\dagger a_j, \quad (8)
\]

where the \( p \) stands for ‘passive’ and the \( a \) for ‘active’. The passive generators are also frequently called linear transformations in the literature (see [Kok+07]). We can now define a set of generators of the symplectic group \( \text{Sp}(2n) \) by using the set of metaplectic generators \( G_j \) above and take corresponding generators \( g_{ij} \) in the Lie algebra \( \mathfrak{sp}(2n) \) in a consistent way. As one would expect from the name, the passive metaplectic generators correspond to a set of passive generators of \( \mathfrak{t}(n) \) and the set of active metaplectic generators corresponds to a set of active generators of \( \pi(n) \). The details of the correspondence are irrelevant (they are explicitly spelled out in equation (6.6b) in [Arv+95a]), except for the fact that the set \( G_i^{a(3)}, i = 1, \ldots, n \) corresponds to the generators \( g_{ii}^{a(3)} \) generating matrices in \( Z_n \).

Given a Hamiltonian, the associated time evolution corresponds to a path on the Lie group: for a (sufficiently regular) path \( \gamma : [0, 1] \rightarrow \text{Sp}(2n) \) we can find a function \( A(t) \in \mathfrak{sp}(2n) \) such that
Instead of directly studying Hamiltonians with time-dependent coefficients as in equation (1), it is equivalent to study functions $A : [0, 1] \to \text{sp}(2n)$.

There are a number of decompositions of the Lie group and its subgroup in addition to the polar decomposition. We will mostly be concerned with the so called Euler decomposition (sometimes called Bloch–Messiah decomposition) and Williamson’s decomposition:

**Proposition 2.8** (Euler decomposition [Arv+95a]). Let $S \in \text{Sp}(2n)$, then there exist $K, K' \in K(n)$ and $A \in A(n)$ such that $S = KAK'$.

**Proposition 2.9** (Williamson’s theorem [Wil36]). Let $M \in \mathbb{R}^{2n \times 2n}$ be a positive definite matrix, then there exists a symplectic matrix $S \in \text{Sp}(2n, \mathbb{R})$ and a diagonal matrix $D \in \mathbb{R}^{n \times n}$ such that $M = S^T \tilde{D} S$, where $\tilde{D} = \text{diag}(D, D)$ is diagonal. The entries of $D$ are also called symplectic eigenvalues.

In particular, for $M \in \Pi(n)$, this implies that $M$ has a symplectic square root. Since covariance matrices are always positive definite, this implies also that a Gaussian state is pure if and only if its covariance matrix is symplectic. Heisenberg’s uncertainty principle has also a Williamson version:

**Corollary 2.10.** A positive definite matrix $M$ is a covariance matrix of a quantum state if and only if all symplectic eigenvalues are larger or equal to one.

The proof is simple and therefore omitted.

### 2.3. Quantum optical operations and squeezing

We have already noted that an important class of operations are those, which leave the CCR-relations invariant, namely the symplectic transformations. Given a quantum state $\rho$, the action of the symplectic group on the canonical variables $R$ descends to a subgroup of unitary transformations on $\rho$ via the metaplectic representation (see [Arv+95b]). Its action on the covariance matrix $\gamma_\rho$ of $\rho$ is even easier: Given $S \in \text{Sp}(2n)$,

$$\gamma_\rho \mapsto S^T \gamma_\rho S. \quad (10)$$

In quantum optics, symplectic transformations can be implemented by the means of

1. beam splitters and phase shifters, implementing operations in $K(n)$ ([Rec+94])
2. single-mode squeezers, implementing operations in $Z(n)$.

Via the Euler decomposition, this implies that any symplectic transformation can be implemented (approximately) by a combination of those three elements.

**Definition 2.11.** An $n$-mode bosonic state $\rho$ is called squeezed, if its covariance matrix $\gamma_\rho$ possesses an eigenvalue $\lambda < 1$.

Especially in the early literature, squeezing is usually defined differently: a state $\rho$ is squeezed if there exists a unitary transformation $K \in K(n)$ such that $K^T \gamma_\rho K$ has a diagonal entry smaller than one. This again comes from the physical definition of squeezed states being...
3. An operational squeezing measure for symplectic transformations

Throughout this section, we will always use $\sigma$ as our standard symplectic form.

3.1. Definition and basic properties

We will now define a first operational squeezing measure for symplectic transformations, which will later be used to define a measure for operational squeezing.

**Definition 3.1.** Define the function $F : \mathbb{R}^{2n \times 2n} \to \mathbb{R}$

$$F(A) = \sum_{i=1}^{n} \log(s_i^1(A)),$$

where $s_i^1$ are the decreasingly ordered singular values of $A$.

Note that we sum only over half of the singular values. Restricting this function to symplectic matrices will yield an operational squeezing measure for symplectic transformations: recall that the symplectic group is generated by symplectic orthogonal matrices and single-mode squeezers. The orthogonal matrices are easy to implement and therefore will be considered a free resource. The squeezers have singular values $s$ and $s^{-1}$ and they are experimentally hard to implement and should therefore be assigned a cost that depends on the squeezing parameter $s$. Using this, the amount of squeezing seems to be characterised by the largest singular values. Here, we quantify the amount of squeezing by a cost $\log(s)$, which can be seen as the interaction strength of the Hamiltonian needed to implement the squeezing.

Let us make this more precise: define the map

$$\Delta : \text{Sp}(2n) \to \bigcup_{m \in \mathbb{N}} \text{Sp}(2n)^{m}$$

$$S \mapsto \bigcup_{m \in \mathbb{N}} \{ (S_1, \ldots, S_m) | S = S_1 \cdots S_m, S_i \in K(n) \cup Z(n) \}.$$

The image of $\Delta$ for a given symplectic matrix contains all possible ways to construct $S$ as a product of matrices from $K(n)$ or $Z(n)$. We define:

**Definition 3.2.** Let $\mathcal{F} : \text{Sp}(2n) \to \mathbb{R}$ be a map defined via

$$\mathcal{F}(S) := \log \inf \left\{ \prod_{i=1}^{m} s_i^1(S_i) \mid (S_1, \ldots, S_m) \in \Delta(S) \right\}.$$  

(12)

**Proposition 3.3.** If $S \in \text{Sp}(2n)$ then $F(S) = \mathcal{F}(S)$.

**Proof.** Let $S = K A K'$ be the Euler decomposition of $S$ with $K, K' \in K(n)$ and $A \in A(n)$. Assume without loss of generality that $A = \text{diag}(a_1, a_1^{-1}, \ldots, a_n, a_n^{-1})$ and $a_1 \geq a_2 \geq \ldots \geq a_n \geq 1$ and define $A_i = \text{diag}(1, \ldots, 1, a_i, a_i^{-1}, 1, \ldots, 1)$. By construction $A = A_1 \cdots A_n$ and $A_i \in Z(n)$. Since $K, K' \in K(n)$, $(K, A_1, \ldots, A_n, K') \in \Delta(S)$. Using that $s_i^1(K) = s_i^1(K') = 1$ and the fact that the Euler decomposition is actually equivalent to the
singular value decomposition of $S$, we obtain:

$$
F(S) \leq \log \left( s_1^i(K) \prod_{i=1}^{n} s_i^j(A_i) s_i^j(K') \right) = \log \prod_{i=1}^{n} s_i^j(S) = F(S).
$$

Conversely, consider $(S_1, \ldots, S_m) \in \Delta(S)$. Using that by definition for each $S_j \in K(n) \cup Z(n)$ we have $\prod_{i=1}^{m} s_i^j(S_j) = s_i^j(S_j)$, we conclude:

$$
F(S) = \log \left( \prod_{i=1}^{n} s_i^j(S) \right) \leq \log \left( \prod_{i=1}^{m} \prod_{j=1}^{n} s_i^j(S_j) \right) = \log \left( \prod_{j=1}^{m} s_i^j(S_j) \right),
$$

where in $(*)$ we used a special case of a theorem by Gel’fand and Naimark ([Bha96], theorem III.4.5 and equation (III.19)). Taking the infimum on the right-hand side gives $F(S) \leq F(S)$.

Let us write the last observation in $(*)$ as a small lemma for later use:

**Lemma 3.4.** Let $S, S' \in Sp(2n)$. Then $F(SS') \leq F(S) + F(S')$.

### 3.2. Lie algebraic definition

Up to now, we have only considered products of symplectic matrices, which would correspond to a chain of beam splitters, phase shifters and single-mode squeezers. The goal of this section is to prove that one cannot improve the results with arbitrary paths on $Sp(2n)$, corresponding to general Hamiltonians of the form of equation (1) as described in section 2.

Let $C'(S)$ be the set of absolutely continuous paths $\alpha : [0, 1] \to Sp(2n)$ with a derivative which is bounded almost everywhere such that $\alpha(0) = 1$ and $\alpha(1) = S$. Such paths seem to capture most if not all physically relevant cases.

Recall the set of generators $g$ of $\mathfrak{sp}(2n)$ defined in section 2 and order them in a single vector. Using equation (9), any $\alpha \in C'(S)$ corresponds to a $A \in L^\infty([0, 1], \mathfrak{sp}(2n))$. Since the generators $g$ form a basis, we can write $A(t) = c_\alpha(t) \cdot g$ with a function $c_\alpha \in L^\infty([0, 1], \mathfrak{sp}(2n))$. Both $A$ or $c_\alpha$ together with the condition $\alpha(0) = 1$ uniquely define $\alpha$.

The goal of this section is to prove that this does not give us any better way to avoid squeezing:

**Theorem 3.5.** For any $S \in Sp(2n)$, we have

$$
F(S) = \inf \left\{ \int_0^1 \| \tilde{c}_\alpha(t) \| \, dt \mid \alpha \in C'(S), \dot{\alpha}(t) = (\tilde{c}_\alpha(t) g^p(\alpha(t)), \tilde{c}_\alpha(t) g^q(\alpha(t)))^T \right\},
$$

where we introduced the notation $\tilde{c}$ to clarify that $g^p/\alpha$ are actually vectors containing a set of generators each, and the coefficients might differ for each of these generators.

The proof of this theorem is quite lengthy in details, thus we split it up into several lemmata. The general idea is easy to relate: we first show that paths corresponding to products of symplectic matrices of type $Z(n)$ or $K(n)$ produce the same outcome in (13) and (12). We then use an approximation argument: given any path, we can approximate it by a path of products of symplectic matrices to arbitrary precision.

To start, we prove the following lemma:
Lemma 3.6. Let $A \in \mathfrak{sp}(2n)$ and write $A = 1/2(A + A^T) + 1/2(A - A^T) =: A_+ + A_-$. Then $A_+ \in \pi(2n)$ and $A_- \in \mathfrak{t}(n)$ and we have $F(\exp(A)) \leq F(\exp(A_+))$.

Proof. First note that $F$ is continuous in $S$ since the singular values are. Using the Trotter-formula, we obtain:

$$F(\exp(A)) = F(\lim_{n \to \infty} (\exp(A_+/n)\exp(A_-/n))^n)$$

$$\leq \lim_{n \to \infty} (nF(\exp(A_+/n)) + nF(\exp(A_-/n)))$$

$$= \lim_{n \to \infty} nF(\exp(A_+/n)) = F(\exp(A_+)),$$

where we used that $F(\exp(A_-)) = 0$ since $A_- \in \mathfrak{t}(n)$ and in $(\ast)$, we used a version of a theorem by Gel’fand and Naimark again (see [Bha96], equation (III.20)).

Let us define yet another version of $F$ which we call $\hat{F}$ in the following way:

$$C^N(S) := \left\{ (c_1^a, c_1^b, \ldots, c_N^a, c_N^b) \left| S = \prod_{j=1}^N \exp(c_j^a g^a + \hat{c}_j^b g^b), \hat{c}_j \in \mathbb{R}^{4n^2} \right. \right\},$$

$$C(S) := \bigcup_{N \in \mathbb{N}} C^N(S),$$

$$\hat{F}(S) := \inf \left\{ \sum_i \| \hat{c}_i^a \| \left| \hat{c} \in C(S) \right. \right\}.$$

This definition is of course reminiscent of the definition of $F$ in equation (12):

Lemma 3.7. For $S \in \mathfrak{sp}(2n)$, we have $\hat{F}(S) = F(S)$.

Proof. To prove $\hat{F} \leq F$, consider the Euler decomposition $S = K_1 A_1 \ldots A_n K_2$ with $A_i \in \mathfrak{z}(n)$ and $K_1, K_2 \in \mathfrak{sp}(2n)$. Since $K(n)$ is compact, the exponential map is surjective and there exist $\hat{c}_i^a$ and $\hat{c}_i^b$ such that $\exp(\hat{c}_i^a g^a) = K_1$ and $\exp(\hat{c}_i^b g^b) = K_2$. Recall that we ordered the vector $g^a$ in such a way that the generators $g^a$ generate the matrices in $\mathfrak{z}(n)$ for $i = 1, \ldots, n$, hence we know that there exist $\hat{c}_i^a = (0, \ldots, 0, (\hat{c}_i^a)_{(i)}, \ldots, 0)$ for $i = 1, \ldots, n$ such that

$$S = \exp(\hat{c}_i^a g^a) \prod_{i=1}^n \exp(\hat{c}_i^b g^b).$$

This implies

$$\hat{F}(S) \leq \sum_i \| \hat{c}_i^a \| = \sum_i F(\exp(\hat{c}_i^a g^a)) = \sum_i F(\exp((\hat{c}_i^a)_{(i)} g_i^a)) = \sum_i \log s_i(A_i) = F(S).$$

Here we used that $(\hat{c}_i^a)_{(i)}$ is also the largest singular value of $\exp((\hat{c}_i^a)_{(i)} g_i^a) \in \mathfrak{z}(n)$, as $F(\exp((\hat{c}_i^a)_{(i)} g_i^a)) = (\hat{c}_i^a)_{(i)}$ by normalisation of $g$.

For the other direction $F \geq \hat{F}$, let $S$ be arbitrary. Let $c \in C(S)$ and consider each vector $\hat{c}_i$ separately. We drop the index $i$ for readability, since we need to consider the entries of the vector $\hat{c}_i$. To make the distinction clear, we denote the $j$th entry of the vector $\hat{c}$ by $\hat{c}_{(j)}$. Recall that the active generators are exactly those generating the positive matrices. Then:
where we basically redid the calculations we used to prove lemma 3.6, using the continuity of $F$ and the Trotter formula from matrix analysis. Until now, we have considered only one $c_i$ of $c \in C(S)$. Now, if we define $S_i = \exp(\tilde{c}_i g)$, then we have $\prod_i S_i = S$ and hence, using lemma 3.4, we find:

$$F(S) \leq \sum_i F(S_i) \leq \sum_i \|c_i\|_{i},$$

But this means $F(S) \leq \hat{F}(S)$, as we claimed. 

We can now prove the first half of the theorem:

**Lemma 3.8.** For $S \in \text{Sp}(2n)$ we have

$$F(S) \geq \inf \left\{ \int_0^1 \|\tilde{c}_n^a(t)\|_i \, dt \mid \alpha(t) \in C'(S), \, \alpha(t) \right\}.$$  

**Proof.** Let $S \in \text{Sp}(2n)$ and consider the Euler decomposition $S = K_1 S_1 \cdots S_n K_2$. We can define a function $A : [0, 1] \to \text{sp}(2n)$ via:

$$A(t) = \begin{cases} 
(n + 2) \cdot \tilde{c}_i^p g^p & t \in [0, 1/(n + 2)), \\
(n + 2) \cdot (\tilde{c}_n^a_{(i)})_{(i)} g^a & t \in [i/(2(n + 2)), (i + 1)/(n + 2)), i = 1, \ldots, n, \\
(n + 2) \cdot \tilde{c}_n^a & t \in [(n + 1)/(n + 2), 1], 
\end{cases}$$

where $(\tilde{c}_i^p, 0, \tilde{c}_n^a, \ldots, 0, \tilde{c}_n^a_{(i)}, 0)_{(i)}$ denotes the element in $C^{n+2}(S)$ for the Euler decomposition and vector indices are denoted by a subscript $i$ as before. Let $U(s, t)$ be the propagator corresponding to $A$, then for $t \in [0, 1/(n + 2))$ according to proposition A.1, since $A$ does not depend on $t$ on this interval, it is given by $U(t, s) = \exp((t - s)A)$. In particular, $U(1/(n + 2), 0) = \exp(\tilde{c}_n^a g^a) = K_2$.

Iterating the procedure above, using $U(0, 1) = U(0, 1/(n + 2)) \cdots U((n + 1)/(n + 2), 1)$, we can see that by construction, $U(0, 1) = K_1 S_1 \cdots S_n K_2 = S$. Hence $A$ defines a continuous path on $\text{Sp}(2n)$ via $U(s, t)$. We can calculate:

$$\int_0^1 \|\tilde{c}_n^a(t)\|_i \, dt = \sum_{i = 1}^n \int_{t/(n + 2)}^{(i+1)/(n + 2)} |(n + 2) \cdot (\tilde{c}_n^a_{(i+1)})_{(i)}| \, dt = \sum_{i = 1}^n |(\tilde{c}_n^a_{(i)})_{(i)}| \cdot F(S),$$

where we used that the integral over the interval $[0, 1/(n + 2))$ and $[(n + 1)/(n + 2), 1]$ is empty due to the fact that all active components are zero. In the last step, we used that for the Euler decomposition, which takes the minimum in $\hat{F}$, this value is exactly $\sum_i |(\tilde{c}_n^a)_{(i)}| = \sum_i |\tilde{c}_n^a| \cdot i$, since $(\tilde{c}_n^a)_{(j)} = 0$ for $j \neq i$. Taking the infimum on the left-hand side only decreases the value. 

J. Phys. A: Math. Theor. 49 (2016) 445304 M Igel et al
For the other direction, we need some facts about ordinary differential equations that are collected in appendix A.

**Lemma 3.9.** For $S \in Sp(2n)$ we have

$$F(S) \leq \inf \left\{ \int_0^1 \| \phi^\alpha_a(t) \|_1 \, dt \mid \alpha \in C^1(S), \dot{\alpha}(t) \right\}.$$  

(15)

**Proof.** Let $S \in Sp(2n)$ be arbitrary. Combining the proof of lemma 3.8 with proposition 3.3 and lemma 3.7 we have already proved:

$$F(S) = \inf \left\{ \int_0^1 \| \phi^\alpha_a(t) \|_1 \, dt \mid \alpha \in C^1(S), \dot{\alpha}(t) = (\phi^\alpha_a(t)g^\alpha(\alpha(t)), \phi^\alpha_a(t)g^a(\alpha(t)))^T, \text{c step fct.} \right\}.$$  

The only thing left to prove is that we can drop the step-function assumption. This will be done by a standard approximation argument: let $\tilde{F}(S)$ denote the right-hand side of equation (15). Let $\varepsilon > 0$ and consider an arbitrary $A \in L^\infty$ such that

$$\left| \int_0^1 \| \phi^\alpha_a(t) \|_1 \, dt - \tilde{F}(S) \right| < \varepsilon$$  

(16)

i.e. $A$ corresponds to a path that is close to the infimum in the definition of $\tilde{F}$. We can now approximate $c^\alpha_a$ by step-functions $c^\alpha_{a'}$ (corresponding to a function $A'$, see lemma A.2) such that

$$\left| \int_0^1 \| c^\alpha_a - c^\alpha_{a'} \|_1 \, dt \right| < \varepsilon.$$  

(17)

Using the fact that the propagators $U_A$, $U_{A'}$ are differentiable almost everywhere (proposition A.1) and absolutely continuous when one entry is fixed, we can define a function $f(s) := U_A(0, s)U_{A'}(s, t)$, which is also differentiable almost everywhere. Furthermore, the fundamental theorem of calculus holds for $f(s)$ (see [Rud87], theorems 6.10 and 7.8).

$$\frac{d}{ds}f(s) = -U_A(0, s)A(s,t)U_{A'}(s, t) + U_A(0, s)A'(s)U_{A'}(s, t)$$

almost everywhere, which implies:

$$U_{A'}(0, t) - U_A(0, t) = f(t) - f(0) = \int_0^t \frac{d}{ds}f(s) \, ds$$

$$= \int_0^t U_A(0, s)(A'(s) - A(s))U_{A'}(s, t) \, ds.$$  

Since $U$ and $g$ are bounded in $\| \|_\infty$, we obtain

$$\| U_{A'}(0, t) - U_A(0, t) \|_1 \leq M \int_0^t \| c^\alpha_{a'}(s) - c^\alpha_a(s) \|_1 \, ds \leq M \varepsilon.$$  

(18)

$M$ can explicitly be computed by the bounds given in proposition A.1.

Up to now, we have taken a path $\alpha$ to $S$ close to the infimum and approximated it by a path $\alpha'$. It is immediate by equations (16) and (17) that


\[ \int_{0}^{1} \|c_{\alpha}(t)\|_{l} \, dt - \tilde{F}(S) \leq 2\varepsilon. \]  \hspace{1cm} (19)

Since \( c_{\alpha} \in C^{N}(S') \) for some \( N \in \mathbb{N} \) and \( S' = U_{A}(0, 1) \), we would be done if \( S' = S \). To remedy this, we want to extend \( \alpha' \) to a path \( \hat{\alpha} \) such that it ends at \( S \). This is where equation (18) enters: set \( \hat{S} := U_{A}(0, 1)^{-1}U_{A}(0, 1) \), then

\[ \|\hat{S} - I\|_{l} \leq \frac{M}{\|S\|_{l}} \varepsilon \]  \hspace{1cm} (20)

hence \( \hat{S} \approx I \) for \( \varepsilon \) small enough. Using the polar decomposition, we can write \( \hat{S} = \exp(\hat{c}_{N+1}^{p})\exp(\hat{c}_{N+2}^{n}) \). A quick calculation yields

\[ \| \log \hat{S} \| \leq n \log \left( \frac{M}{\|S\|_{l}} \right) \leq n\varepsilon \frac{M}{\|S\|_{l}} = C\varepsilon. \]  \hspace{1cm} (21)

This lets us construct a new \( \hat{A} : [0, 2] \rightarrow \text{sp}(2n) \):

\[ t \mapsto \begin{cases} A'(t) & t \in [0, 1], \\ 2 \cdot \hat{c}_{N+1}^{p} & t \in (1, 3/2), \\ 2 \cdot \hat{c}_{N+2}^{n} & t \in (3/2, 2]. \end{cases} \]

By construction, for the corresponding propagator we have \( U_{\hat{A}}(0, 2) = S \) and \( \hat{\alpha} \) is a valid path for \( \hat{F}(S) \) (at least after reparameterisation) fulfilling:

\[ \int_{0}^{1} \|c_{\hat{\alpha}}^{n}(t)\| \, dt - \hat{F}(S) \leq \int_{0}^{1} \|c_{\alpha}^{n}(t)\| \, dt - \tilde{F}(S) + \int_{0}^{1} \|c_{\alpha}^{n}(t)\| \, dt \leq (19) + (21) \leq (2 + C)\varepsilon. \]

Since, \( c_{\alpha} \in C^{N+2}(S) \), \( \hat{\alpha} \) is a valid path for \( \hat{F}(S) \), which implies that for any \( \varepsilon > 0 \), choosing \( \varepsilon := \varepsilon / (2 + C) \), we have seen:

\[ \hat{F}(S) < \tilde{F}(S) + \varepsilon. \]  \hspace{1cm} (22)

For \( \varepsilon \to 0 \), \( \hat{F}(S) \leq \tilde{F}(S) \), which implies the lemma via lemma 3.7. \( \Box \)

### 4. A mathematical measure for squeezing of arbitrary states

Throughout this section, for convenience, we will switch to using \( J \) as symplectic form. Having defined the measure \( F \), we will now proceed to define a squeezing measure for creating an arbitrary (mixed) state:

**Definition 4.1.** Let \( \rho \) be an \( n \)-mode bosonic quantum state with covariance matrix \( \Gamma \). We then define:

\[ G(\rho) \equiv G(\Gamma) := \inf \{ F(S) | |\Gamma| \geq S^{T}S, S \in \text{Sp}(2n) \}. \]  \hspace{1cm} (23)

Note that \( G \) is always finite: for any given covariance matrix \( \Gamma \), by Williamson’s theorem and corollary 2.10, we can find \( S \in \text{Sp}(2n) \) and \( D \geq 1 \) such that \( \Gamma = S^{T}DS \geq S^{T}S \). Furthermore \( G \) is also non-negative since \( F \) is non-negative for symplectic \( S \). We will prove in section 5 that this is indeed an operational measure.
4.1. Different reformulations of the measure

We will now give several reformulations of the squeezing measure and prove some of its properties. In particular, $G$ is convex and one of the crucial steps towards proving convexity of $G$ is given by a reformulation of $G$ with the help of the Cayley transform. For the reader unfamiliar with the Cayley transform, a definition and basic properties are provided in appendix B.

**Proposition 4.2.** Let $\Gamma \geq iJ$ and $\Gamma \in \mathbb{R}^{2n \times 2n}$ symmetric. Then:

\[
G(\Gamma) = \inf \{ F(\Gamma_i^{1/2}) | \Gamma \geq \Gamma_0 \geq iJ \}
\]

\[
= \inf \left\{ \frac{1}{2} \sum_{i=1}^{n} \log \left( \frac{1 + s_i(A + iB)}{1 - s_i(A + iB)} \right) | C^{-1}(\Gamma) \geq H, H \in \mathcal{H} \right\},
\]

where $\mathcal{H}$ is defined via:

\[
\mathcal{H} = \left\{ H = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \in \mathbb{R}^{2n \times 2n} \mid A^T = A, B^T = B, \text{spec}(H) \subset (-1, 1) \right\}.
\]

**Proof.** First note that the infimum in all three expressions is actually attained. We can see this most easily in the definition (23): the matrix inequalities $\Gamma \geq S^2 S (\geq iJ)$ imply that the set of feasible $S$ in the minimisation is compact, hence its minimum is attained. To see $\Gamma \geq S^2 S \geq iJ$, first note that $\Gamma \geq 1$ since any $S \in Sp(2n)$ also fulfills $S^2 S \geq iJ$, hence $\Gamma \geq S^2 S \geq iJ$. For equality, note that for any $\Gamma \geq \Gamma_0 \geq iJ$, using Williamson’s theorem we can find $S \in Sp(2n)$ and a diagonal $D \geq 1$ (via corollary 2.10) such that $\Gamma_0 = S^2 DS \geq S^2 S \geq iJ$. But since $F(\Gamma_0^{1/2}) \geq F((S^2 S)^{1/2}) = F(S)$ via the Weyl monotonicity principle, the infimum is achieved on symplectic matrices.

Finally, let us prove equality with (25). First observe that we can replace $Sp(2n)$ by $\mathcal{H}$ using proposition B.1(4).

Using the fact that $s_i(S) = \lambda_i((S)^{1/2}) = \lambda_i(C(H))$ and the fact that $H$ is diagonalised by the same unitary matrices as $C(H) = (I + H) \cdot (I - H)^{-1}$ whence its eigenvalues are

\[
\lambda_i^1(C(H)) = \frac{1 + \lambda_i^1(H)}{1 - \lambda_i^1(H)},
\]

we have:

\[
\inf \{ F(S) | \Gamma \geq S^2 S, S \in Sp(2n) \} = \inf \left\{ \log \prod_{i=1}^{n} \left( \frac{1 + \lambda_i^1(H)}{1 - \lambda_i^1(H)} \right)^2 | \Gamma \geq C(H), H \in \mathcal{H} \right\}.
\]

Next we claim $\lambda_i^1(H) = s_i(A + iB)$ for $i = 1, \ldots, n$. To see this note:

\[
\frac{1}{2} \begin{pmatrix} 1 & i \beta \\ 1 & -i \beta \end{pmatrix} \cdot \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \cdot \begin{pmatrix} 1 & -i \beta \\ 1 & i \beta \end{pmatrix} = \begin{pmatrix} 0 & A + iB \\ A - iB & 0 \end{pmatrix}.
\]

The singular values of the matrix on the right-hand side of equation (27) are the eigenvalues of $\text{diag}(A + iB)(A + iB)^{1/2}$, which are the singular values of $A + iB$ with double multiplicity. From the structure of $H$, it is immediate that the eigenvalues of the right-hand side of equation (27) and thus of $H$ come in pairs $\pm s_i(A + iB)$. Hence
\( \lambda_i^j(H) = s_i^j(A + iB) \) for \( i = 1, \ldots, n \) and we have:

\[
\inf \{ F(S) | \Gamma \geq S^T S, S \in \text{Sp}(2n) \} = \inf \left\{ \frac{1}{2} \sum_{i=1}^n \log \left( \frac{1 + s_i(A + iB)}{1 - s_i(A + iB)} \right) | \Gamma \geq \mathcal{C}(H), H \in \mathcal{H} \} \right\}
\]

To see that the right-hand side equals (25), we only need to use the fact that \( \mathcal{C}(H) \equiv \mathcal{C}^{-1}(\Gamma) \geq H \) for all \( H \in \mathcal{H} \) and \( \Gamma \geq iJ \) since the Cayley transform and its inverse are operator monotone.

### 4.2. Convexity

The reformulation (25) will allow us to prove:

**Theorem 4.3.** \( G \) is convex on the set of covariance matrices \( \{ \Gamma \in \mathbb{R}^{2n \times 2n} | \Gamma \geq iJ \} \).

The crucial part of the proof is the following lemma:

**Lemma 4.4.** Consider the map \( f : \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \to \mathbb{R} \):

\[
f(A, B) = \frac{1}{2} \sum_{i=1}^n \log \left( \frac{1 + s_i(A + iB)}{1 - s_i(A + iB)} \right)
\]

If we restrict \( f \) to symmetric matrices \( A \) and \( B \) such that \( s_i(A + iB) < 1 \) for all \( i = 1, \ldots, n \), \( f \) is jointly convex in \( A, B \), i.e.

\[
f(tA + (1-t)A', tB + (1-t)B') \leq tf(A, B) + (1-t)f(A', B') \quad \forall t \in [0, 1].
\]

**Proof.** Let \( \tilde{A} := tA + (1-t)A' \) and \( \tilde{B} := tB + (1-t)B' \). Note that \( \tilde{A} \) and \( \tilde{B} \) are also symmetric, and the largest singular value of \( A + iB \) fulfills \( s_i(A + iB) \leq s_i(\tilde{A} + i\tilde{B}) \). Therefore, the singular values of any convex combination of \( A + iB \) and \( A' + iB' \) also lie in the interval \( [0, 1] \). This makes our restriction well-defined under convex combinations.

For any \( j = 1, \ldots, n \), by Thompson’s theorem (see [Tho76]), which states that for every complex \( A, B \), there exist unitary matrices \( U, V \) such that \( |A + B| \leq U|A|U^* + V|B|V^* \), we have

\[
s_j(\tilde{A} + i\tilde{B}) = \lambda_j(|\tilde{A} + i\tilde{B}|) \leq \lambda_j(U|t(A + iB)|U^* + V|(1-t)(A' + iB')|V^*)
\]

Using Lidskii’s theorem ([Bha96], chapter III with explicit formulation in exercise III.4.3), we have

\[
s_j(\tilde{A} + i\tilde{B}) \leq \lambda_j(U|t(A + iB)|U^*) + \sum_{\pi} p_{\pi} \lambda_{\pi j}(V|(1-t)(A' + iB')|V^*)
\]

with \( p_{\pi} \geq 0 \) and \( \sum_{\pi} p_{\pi} = 1 \). In (29), we used that unitaries do not change the spectrum. Now each summand in equation (28) is the Cayley transform of a singular value. We can use the
log-convexity of the Cayley-transform to prove the joint convexity of $f$:

$$f(\bar{\mathbf{A}}, \bar{\mathbf{B}}) = \sum_{i=1}^{n} \log \mathcal{C}[\mathcal{C}(\bar{A} + i\bar{B})]$$

$$\leq \sum_{i=1}^{n} \log \left[ t \log \mathcal{C}[\lambda_i(A + iB)] + (1 - t) \sum_{\pi} p_{\pi} \mathcal{C}[\lambda_{\pi(i)}(A' + iB')] \right]$$

$$= \sum_{i=1}^{n} \log \mathcal{C}[\lambda_i(A + iB)] + (1 - t) \sum_{\pi} p_{\pi} \left( \sum_{i=1}^{n} \log \mathcal{C}[\lambda_{\pi(i)}(A' + iB')] \right)$$

$$\leq tf(A, B) + (1 - t)f(A', B'),$$

where in $(**)$ we use that the sum of all eigenvalues is of course not dependent on the order of the eigenvalues.

This lemma will later allow us to calculate $G$ as a convex programme.

**Proof of theorem 4.3.** We can now finish the proof of the convexity of $G$.

First note that using the definition of $f$ in lemma 4.4 we can reformulate (25) to

$$G(\Gamma) = \inf \{ f(A, B) | \mathcal{C}^{-1}(\Gamma) \geq H, H \in \mathcal{H} \}. \quad (30)$$

Let $\Gamma \geq \mathcal{I}$, $\Gamma' \geq \mathcal{I}$ be two covariance matrices and let $H, H' \in \mathcal{H}$ be the matrices that attain the minimum of $G(\Gamma)$, $G(\Gamma')$ respectively. Then, in particular, $tH + (1 - t)H' \in \mathcal{H}$.

Furthermore, since $\mathcal{C}^{-1}(\Gamma) \geq H$ and $\mathcal{C}^{-1}(\Gamma') \geq H'$ we have

$$\mathcal{C}^{-1}(t\Gamma + (1 - t)\Gamma') \overset{(**)}{=} t\mathcal{C}^{-1}(\Gamma) + (1 - t)\mathcal{C}^{-1}(\Gamma') \geq tH + (1 - t)H',$$

where we used the operator concavity of $\mathcal{C}^{-1}$ in $(*)$. This means that $tH + (1 - t)H'$ is a feasible matrix for the minimisation in $G$, which implies using equation (30)

$$G(t\Gamma + (1 - t)\Gamma') \leq f(tA + (1 - t)A', tB + (1 - t)B').$$

The convexity now follows directly from lemma 4.4 and the fact that we chose $H$ and $H'$ to attain $G(\Gamma)$ and $G(\Gamma')$.

4.3. Continuity properties

From the convexity of $G$ on the set of covariance matrices, it follows from general arguments in convex analysis that $G$ is continuous on the interior of the set of covariance matrices (see [Roc97], theorem 10.1). What more can we say about the boundary?
Theorem 4.5. $G$ is lower semicontinuous on the set of covariance matrices \{\Gamma \in \mathbb{R}^{2n \times 2n} | \Gamma \succeq iJ\} and continuous on its interior. Moreover, $G(\Gamma + \varepsilon I) \rightarrow G(\Gamma)$ for $0 < \varepsilon \rightarrow 0$ for any $\Gamma \succeq iJ$.

The ultimate goal is to extend continuity from the interior to the exterior, which we do not know how to do at present. The proof will need a few notions from set-valued analysis that we review in appendix C.

Proof of theorem 4.5. As already observed, $G$ is continuous on the interior. Let $\Gamma_0 \succeq iJ$ be arbitrary and suppose

$$A(\Gamma) = \{\hat{\Gamma} | (\Gamma - 2\|\Gamma_0\|I) \leq \hat{\Gamma} \leq \Gamma\}.$$ 

By definition, $A$ is compact and convex for any $\Gamma$. Moreover, it defines a set-valued function on the set of covariance matrices with non-empty values. Let $\varepsilon > 0$, then for all $\Gamma_0 \succeq iJ$ with $\|\Gamma - \Gamma_0\| < \varepsilon$, we have that for any $\hat{\Gamma} \in A(\Gamma)$, $\hat{\Gamma} = \hat{\Gamma} + (\Gamma - \Gamma_0) \in A(\Gamma_0)$ and $\|\Gamma - \hat{\Gamma}\| < \varepsilon$. This is the condition in lemma C.2 hence the set-valued function defined by $A$ is upper semicontinuous at $\Gamma_0$, which implies that $A(\Gamma) \cap \{X|iJ \leq X\}$ is also upper semicontinuous by proposition C.3. If $\varepsilon$ is small enough (e.g. $\varepsilon < 1$), this implies

$$A(\Gamma) \cap \{X|iJ \leq X\} = \{X|iJ \leq X \leq \Gamma\} = \mathcal{G}(\Gamma),$$

hence this set is upper semicontinuous at $\Gamma_0$.

Since $F$ is continuous on positive definite matrices, it is absolutely continuous if we restrict to a small neighbourhood of the covariance matrix $\Gamma_0$. This means that for every $\varepsilon > 0$ there exists an $\delta > 0$ such that

$$F(\hat{\Gamma}) - \varepsilon < F(\hat{\Gamma}) < F(\hat{\Gamma}) + \varepsilon$$

for all $\|\hat{\Gamma} - \hat{\Gamma}\| < \varepsilon$ and all $\hat{\Gamma}, \hat{\Gamma} \in \bigcup_{\|\Gamma - \Gamma_0\| < \delta} A(\Gamma)$.

Assuming without loss of generality that $\|\Gamma - \Gamma_0\| < 1$, the set $\mathcal{G}(\Gamma)$ is exactly the set for the minimisation in the definition of $G$. The upper semicontinuity of $\mathcal{G}(\Gamma)$ implies by lemma C.2 that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $\|\Gamma - \Gamma_0\| < \delta$ we have: for all $\hat{\Gamma} \in \mathcal{G}(\Gamma)$ there exists a $\hat{\Gamma} \in \mathcal{G}(\Gamma_0)$ such that $\|\hat{\Gamma} - \hat{\Gamma}\| < \varepsilon$. In particular, this is true for all minimisers $\hat{\Gamma}$ with $G(\Gamma) = F(\hat{\Gamma})^{1/2}$, where $\hat{\Gamma}$ and $\hat{\Gamma} \in \bigcup_{\|\Gamma - \Gamma_0\| < \delta} A(\Gamma)$. Using equation (31) we obtain: for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $\|\Gamma - \Gamma_0\| < \delta$, we have a pair $\hat{\Gamma}, \hat{\Gamma}$ with $\hat{\Gamma} \in \mathcal{G}(\Gamma_0)$ minimising $G(\Gamma)$ and $\hat{\Gamma} \in \mathcal{G}(\Gamma_0)$ such that

$$F(\hat{\Gamma}) - \varepsilon < F(\hat{\Gamma}) = G(\Gamma).$$

This implies that for all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$G(\Gamma_0) \leq G(\Gamma) + \varepsilon$$

for all $\|\Gamma - \Gamma_0\| < \delta$.

Taking the limit inferior on both sides implies that $G$ is lower semicontinuous at $\Gamma_0$. Upper semicontinuity would follow for instance if $\mathcal{G}(\Gamma_0)$ is also lower semicontinuous.

Finally, let us prove that $G(\Gamma_0 + \varepsilon I) \rightarrow 0$ for $\varepsilon \rightarrow 0$. To see this, consider the closed sets

$$C_n = \bigcup_{0 \leq \zeta \leq 1/n} \mathcal{G}(\Gamma_0 + \zeta I)$$

for any $n \in \mathbb{N}$. It is easy to see that $C_{n+1} \subseteq C_n$ and that $\bigcap_{n \in \mathbb{N}} C_n = \mathcal{G}(\Gamma_0)$. Moreover, $C_n$ is compact. Now let $\Gamma_n$ be the sequence of minimisers for $G(\Gamma_0 + 1/n I)$, then $\Gamma_n \in C_n$ for all $n \in \mathbb{N}$. By compactness, a subsequence will converge to
Therefore, \( G(\Gamma_0) \leq \lim_{\varepsilon \to 0} G(\Gamma_0 + \varepsilon I) \), but since \( \Gamma_0 \leq \Gamma_0 + \varepsilon I \) for all \( \varepsilon > 0 \) we also have \( G(\Gamma) \geq \lim_{\varepsilon \to 0} G(\Gamma_0 + \varepsilon I) \).

\[ \square \]

### 4.4. Additivity properties

Now we consider additivity properties of \( G \). We switch our basis again and use \( \gamma \) and \( \sigma \).

**Proposition 4.6.** For any covariance matrices \( \gamma_A \in \mathbb{R}^{2n_A \times 2n_A} \) and \( \gamma_B \in \mathbb{R}^{2n_B \times 2n_B} \), we have

\[
\frac{1}{2} (G(\gamma_A) + G(\gamma_B)) \leq G(\gamma_A + \gamma_B) \leq G(\gamma_A) + G(\gamma_B).
\]

In particular, \( G \) is subadditive.

**Proof.** For subadditivity, let \( S^T S \leq \gamma_A \) and \( S'^T S' \leq \gamma_B \) obtain the minimum in \( G(\gamma_A) \) and \( G(\gamma_B) \) respectively. Then \( S \oplus S' \) is symplectic and \( (S \oplus S')^T (S \oplus S') \leq \gamma_A + \gamma_B \) hence, \( G(\gamma_A + \gamma_B) \leq G(A) + G(B) \).

To prove the lower bound, we need the following equation that we will only prove later on (see equation (46)): 

\[
a \geq 1 : \quad G(\gamma_A) \leq G(\gamma_A + a 1_{n_A}). \quad (32)
\]

Assuming this inequality, let \( a \geq 1 \) be such that \( a 1_{n_A} \geq \gamma_B \), then

\[
G(\gamma_A + a 1_{n_A}) \leq G(\gamma_A + \gamma_B)
\]

hence \( G(\gamma_A) \leq G(\gamma_A + \gamma_B) \) and since we can do the same reasoning for \( \gamma_B \), we have \( G(\gamma_A) + G(\gamma_B) \leq 2G(\gamma_A + \gamma_B) \).

\[ \square \]

We do not know whether \( G \) is also superadditive, which would make it additive. At present, we can only prove:

**Corollary 4.7.** Let \( \gamma_A \in \mathbb{R}^{2n_A \times 2n_A} \) and \( \gamma_B \in Sp(2n_B) \), be two covariance matrices (i.e. \( \gamma_B \) is a covariance matrix of a pure state). Then \( G \) is additive.

**Proof.** Subadditivity has already been proven in the lemma. For superadditivity, we use the second reformulation of the squeezing measure in equation (24); note that there is only one matrix \( \gamma_B \geq \gamma \geq i\sigma \), namely \( \gamma_B \) itself. Now write

\[
\gamma_A \oplus \gamma_B \geq \begin{pmatrix} \tilde{A} & C \\ C^T & \tilde{B} \end{pmatrix} \geq i\sigma
\]

for \( \tilde{A} \in \mathbb{R}^{2n_A \times 2n_A} \) and \( \tilde{B} \in \mathbb{R}^{2n_B \times 2n_B} \). Then in particular \( \gamma_B - \tilde{B} \geq 0 \), but also \( \tilde{B} \geq i\sigma \), hence \( \gamma_B \geq \tilde{B} \geq i\sigma \) and therefore \( \tilde{B} = \gamma_B \). But then

\[
\gamma_A \oplus \gamma_B = \begin{pmatrix} \tilde{A} & C \\ C^T & \tilde{B} \end{pmatrix} = \begin{pmatrix} \gamma_A - \tilde{A} & C \\ C^T & \tilde{B} \end{pmatrix}
\]

hence also \( C = 0 \) and the matrix that takes the minimum in \( G(\gamma_A + \gamma_B) \) must be block-diagonal. Then \( \gamma_A \oplus \gamma_B \geq \tilde{A} \oplus \gamma_B \geq 0 \) and \( \tilde{A} \) is in the feasible set of \( G(\gamma_A) \). 

\[ \square \]
Corollary 4.8. For any covariance matrices $\gamma_A \in \mathbb{R}^{2n \times 2n}$ and $\gamma_B \in \mathbb{R}^{2n \times 2n}$,

$$ G(\gamma_A) + G(\gamma_B) \leq 2G\left( \begin{pmatrix} \gamma_A & C \\ C^T & \gamma_B \end{pmatrix} \right) $$

If $G$ is superadditive, then this inequality holds without the factor of two.

Proof.

$$ G(\gamma_A) + G(\gamma_B) \leq 2G\left( \begin{pmatrix} \gamma_A & 0 \\ 0 & \gamma_B \end{pmatrix} \right) = 2G\left( \begin{pmatrix} \frac{1}{2} \gamma_A & \frac{1}{2} C \\ \frac{1}{2} C^T & \frac{1}{2} \gamma_B \end{pmatrix} \right) $$

$$ \leq G\left( \begin{pmatrix} \gamma_A & C \\ C^T & \gamma_B \end{pmatrix} \right) + G\left( \begin{pmatrix} \gamma_A & -C \\ -C^T & \gamma_B \end{pmatrix} \right) \overset{(**)\text{ (**)}}{=} 2G\left( \begin{pmatrix} \gamma_A & C \\ C^T & \gamma_B \end{pmatrix} \right). $$

Here we used proposition 4.6 and then convexity of $G$ in $(\ast)$. Finally, in $(**)$ we used that for every

$$ \begin{pmatrix} \gamma_A & C \\ C^T & \gamma_B \end{pmatrix} \succ \begin{pmatrix} S_A & \hat{C} \\ \hat{C}^T & S_B \end{pmatrix} \begin{pmatrix} S_A & \hat{C} \\ \hat{C}^T & S_B \end{pmatrix}^T \in \text{Sp}(2(n_1 + n_2)) \quad (33) $$

we also have:

$$ \begin{pmatrix} \gamma_A & -C \\ -C^T & \gamma_B \end{pmatrix} \succ \begin{pmatrix} S_A & -\hat{C} \\ -\hat{C}^T & S_B \end{pmatrix} \begin{pmatrix} S_A & -\hat{C} \\ -\hat{C}^T & S_B \end{pmatrix}^T \in \text{Sp}(2(n_1 + n_2)) \quad (34) $$

and vice versa. Since the two matrices on the right-hand side of equations (33) and (34) have equal spectrum, the two squeezing measures of the matrices on the left-hand side need to be equal.

4.5. Bounds

Let us give a few simple bounds on $G$.

Proposition 4.9 (Spectral bounds). Let $\Gamma \succeq ij$ be a valid covariance matrix and $\lambda^i(\Gamma)$ be the vector of eigenvalues in decreasing order. Then:

$$ -\frac{1}{2} \sum_{\lambda^i(\Gamma) < 1} \log(\lambda^i(\Gamma)) \leq G(\Gamma) \leq \frac{1}{2} \sum_{i=1}^n \log \lambda^i(\Gamma) = F(\Gamma^{1/2}). \quad (35) $$

Proof. According to the Euler decomposition, a symplectic positive definite matrix has positive eigenvalues that come in pairs $s, s^{-1}$ and we can find $O \in SO(2n)$ such that for any $S^{-1}S \leq \Gamma$

$$ O^T\Gamma O \succeq \text{diag}(s_1, \ldots, s_n, s_1^{-1}, \ldots, s_n^{-1}). $$

But then, $\lambda^i(\Gamma) \geq \lambda^i(\text{diag}(s_1, \ldots, s_n, s_1^{-1}, \ldots, s_n^{-1}))$ via the Weyl inequalities $\lambda^i(A) \geq \lambda^i(B)$ for all $i$ and $A \succeq B \succeq 0$ (see [Bha96], theorem III.2.3). This implies:
\[ G(\Gamma) \leq \sum_{i=1}^{n} \log(\max \{s_i, s_i^{-1}\}) \leq \sum_{i=1}^{n} \log \lambda_i(\Gamma)^{1/2}. \]

For the lower bound, given an optimal matrix \( S \) with eigenvalues \( s_i \), we have
\[ G(\Gamma) = \sum_i \max\{s_i, s_i^{-1}\}. \]

If \( S^T S = O^T \text{diag}(s_1^2, \ldots, s_n^2, s_1^{-2}, \ldots, s_n^{-2})O \) with \( O \in SO(2n) \) is the diagonalisation of \( S^T S \), we can write:
\[ O^{-T} \Gamma^{-1} O^{-1} \leq \text{diag}(s_1^2, \ldots, s_n^2, s_1^{-2}, \ldots, s_n^{-2}) \]
and again by Weyl’s inequalities, we can find for all \( k \leq n \):
\[ -\frac{1}{2} \sum_{i=2n-k+1}^{2n} \log(\lambda_i(\Gamma)) \leq \frac{1}{2} \sum_{i=1}^{k} \log \lambda_i(\text{diag}(s_1^2, \ldots, s_n^2, s_1^{-2}, \ldots, s_n^{-2})) \leq G(\Gamma). \tag{36} \]

Now, \( -\frac{1}{2} \sum_{i=2n-k+1}^{2n} \lambda_i(\Gamma) \) can be upper bounded by restricting to eigenvalues \( \lambda_i(\Gamma) < 1 \). This implies
\[ -\frac{1}{2} \sum_{\lambda_i(\Gamma)<1} \log(\lambda_i(\Gamma)) \leq G(\Gamma) \]
using that the number of eigenvalues \( \lambda_i(\Gamma) < 1 \) can at most be \( n \) (hence \( k \leq n \) in the inequality of equation (36)), since \( \Gamma \geq S^T S \) and \( S^T S \) has at least \( n \) eigenvalues bigger than one.

Numerics suggest that the lower bound is often very good for low dimensions. In fact, it can sometimes be achieved:

**Proposition 4.10.** Let \( \Gamma \geq i \) be a covariance matrix, then \( G \) achieves the lower bound in equation (35) if there exists an orthonormal eigenvector basis \( \{v_i\}_{i=1}^{2n} \) of \( \Gamma \) with \( v_i^T J v_j = \delta_{i+n,j} \). Conversely, if \( G \) achieves the lower bound, then \( v_i^T J v_j = 0 \) for all normalised eigenvectors \( v_i, v_j \) of \( \Gamma \) with \( \lambda_i, \lambda_j < 1 \).

**Proof.** Suppose that the lower bound in equation (35) is achieved. Via Weyl’s inequalities (see [Bha96] theorem III.2.3), for all \( S^T S \leq \Gamma \) in the definition of \( G \) we have \( \lambda_i(S^T S) \leq \lambda_i(\Gamma) \). For the particular \( S \) achieving \( G \), this implies that for all \( \lambda_i(\Gamma) < 1 \) we have \( \lambda_i(S^T S) = \lambda_i(\Gamma) \). But then \( \Gamma \geq S^T S \) implies that \( S^T S \) and \( \Gamma \) share all eigenvectors to the smallest eigenvalue. Iteratively, every eigenvector of \( \Gamma \) with \( \lambda_i(\Gamma) < 1 \) must be an eigenvector of \( S^T S \) with the same eigenvalue.

Since the matrix diagonalising \( S^T S \) also diagonalises \( C^{-1}(S^T S) \), the eigenvectors of the two matrices are the same. Now, since \( C^{-1}(S^T S) \in \mathcal{H} \) by reformulation (25), for any eigenvector \( v_i \) of any eigenvalue \( \lambda_i \) \( C^{-1}(\lambda_i) < 0 \), \( Jv_i \) is also an eigenvector of \( C^{-1}(S^T S) \) to the eigenvalue \( -C^{-1}(\lambda_i) \), implying \( v_i^T J v_i = 0 \) for all \( i, j \). By definition, this means that \( \{v_i, Jv_i\} \) forms a symplectic basis. Above, we already saw that the eigenvectors of \( \Gamma \) for \( \lambda_i(\Gamma) < 1 \) are also eigenvectors of \( S^T S \), hence \( v_i^T J v_i = 0 \) for all \( i \) such that \( \lambda_i(\Gamma) < 1 \).

Conversely, suppose we have an orthonormal basis \( \{v_i\}_{i=1}^{2n} \) such that \( v_i^T J v_j = \delta_{i+n,j} \) (modulo \( 2n \) if necessary) for all eigenvectors of \( \Gamma \), i.e. \( \Gamma \) is diagonalisable by a symplectic orthonormal matrix \( \tilde{O} \in U(n) \). Then
\[ \tilde{\Omega} \Omega^T = \text{diag}(\lambda_1, \ldots, \lambda_{2n}). \]

Since \( \Gamma \geq iJ \) we have \( \lambda_1, \lambda_2 \geq 1 \). Assume that \( \lambda_i \geq \lambda_{n+i} \) for all \( i = 1, \ldots, n \) and the \( \lambda_{n+i} \) are ordered in decreasing order. Then \( \lambda_{n+r} < 1 \leq \lambda_{n+r-1} \) for some \( r \leq n \) and

\[ S^T S = \tilde{\Omega}^T \text{diag}(1, \ldots, 1, \lambda_1^{-1}, \ldots, \lambda_{n-1}^{-1}, 1, \ldots, 1, \lambda_{n-1}, \ldots, \lambda_{2n}) \tilde{\Omega} \]

fulfills \( S^T S \leq \Gamma \) and obviously achieves the lower bound in equation (35).

\[ \text{Proof.} \quad \text{Since } \Gamma \geq iJ \text{ via } \Gamma \geq iJ, \text{ the upper bound follows directly from the definition. Also, } F(S) \leq F((\Gamma^2)^{1/2}), \text{ which makes this bound trivially better than the spectral upper bound in equation (35).} \]

The lower bound follows from:

\[ G(\Gamma) \geq \frac{1}{2} \log \left( \prod_{i=n+1}^{2n} \lambda_i^2(\Gamma)^{-1} \right) = \frac{1}{2} \log \left( \prod_{i=1}^{n} \lambda_i^2(\Gamma) \right) \]

\[ = F(\Gamma^2) - \log(\det(\Gamma)^{1/2}) \geq F((S^T S)^{1/2}) - \log(\sqrt{\det(\Gamma)}) \]

\[ = F(S) - \log(\sqrt{\det(\Gamma)}) \]

using Weyl’s inequalities once again, implying that since \( S^T S \leq \Gamma \), we also have \( F(S)^2 = F(S^T S) \leq F(\Gamma) \).

\[ \text{The upper bound here can also be arbitrarily bad. One just has to consider } \Gamma = S^T (N \cdot I) S \text{ with } S^2 = \text{diag}(N - 1, \ldots, N - 1, (N - 1)^{-1}, \ldots, (N - 1)^{-1}) \in \text{Sp}(2n). \]

Then \( \Gamma \geq I \), i.e. \( G(\Gamma) = 0 \), but \( F(S) \rightarrow \infty \) for \( N \rightarrow \infty \).

\[ \text{Proposition 4.12.} \quad \text{Let } \Gamma \geq iJ \text{ be a covariance matrix. Then} \]

\[ G(\Gamma) \geq \frac{1}{4} \inf \{ ||\gamma_0||; |\log \Gamma \geq \gamma_0, \gamma_0 \in \pi(n) \}, \]

where \( \pi(n) \) was defined in proposition 2.5 as the Lie algebra of the positive semidefinite symplectic matrices. This infimum can be computed efficiently as a semidefinite programme.

\[ \text{Proof.} \quad \text{Recall that the logarithm is operator monotone on positive definite matrices. Using this, we have:} \]

21
The last step is valid, because the eigenvalues of matrices $\gamma_0 \in \pi(n)$ come in pairs $\pm \lambda_i$. Since the sum of all the singular values is just the trace-norm, we are done.

It remains to see that this can be computed by a semidefinite programme. First note that since the matrices $H \in \pi(n)$ are those symmetric matrices with $HH^\dagger = 0$, the constraints are already linear semidefinite matrix inequalities. The trace norm is an SDP by standard reasoning \cite{RFP10, VB96}:

$$G(\Gamma) = \log \inf \left\{ \prod_{i=1}^n \lambda_i^{1/2} | \Gamma \geq S^T S \right\}$$

$$\geq \inf \left\{ \sum_{i=1}^n \log \lambda_i^{1/2} (\exp(\gamma_0))^{1/2} | \log \Gamma \geq \gamma_0, \gamma_0 \in \pi(n) \right\}$$

$$= \inf \left\{ \frac{1}{2} \sum_{i=1}^n \lambda_i^{1/2} (\gamma_0) | \log \Gamma \geq \gamma_0, \gamma_0 \in \pi(n) \right\}$$

$$= \inf \left\{ \frac{1}{4} \sum_{i=1}^n \lambda_i^{1/2} (\gamma_0) | \log \Gamma \geq \gamma_0, \gamma_0 \in \pi(n) \right\}.$$

Numerics for small dimensions suggest that this bound is mostly smaller than the spectral lower bounds.

5. An operational definition of the squeezing measure

We claim that $G$ answers the question: given a state, what is the minimal amount of single-mode squeezers needed to prepare it? In other words, it quantifies the amount of squeezing needed for the preparation of a state.

5.1. Operations for state preparation and an operational measure for squeezing

We first specify the preparation procedure. Since we want to quantify squeezing, it seems natural that we allow to freely draw states from the vacuum or a thermal bath to start with. Furthermore, we can perform an arbitrary number of the following operations for free:

(1) Add ancillary states also from a thermal bath or the vacuum.
(2) Add Gaussian noise.
(3) Implement any gate from linear optics.
(4) Perform Weyl-translations of the state.
(5) Perform selective or non-selective Gaussian measurements such as homodyne or heterodyne detection.
(6) Forget part of the state.
(7) Create convex combinations of ensembles.

In addition, the following operation comes with an associated cost:

(8) Implement single-mode squeezers at a cost of $\log(s)$, where $s$ is the squeezing parameter.
All these operations are standard operations in quantum optics and they should capture all important Gaussian operations except for squeezing.

It is well-known that all of these operations are captured by the following set of operations on the covariance matrix (for a justification, see appendix D):

(O0) We can always draw $N$-mode states with $\gamma \in \mathbb{R}^{2N \times 2N}$ for any dimension $N$ from the vacuum $\gamma = 1$ or a bath fulfilling $\gamma \geq 1$.

(O1) We can always add ancillary modes from the vacuum $\gamma_{\text{anc}} = 1$ or a bath $\gamma \geq 1$ and consider $\gamma \oplus \gamma_{\text{anc}}$.

(O2) We can freely add noise with $\gamma_{\text{noise}} \geq 0$ to our state, which is simply added to the covariance matrix of a state.

(O3) We can perform any beam splitter or phase shifter and in general any operation $S \in K(n)$, which translates to a map $\gamma \mapsto S^T \gamma S$ on covariance matrices of states.

(O4) We can perform any single-mode squeezer $S_{\text{ss}} = \frac{1}{1 + \gamma}$ for some $\gamma \in \mathbb{R}^+$. $S_{\text{ss}}$ is defined as

\[
S_{\text{ss}} = \text{diag}(1, 1, \ldots, 1, \frac{1}{\sqrt{1 + \gamma}}, \frac{1}{\sqrt{1 + \gamma}}, \ldots, \frac{1}{\sqrt{1 + \gamma}}).
\]

(O5) We can perform any Weyl-translation leaving the covariance matrix invariant.

(O6) Given two states with covariance matrices $\gamma_1$ and $\gamma_2$, we can always take their convex combination $\gamma = p \gamma_1 + (1 - p) \gamma_2$ for any $p \in [0, 1]$.

(O7) At any point, we can perform a selective measurement of the system corresponding to a projection into a finitely or infinitely squeezed state. Given a state with covariance matrix $\gamma = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$, this maps $\gamma \mapsto A - C (B - \gamma_0) MP C^T$, where $MP$ denotes the Moore–Penrose pseudoinverse.

Only operation (O4) comes at a cost of $s \log(s)$, all other operations are free.

We are now ready to state our main theorem, which states that the minimal squeezing cost for any possible preparation procedure consisting of operations (1)–(8) is given by $G$.

\begin{theorem}
Let $\rho$ be a quantum state with covariance matrix $\gamma$. Consider arbitrary sequences

\[
\tilde{\gamma}_N : \gamma_0 \rightarrow \gamma_1 \rightarrow \cdots \rightarrow \gamma_N,
\]

where $\gamma_0$ fulfills (O0) and every arrow corresponds to an arbitrary operation (O1)–(O5) or (O7). Using (O6), we can merge two sequences $\tilde{\gamma}_{N_1}$ and $\tilde{\gamma}_{N_2}$ to one resulting tree with $\gamma_{N_1 + N_2 + 1} = \lambda \gamma_{N_1} + (1 - \lambda) \gamma_{N_2}$ for some $\lambda \in (0, 1)$. Iteratively, we can construct trees of any depth and width using operations (O1)–(O7).

Let $\mathcal{D}_N(\gamma)$ be the set of such trees with $N$ operations ending with $\gamma$ (i.e. $\gamma_N = \gamma$). Let $\mathcal{D}(\gamma) = \bigcup_{N=1}^{\infty} \mathcal{D}_N(\gamma)$.

Furthermore, for any tree $\hat{\gamma} \in \mathcal{D}_N(\gamma)$, let $\hat{s} = \{s_i\}_{i=1}^{M}$ be the sequence of the largest singular values of any single-mode squeezer (O4) implemented along the sequence (in particular, $M \leq N$). Then

\[
G(\gamma) = \inf \left\{ \sum_i \log s_i | s_i \in \hat{s}, \hat{\gamma} \in \mathcal{D}(\gamma) \right\}.
\]

(39)
\end{theorem}
5.2. Proof of the main theorem

Since we consider many different operations, the proof is rather lengthy, where the main difficulties will be in showing that measurements do not squeeze. In order to increase readability, the proof will be split into several lemmata.

**Lemma 5.2.** Let \( \gamma \in \mathbb{R}^{2n \times 2n} \) be a covariance matrix, \( \gamma_0 \geq 1 \), let \( N \in \mathbb{N} \) and

\[
\gamma_0 \rightarrow \gamma_1 \rightarrow \cdots \rightarrow \gamma_N = \gamma
\]

be any sequence of actions (O1)-(O5) or (O7). If we denote the cost (sum of the logarithm of the largest singular values of any symplectic matrix involved) of this sequence by \( c \), then one can replace this sequence by:

\[
\gamma_0 \rightarrow \gamma_0 \oplus \gamma_{\text{anc}} \rightarrow \gamma_0 \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}} \rightarrow \gamma_0 \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}} S \rightarrow \mathcal{M}(S^T(\gamma_0 \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}})S)
\]

with \( \gamma_{\text{anc}} \geq 1, \gamma_{\text{noise}} \geq 0, S \in \text{Sp}(2n) \) and \( \mathcal{M} \) a partial Gaussian measurement of type specified in (O7). For this sequence, \( c \geq F(S) \).

**Proof.** We prove the proposition by proving that given any chain \( \gamma_0 \rightarrow \gamma_1 \rightarrow \cdots \rightarrow \gamma_N = \gamma \) as in (40), we can interchange all operations and obtain a chain as in equation (41). For readability, we will not always specify the size of the matrices and we will assume that \( \gamma \geq i\sigma, \gamma_{\text{anc}} \geq 1, \gamma_{\text{noise}} \geq 0, \) and \( S \) a symplectic matrix, whenever the symbols arise:

1. We can combine any sequence \( \gamma_i \rightarrow \gamma_{i+1} \rightarrow \cdots \rightarrow \gamma_{i+m} \) for some \( m \in \mathbb{N} \) where each of the arrows corresponds to a symplectic transformation \( S_j, j = 1, \ldots, m \) as in (O3) or (O4), into a single symplectic matrix \( S \in \text{Sp}(2n) \) such that \( \gamma_{i+m} = S^T\gamma_i S \). Furthermore lemma 3.4 implies \( F(S) \leq \sum_i s_i^S(\gamma_i) \), hence this recombination of steps only lowers the amount of squeezing.

2. Any sequence \( \gamma \rightarrow S^T\gamma S \rightarrow S^T\gamma S + \gamma_{\text{noise}} \) can be converted into a sequence \( \gamma \rightarrow S^T(\gamma + \gamma_{\text{noise}})S \) with the same \( S \) and hence the same costs by setting \( \gamma_{\text{noise}} := S^{-T}s_{\text{noise}}S^{-1} \geq 0 \).

3. Any sequence \( \gamma \rightarrow S^T\gamma S \rightarrow S^T\gamma S \oplus \gamma_{\text{anc}} \) can be converted into a sequence \( \gamma \rightarrow S^T(\gamma \oplus \gamma_{\text{anc}})S \) by setting \( \tilde{S} = S \oplus 1 \) with 1 of the same dimension as \( \gamma_{\text{anc}} \). Since we only add the identity, we have \( F(\tilde{S}) = \sum_i \log s_i^{\tilde{S}} = F(S) \) and the costs do not increase.

4. Any sequence \( \gamma \rightarrow \gamma + \gamma_{\text{noise}} \rightarrow (\gamma + \gamma_{\text{noise}}) \oplus \gamma_{\text{anc}} \) can be converted into a sequence \( \gamma \rightarrow \gamma \oplus \gamma_{\text{anc}} \rightarrow \gamma \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}} \) by setting \( s_{\text{noise}} := s_{\text{noise}} \oplus 0 \geq 0 \), which is again a valid noise matrix. As no operation of type (O4) is involved, the squeezing costs do not change.

In a next step we consider measurements. We will only consider homodyne detection, since the proof is exactly the same for arbitrary Gaussian measurements of type (O7). Given a covariance matrix \( \gamma \), we assume a decomposition

\[
\gamma = \begin{pmatrix}
A & C \\
C^T & B
\end{pmatrix}; \quad \mathcal{M}(\gamma) = A - C(\pi B \pi)^{MC^T}
\]

as in the definition of (O7) with \( \pi = \text{diag}(1, 0) \).
Any sequence $\gamma \to \mathcal{M}(\gamma) \to S^{T} \mathcal{M}(\gamma) S$ can be converted into a sequence $\gamma \to S^{T} \gamma S \to \mathcal{M}(S^{T} \gamma S)$ by setting $\hat{S} = S \oplus I_{2}$. To see this, write $S^{T} \mathcal{M}(\gamma) S = S^{T} A S - S^{T} C(\pi B \pi)^{M} C^{T} S$ and

$$
\mathcal{M} \left( \begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix} \right) \mathcal{M} \left( \begin{pmatrix} A & C \\ C^{T} & B \end{pmatrix} \right) \mathcal{M} \left( \begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix} \right) = \mathcal{M} \left( \begin{pmatrix} S^{T} A S & S^{T} C \\ C^{T} S & B \end{pmatrix} \right) = S^{T} A S - S^{T} C(\pi B \pi)^{M} C^{T} S
$$

hence the final covariance matrices are the same. By the same reasoning as in (3), the costs are equivalent.

Any sequence $\gamma \to \mathcal{M}(\gamma) \to \mathcal{M}(\gamma) + \gamma_{\text{noise}}$ can be converted into a sequence $\gamma \to \gamma + \gamma_{\text{noise}} \to \mathcal{M}(\gamma + \gamma_{\text{noise}})$ by setting $\gamma_{\text{noise}} = \gamma_{\text{noise}} \oplus 0$, with 0 on the last mode being measured. Since no symplectic matrices are involved, the costs are equivalent.

Any sequence $\gamma \to \mathcal{M}(\gamma) \to \mathcal{M}(\gamma) \oplus \gamma_{\text{anc}}$ can be changed into a sequence $\gamma \to \gamma \oplus \gamma_{\text{anc}} \to \hat{\mathcal{M}}(\gamma \oplus \gamma_{\text{anc}})$, where the measurement $\hat{\mathcal{M}}$ measures the last mode of $\gamma$, i.e.

$$
\hat{\mathcal{M}} \left( \begin{pmatrix} A & C & 0 \\ C^{T} & B & 0 \\ 0 & 0 & \gamma_{\text{anc}} \end{pmatrix} \right) = (A \oplus \gamma_{\text{anc}}) - (C \oplus 0)(\pi B \pi)^{M}(C \oplus 0)^{T}.
$$

Clearly, the resulting covariance matrices of the two sequences are the same and the costs are equivalent.

We can now easily prove the lemma. Let $\gamma_{0} \to \cdots \to \gamma_{m}$ be an arbitrary sequence with operations of type (O1)–(O5) or (O7). We can first move all measurements to the right of the sequence, i.e. we first perform all operations of type (O1)–(O5) and then all measurements. This is done using the observations above. Note also that this step is similar to the quantum circuit idea to ‘perform all measurements last’ (see [NC00], chapter 4).

Similarly, we can combine operations of type (O3) and (O4) and rearrange the other operations to obtain a new sequence as in equation (41) with at most the costs of the sequence $\gamma_{1} \to \cdots \to \gamma_{m}$ we started with.

We can now slowly work towards theorem 5.1:

Lemma 5.3. Let $\gamma \in \mathbb{R}^{2n \times 2n}$ be a covariance matrix, then

$$
G(\gamma) = \inf \{ F(S) | \gamma = S^{T}(\gamma_{0} \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}}) S, \ S \in Sp(2n), \ \gamma_{0} \oplus \gamma_{\text{anc}} \geq 1, \ \gamma_{\text{noise}} \geq 0 \}.
$$

Proof. First note that for any $\gamma \geq i \sigma$, we can find $S \in Sp(2n), \ \gamma_{0} \in \mathbb{R}^{2n \times 2n}$ with $\gamma_{0} \geq 1$ and $\gamma_{\text{noise}} \in \mathbb{R}^{2n \times 2n}$ with $\gamma_{\text{noise}} \geq 0$ such that $\gamma = S^{T}(\gamma_{0} + \gamma_{\text{noise}}) S$ by using Williamson’s theorem, hence the feasible set is never empty. The lemma is immediate by observing that for any $\gamma = S^{T}(\gamma_{0} + \gamma_{\text{anc}} + \gamma_{\text{noise}}) S$ since $(\gamma_{0} \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}}) \geq 1$ we have $\gamma \geq S^{T} S$ and conversely, for any $\gamma \geq S^{T} S$, defining $\gamma_{0} := S^{T} \gamma S^{-1} \geq 1$, we have $\gamma = S^{T} \gamma_{0} S$. 

We can now slowly work towards theorem 5.1:
As an intermediate step we introduce the following notation:

\[
G(\gamma) := \inf \{ F(S) \mid \gamma = M(S^{T}(\gamma_{0} \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}})S), \ S \in \text{Sp}(2n), \\
\gamma_{0} \oplus \gamma_{\text{anc}} \geq I_{2n}, \ \gamma_{\text{noise}} \geq 0, \ \mathcal{M} \text{ measurement} \}. 
\]

Then we have:

**Lemma 5.4.** For \( \gamma \in \mathbb{R}^{2n \times 2n} \) a covariance matrix, we have

\[
G(\gamma) = \inf \{ F(\hat{\gamma}^{1/2}) \mid \gamma = M(\hat{\gamma}), \ \hat{\gamma} \geq \hat{\gamma} \geq i\sigma, \ \mathcal{M} \text{ measurement} \}. 
\]

**Proof.** This follows from lemma 5.3:

\[
G(\gamma) = \inf \{ F(S) \mid \gamma = M(S^{T}(\gamma_{0} \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}})S) \}
\]

\[
= \inf \{ F(S) \mid \gamma = M(\hat{\gamma}), \ \hat{\gamma} = S^{T}(\gamma_{0} \oplus \gamma_{\text{anc}} + \gamma_{\text{noise}})S \geq i\sigma \}
\]

Lemma 5.3 \inf \{ G(\hat{\gamma}) \mid \gamma = M(\hat{\gamma}), \ \hat{\gamma} \geq \hat{\gamma} \geq i\sigma \}

Prop. 4.2 \inf \{ F(\hat{\gamma}^{1/2}) \mid \gamma = M(\hat{\gamma}), \ \hat{\gamma} \geq \hat{\gamma} \geq i\sigma \}

by taking the infimum over all measurements last.

Note here, that equation (45) together with the following proposition 5.5 finishes the proof of proposition 4.6 via:

\[
G(\gamma) = \inf \{ G(\hat{\gamma}) \mid \gamma = M(\hat{\gamma}), \ \hat{\gamma} \geq \hat{\gamma} \geq i\sigma \} \leq G(\gamma \oplus aI_{n})
\]

for \( a \geq 1 \), using that measuring the last modes we obtain \( M(\gamma \oplus aI_{n}) = \gamma \) and therefore, \( \gamma \oplus aI_{n} \) is in the feasible set of \( G(\gamma) = G(\gamma) \).

**Proposition 5.5.** For \( \gamma \in \mathbb{R}^{2n \times 2n} \) a covariance matrix we have

\[
G(\gamma) = G(\gamma).
\]

This proposition shows that \( G \) is operational if we exclude convex combinations (and therefore also non-selective measurements).

**Proof.** Using lemma 5.4, the proof of this proposition reduces to the question whether:

\[
\inf \{ F(\hat{\gamma}^{1/2}) \mid \gamma \geq \hat{\gamma} \geq i\sigma, \ M(\hat{\gamma}) = \gamma \} = \inf \{ F(\hat{\gamma}^{1/2}) \mid \gamma \geq \hat{\gamma} \geq i\sigma \}.
\]

Since we do not need to use measurements, \( \leq \) is obvious.

Let \( \gamma \geq \hat{\gamma} \geq i\sigma \) for some \( M(\hat{\gamma}) = \gamma \). Our first claim is that

\[
\gamma \geq M(\hat{\gamma}) \geq i\sigma
\]

\( M(\hat{\gamma}) \geq i\sigma \) is clear from the fact that \( \hat{\gamma} \) is a covariance matrix and a measurement takes states to states. \( \gamma \geq M(\hat{\gamma}) \) is proved using Schur complements. Let \( M \) be a Gaussian measurement as in equation (68) with \( \gamma_{G} = \text{diag}(d, 1/d) \) with \( d \in \mathbb{R}^{+} \). It is well-known that

\[
M(\gamma) = (1 \oplus \text{diag}(1/d, d) \gamma (1 \oplus \text{diag}(1/d, d)) + 0 \oplus I_{2})^{\frac{1}{2}},
\]
where $\hat{g}$ denotes the Schur complement of the block in the lower-right corner of the matrix. For homodyne measurements, we take the limit $d \to \infty$. Since for any $\gamma \geq \hat{g} \geq 0$, the Schur complements of the lower right block fulfill $\hat{g}^S \geq \hat{g}^S \geq 0$ (see [Bha07], exercise 1.5.7), we have $\gamma \geq \mathcal{M}(\hat{g})$ as claimed in equation (48).

Next, we claim

$$F(\mathcal{M}(\gamma)^{1/2}) \leq F(\gamma^{1/2}).$$

To prove this claim, note that via the monotonicity of the exponential function on $\mathbb{R}$, it suffices to prove

$$\prod_{j=1}^m s_j(\mathcal{M}(\hat{g})) \leq \prod_{j=1}^n s_j(\hat{g})$$

when we assume $\hat{g} \in \mathbb{R}^{2n \times 2n}$ and $\mathcal{M}(\hat{g}) \in \mathbb{R}^{2m \times 2m}$ with $m \leq n$. If we write

$$\hat{g} = \begin{pmatrix} \hat{A} & \hat{C} \\ \hat{C}^T & \hat{B} \end{pmatrix}$$

then the state after measurement is given by $\mathcal{M}(\hat{g}) = \hat{A} - \hat{C}(\hat{B} + \text{diag}(d, 1/d)^{-1}\hat{C}^T \hat{B})$ or the limit $d \to \infty$ for homodyne measurements. In any case $\hat{C}(\hat{B} + \text{diag}(d, 1/d)^{-1}\hat{C}^T \hat{B}) \geq 0$ and $\mathcal{M}(\hat{g}) \leq \hat{A}$ and therefore, by Weyl’s inequalities, also

$$\prod_{j=1}^m s_j(\mathcal{M}(\hat{g})) \leq \prod_{j=1}^m s_j(\hat{A}).$$

Now we use Cauchy’s interlacing theorem (see [Bha96], corollary III.1.5): as $\hat{A}$ is a submatrix of $\hat{g}$, we have $\lambda^i(\hat{A}) \leq \lambda^j(\hat{g})$ for all $i = 1, \ldots, 2m$. Since at least $m$ eigenvalues of $\hat{A}$ are bigger or equal one and at least $n$ eigenvalues of $\hat{g}$ are bigger or equal one, this implies

$$\prod_{j=1}^m s_j(\hat{A}) = \prod_{j=1}^m \lambda^j(\hat{A}) \leq \prod_{j=1}^m \lambda^j(\hat{g}) \leq \prod_{j=1}^n \lambda^j(\hat{g}) = \prod_{j=1}^n s_j(\hat{g}).$$

In particular, this proves equation (49).

We can then complete the proof: let $\hat{g} \geq \hat{g} \geq i\sigma$ for some $\mathcal{M}(\hat{g}) = \gamma$ in equation (47). We have just seen that this implies $\gamma \geq \mathcal{M}(\hat{g}) \geq i\sigma$ via equation (48) and furthermore that $F(\gamma^{1/2}) \geq F(\mathcal{M}(\hat{g})^{1/2})$ via equation (49). But this means that we have found $\gamma = \mathcal{M}(\hat{g})$ such that $\gamma \geq \gamma \geq i\sigma$. Hence $\gamma$ is in the feasible set of the right-hand side of (47) and $F(\gamma^{1/2}) \geq F(\gamma^{1/2})$, which implies $\geq$ in equation (47).

Finally, we can prove theorem 5.1 by also covering convex combinations:

**Proof.** Let $\gamma \in \mathbb{R}^{2n \times 2n}$ be a covariance matrix. First consider only sequences $\hat{g}_i$: we replace any sequence by the special type of sequence of lemma 5.3. For these sequences, we have seen that the minimum cost is given by $G(\gamma)$ in proposition 5.5.

However, we explicitly excluded convex combinations (O6) by considering only sequences and not trees $\hat{g}_i$: consider a tree of operations (O1)–(O7) which has $\gamma$ at its root and $\gamma_0 = I$ as leaves. Let us consider any node closest to the leaves. At such a node, we start with two covariance matrices $\gamma_1$ and $\gamma_2$ that were previously constructed without using convex combinations and with costs $G(\gamma_1)$ and $G(\gamma_2)$. The combined matrix would be $\gamma = \lambda \gamma_1 + (1 - \lambda) \gamma_2$ for some $\lambda \in (0, 1)$ and the costs would be $\lambda G(\gamma_1) + (1 - \lambda) G(\gamma_2)$.
By convexity of $G$ (see theorem 4.3):

$$G(\lambda \gamma_1 + (1 - \lambda) \gamma_2) \leq \lambda G(\gamma_1) + (1 - \lambda) G(\gamma_2)$$

which means that we can find a sequence (without any convex combinations) producing $\lambda \gamma_1 + (1 - \lambda) \gamma_2$ which is cheaper than first producing $\gamma_1$ and $\gamma_2$ and then taking a convex combination. Iteratively, this means we can eliminate every node from the tree and replace the tree by a sequence of operations (O1)–(O5) and (O7), which is cheaper than the tree and trees do not matter.

5.3. The squeezing measure as a resource measure

We have now seen that the measure $G$ can be interpreted as a measure of the amount of single-mode squeezing needed to create a state $\rho$. Let us now take a different perspective, which is the analogue of the entanglement of formation for squeezing: consider covariance matrices of the form

$$g_s := \begin{pmatrix} s & 0 \\ 0 & s^{-1} \end{pmatrix}.$$ (51)

These are single-mode squeezed states with squeezing parameter $s \geq 1$. We will now allow these states as resources and ask the question: given a (Gaussian) state $\rho$ with covariance matrix $\gamma$, what is the minimal amount of these resources needed to construct $\gamma$, if we can freely transform the state by the same operations as before excluding squeezing (O1)–(O7) excluding (O4)).

The corresponding measure is once again $G$:

**Theorem 5.6.** Let $\rho$ be an $n$-mode state with covariance matrix $\gamma \in \mathbb{R}^{2n \times 2n}$. Then

$$G(\gamma) = \inf \left\{ \sum_{i=1}^{m} \frac{1}{2} \log(s m_{i}) = T(\bigoplus_{i=1}^{m} \gamma_{i}) \right\},$$ (52)

where $T: \mathbb{R}^{2m \times 2m} \rightarrow \mathbb{R}^{2n \times 2n}$ is a combination of the operations (1)–(6) above.

**Proof.** $\leq$: Note that for any feasible $S \in Sp(2n)$ in $G(\gamma)$, i.e. any $S$ with $S^T S \leq \gamma$, we can find $O \in Sp(2n) \cap O(2n)$ and $D = \bigoplus_{i=1}^{m} \gamma_{i}$ with $S^T S = O^T DO$ via the Euler decomposition. Using that the Euler decomposition minimises $F$, we have $F(S) = \frac{1}{2} F(D) = \sum_{i=1}^{m} \frac{1}{2} \log(s_{i})$. But then, since we can find $\gamma_{\text{noise}} \geq 0$ such that $\gamma = O^T \bigoplus_{i=1}^{m} \gamma_{i} + \gamma_{\text{noise}}$, we have that $D$ is a feasible resource state to produce $\gamma$. This implies $G_{\text{resource}}(\gamma) \leq G(\gamma)$.

$\geq$: For the other direction, the proof proceeds exactly as the proof of theorem 5.1. First, we exclude convex combinations. Then, we realise that we can change the order of the different operations (even if we include adding resource states during any stage of the preparation process) according to lemma 5.2, making sure that any preparation procedure can be implemented via:

$$\gamma = \mathcal{M} \left( O \bigoplus_{i=1}^{m} \gamma_{i} \oplus I_{2m'} + \gamma_{\text{noise}} \right) O^T,$$

where $O \in Sp(2m + 2m') \cap O(2m + 2m')$, $\gamma_{\text{noise}} \in \mathbb{R}^{2m + 2m' \times 2m + 2m'}$ with $\gamma_{\text{noise}} \geq 0$ and $\mathcal{M}$ a measurement. Now the only difference to proof of 5.1 is that we had the vacuum $I$ instead of $\bigoplus_{i=1}^{m} \gamma_{i} \oplus I_{2m'}$ and an arbitrary symplectic matrix $S$ instead of $O$, but the two ways of writing the maps are completely interchangeable, so that the proof proceeds as in theorem 5.1. □
We could call this measure the ‘(Gaussian) squeezing of formation’, as it is the analogue to the Gaussian entanglement of formation. Note also that the measure is similar to the Gaussian entanglement of formation as defined in [Wol+04]. One natural further question would be whether ‘distillation of squeezing’ is possible with Gaussian operations. It is impossible in some sense for the minimal eigenvalue via [Kra+03], while it is possible and has been investigated for non-Gaussian states in many papers (see [Fi13, Hee+06] and references therein). In our case, it is not immediately clear whether extraction of single-mode squeezed states with less squeezing is possible or not. This could be investigated in future work.

6. Calculating the squeezing measure

We have seen that the measure \( G \) is operational. However, to be useful, we need a way to compute it.

6.1. Analytical solutions

\[ \text{Proposition 6.1.} \quad \text{Let } n = 1, \text{ then } G(\Gamma) = -\frac{1}{2} \min_{\lambda} \log(\lambda(\Gamma)) \text{ for all } \Gamma \in \mathbb{R}^{2n \times 2n}. \]

\[ \text{Proof.} \quad \text{Note that this is the lower bound in proposition 4.9, hence } \frac{1}{2} \min_{\lambda} \log(\lambda(\Gamma)) \leq G(\Gamma). \text{ Now consider the diagonalisation } \Gamma = O \text{ diag}(\lambda_1, \lambda_2)O^T \text{ with } O \in SO(2) \text{ and assume } \lambda_1 \geq \lambda_2. \text{ Then, } \lambda_1^{-1} \leq \lambda_1 \text{ since otherwise, } \Gamma \neq iJ. \]

Consider \( \text{diag}(\lambda_1, \lambda_2) \geq O^{-T}S^TOS^{-1} \) for some \( S \in Sp(2) \) with eigenvalues \( s \geq 1 \) and \( s^{-1} \). Since \( \text{diag}(\lambda_1, \lambda_2) \geq O^{-T}S^TOS^{-1} \), this implies in particular that \( s^{-1} \leq \lambda_2 \) by Weyl’s inequality. Since \( F(S^T S) = \log s \), in order to minimise \( F(S) \) over \( S^T S \leq \Gamma \), we need to maximize \( s^{-1} \). Setting \( s^{-1} = \lambda_2 \) we obtain \( s = \lambda_2^{-1} \leq \lambda_1 \) and \( \text{diag}(\lambda_1, \lambda_2) \geq \text{diag}(s, s^{-1}). \) Since \( SO(2) = K(1), S^T S = O^T \text{ diag}(\lambda_1, \lambda_2)O \leq \Gamma \) is the minimising matrix in \( G \) and \( G(\Gamma) = F(S) = \frac{1}{2} \log \lambda_2^{-1}. \)

\[ \text{Proposition 6.2.} \quad \text{Let } \rho \text{ be a pure, Gaussian state with covariance matrix } \Gamma \in \mathbb{R}^{2n \times 2n}. \text{ Then } G(\Gamma) = F(\Gamma^{1/2}). \]

\[ \text{Proof.} \quad \text{From proposition 2.2, we know that } \det(\Gamma) = 1 \text{ in particular. Therefore, the bounds in proposition 4.11 are tight and } G(\Gamma) = F(\Gamma^{1/2}). \]

6.2. Numerical calculations using Matlab

The crucial observation to numerically find the optimal squeezing measure is given in lemma 4.4: if we use \( G \) in the form of equation (25), we know that the function to be minimised is convex on \( \mathcal{H} \). In general, convex optimisation with convex constraints is efficiently implementable and there is a huge literature on the topic (see [BV04] for an overview).

In our case, a certain number of problems occur when performing convex optimisation:

(1) The function \( f \) in equation (28) is highly nonlinear. It is also not differentiable at eigenvalue crossings of \( A + iB \) or \( H \in \mathcal{H} \). In particular, it is not differentiable when one of the eigenvalues becomes zero, which is to be expected at the minimum.
(2) While the constraints $C^{-1}(\gamma) \geq H$ and $1 > H > -1$ are linear in matrices, they are nonlinear in simple parameterisations of matrices.

(3) For $\gamma$ on the boundary of the set of allowed density operators, the set of feasible solutions might not have an inner point.

The first and second problem imply that most optimisation methods are unsuitable, as they are either gradient-based or need more problem structure. It also means that there is no guarantee for good stability of the solutions. The third problem implies that interior point methods become unsuitable on the boundary, which limits applications. For instance, our example of the next section (see equation (53)) lies on the boundary. As a proof of principle implementation, we used the MATLAB-based solver SOLVOPT (for details see the manual [KK97]). We believe our implementation could be made more efficient and more stable, but it seems to work well in most cases for less than ten modes. More information on the programme is provided in appendix E.

### 6.3. Squeezing-optimal preparation for certain three-mode separable states

Let us now work with a particular example that has been studied in the quantum information literature. In [MK08], Mišta Jr and Korolkova define the following three-parameter group of three-mode states where the modes are labelled $A$, $B$, $C$:

$$\gamma = \gamma_{AB} \otimes \mathbb{1}_C + x(q_1q_1^T + q_2q_2^T)$$  \hspace{1cm} (53)$$

with

$$\gamma_{AB} = \begin{pmatrix} e^{2d_a} & 0 & -e^{2d_c} & 0 \\ 0 & e^{-2d_a} & 0 & e^{-2d_c} \\ -e^{2d_c} & 0 & e^{2d_a} & 0 \\ 0 & e^{-2d_c} & 0 & e^{-2d_a} \end{pmatrix},$$

$$q_1 = (0, \sin \phi, 0, -\sin \phi, \sqrt{2}, \sqrt{2})^T,$$

$$q_2 = (\cos \phi, 0, \cos \phi, 0, \sqrt{2}, \sqrt{2})^T,$$

where $a = \cosh(2r)$, $c = \sinh(2r)$, $\tan \phi = e^{-2r} \sinh(2d) + \sqrt{1 + e^{-4r} \sinh^2(2d)}$. The remaining parameters are $d \geq r > 0$ and $x \geq 0$. For

$$x = r \geq \frac{2 \sinh(2r)}{e^{2d} \sin^2 \phi + e^{-2d} \cos^2 \phi},$$

the state becomes fully separable [MK08]. The state as such is a special case of a bigger family described in [Gie+01]. In [MK08], it was used to entangle two systems at distant locations using fully separable mediating ancillas (here the system labelled $C$). Therefore, Mišta Jr and Korolkova considered also an LOCC procedure to prepare the state characterised by (53). For our purposes, this is less relevant and we allow for arbitrary preparations of the state. This was also done in [MK08] by first preparing modes $A$ and $B$ each in a pure squeezed-state with position quadratures $e^{2d(r-r)}$ and $e^{2d(r+r)}$. A vacuum mode in $C$ was added and $x(q_1q_1^T + q_2q_2^T)$ was added as random noise. Therefore, the squeezing needed to produce this state in this protocol is given by

$$c = \frac{1}{2} \log(e^{2d(r-r)} \cdot e^{2d(r+r)}) = 2d.$$  \hspace{1cm} (54)$$

We numerically approximated the squeezing measure for $\gamma_{ABC}$, choosing $x = x_{sep}$, which leaves a two-parameter family of states. We chose parameters $d$ and $r$ according to
with \( \mu_{ij} \), \( 1, 2, 3 \) \{ \} for a total of 900 data points. Since the algorithm is not an interior point algorithm as described above, to check the result, we reprepared the state in the following way:

1. Let \( S \) be the symplectic matrix at the value optimum found by SOLVOPT for a covariance matrix \( \gamma_{ABC} \).
2. Calculate \( S^{-1} \gamma_{ABC} S^{-1} \) and calculate its lowest eigenvalue \( \lambda_{2n} \).

\[
\begin{align*}
\text{Figure 1. Results of numerical calculations (formulas for } d \text{ and } r \text{ in equation (55)). On the upper figure, the lower range of points (green in the online version) are the best lower bound, the middle points (blue) denote the value of the objective function at the minimum found by SOLVOPT and the upper points (red) denote the squeezing costs of the preparation protocol of [MK08] (equation (54)). The lower figure shows the preparation error. It is mostly below } 10^{-6}. \\

r = 0.1 + j \cdot 0.05, \quad d = r + i \cdot 0.03 \\
\end{align*}
\]

with \( i, j \in \{1, \ldots, 30\} \) for a total of 900 data points. Since the algorithm is not an interior point algorithm as described above, to check the result, we reprepared the state in the following way:

1. Let \( S \) be the symplectic matrix at the value optimum found by SOLVOPT for a covariance matrix \( \gamma_{ABC} \).
2. Calculate \( S^{-1} \gamma_{ABC} S^{-1} \) and calculate its lowest eigenvalue \( \lambda_{2n} \).
(3) Define \( \tilde{\gamma} := S^{-T} \gamma_{ABC} S^{-1} + (1 - \min\{1, \lambda_{2n}\}) I \succeq I \). Calculate the largest singular value of \( S^T \tilde{\gamma} S - \gamma \).

If \( S \) was a feasible point, then \( S^T \tilde{\gamma} S = \gamma \). Since it is obvious how to prepare \( \tilde{\gamma} \) with operations specified in section 5, the largest singular value of \( S^T \tilde{\gamma} S - \gamma \) is an indicator of how well we can approximate the state we want to prepare by a state with comparably low squeezing costs.

The results of the numerical computation are shown in figure 1. We computed the minimum both with the help of numerical and analytical subgradients and took the value with a better approximation error. At rare occasions, one algorithm failed to obtain a minimum. Possible reasons for this are discussed in appendix E. The optimal values computed by the algorithm are close to the lower bound and a lot better than the upper bound and the costs obtained by equation (54). One can easily see that \( \gamma_{ABC} \) cannot achieve the spectral lower bound as the assumptions of lemma 4.10 are not met.

7. Discussion of modifications to allowed operations

In experiments, squeezing of a state is most commonly measured by the logarithm of the smallest eigenvalue (up to a constant) and the unit is usually referred to as decibel (dB) [Lvo15]. We know of no operational interpretation for this measure that is similar to the interpretation given in section 5 and the measure is not natural for multimode states.

In contrast, \( G \) is a natural measure for multimode states. However, squeezing is not just experimentally challenging, it gets much harder if we want to achieve a larger amount of single-mode squeezing. Currently, the highest amount of squeezing obtained in quantum optical systems seems to be about 13 dB (see [And+15]). In other words, the two states \( \rho \) and \( \rho' \) with covariance matrices

\[
\gamma = \text{diag}(s, s^{-1}, s, s^{-1}), \quad \gamma' = \text{diag}(s^2, s^{-2}, 1, 1)
\]

will not be equally hard to prepare although \( G(\gamma) = G(\gamma') \). This is due to the fact that we quantified the cost of a single-mode squeezer by \( \log s \).

To amend this, one could propose an easy modification to the definition of \( F \) in equation (11):

\[
F_g(\gamma) = \sum_{i=1}^{n} \log(g(s_i'(S)))
\]

by inserting another function \( g : \mathbb{R} \to \mathbb{R} \) to make sure that for the corresponding measure \( G_g(\rho) \equiv G_g(\gamma) \), we have \( G_g(\gamma) = G_g(\gamma') \) in equation (56). We pose the following natural restrictions on \( g \):

- We need \( g(1) = 1 \) since \( G_g(\rho) \) should be zero for unsqueezed states.
- Squeezing should get harder with larger parameter, hence \( g \) should be monotonously increasing.
- For simplicity, we assume \( g \) to be differentiable.

Let us first consider squeezing operations and the measure \( F_g \). We proved in proposition 3.3 and theorem 3.5 that \( F \) is minimised by the Euler decomposition. A crucial part was given by lemma 3.4. In order to be useful for applications, we must require the same to be true for \( F_g \), i.e.
This puts quite strong restraints on $g$: considering $n = 1$ and assuming that $S$ and $S'$ are diagonal with ordered singular values, this implies that $g$ must fulfill $g(xy) \leq g(x)g(y)$ for $x, y \geq 1$. This submultiplicativity restraint rules out all interesting classes of functions: Assume for instance that $g(2) = c$, then $g(2^n) \leq c^n$, where equality is attained if $g(x) = c \cdot x$. Therefore, all submultiplicative functions $g(x)$ for $x \geq 1$ must lie below $g(x) = c \cdot x$ at least periodically. Hence, lemma 3.4 does not hold if we consider increasingly growing functions $g$. This implies that one could make the measure arbitrarily small by splitting the single-mode squeezer into many successive single-mode squeezers with smaller squeezing parameter, which does not reflect experimental reality.

A way to circumvent the failure of lemma 3.4 would be to work with the ‘squeezing of formation’ measure. Likewise, one could require that there was only one operation of type (O4) as specified in section 5 in any preparation procedure. In that case we have:

**Proposition 7.1.** If $g : \mathbb{R} \to \mathbb{R}$ fulfills

\begin{enumerate}
\item $(1) \log \circ g \circ C$ is convex on $(1, \infty)$,
\item $(2) \log(g(\exp(t)))$ is convex and monotone increasing in $t$,
\end{enumerate}

then the squeezing of formation measure $G_g$ is still operational, i.e. theorem 5.1 still holds.

**Proof.** The first condition replaces the log-convexity of the Cayley transform in the proof of theorem 4.3, making the measure convex. Using [Bha96], II.3.5 (v), the second condition makes sure that equation (50) still holds. The second condition can probably be relaxed while the proof of theorem 5.1 is still applicable. A function $g$ fulfilling these prerequisites is $g(x) = \exp(x)$, which would correspond to a squeezing cost increasing linearly in the squeezing parameter. One could even introduce a cutoff after which $g$ would be infinite. □

A simpler way to reflect the problems of equation (56) would be to consider the measures $G$ and $G_{\text{minEig}}$ together (calculating $G_{\text{minEig}}$ of both the state and the minimal preparation procedure in $G$).

Another problem is associated with the form of the Hamiltonian (1). In the lab, the Hamiltonians that can be implemented might not be single-mode squeezers, but other operations such as symmetric two-mode squeezers (e.g. [SZ97], chapter 2.8). It is clear how to define a measure $G'$ for these kinds of squeezers. Using the Euler decomposition, $G$ is a lower bound to $G'$, but we did not investigate this any further.

**Acknowledgments**

MI thanks Konstantin Pieper for discussions about convex optimisation and Alexander Müller-Hermes for discussions about MATLAB. MI is supported by the Studienstiftung des deutschen Volkes.

**Appendix A. Preliminaries for the proof of theorem 3.5**

Let us collect facts about ordinary differential equations needed in the proof:
Proposition A.1. Consider the following system of differential equations for $x : [0, 1] \rightarrow \mathbb{R}^{2n}$:
\[
\begin{align*}
\dot{x}(t) & = x(t)^T A(t) \quad \forall t \in [0, 1], \quad \text{for some } x_0 \in \mathbb{R}^{2n}, s \in [0, 1], \\
x(s) & = x_s
\end{align*}
\]
where $A \in L^\infty([0, 1], \mathfrak{sp}(2n))$. Then this system has a unique solution, which is linear in $x_s$ and defined on all of $[0, 1]$ such that we can define a map
\[
\forall s, t \in [0, 1]: \quad (s, t) \mapsto U(s, t) \in B(\mathbb{R}^{2n})
\]
via $x(t)^T = x_s^T U(s, t)$ called the propagator of (58) that fulfills:

1. $U$ is continuous and differentiable almost everywhere.
2. $U(s, t)$ is absolutely continuous in $t$.
3. $U(s, t) = 1$ and $U(s, r)U(r, t) = U(s, t)$ for all $s, t \in [0, 1]$.
4. $U(s, t)^{-1} = U(t, s)$ for all $s, t \in [0, 1]$.
5. $U$ is the unique generalised (i.e. almost everywhere) solution to the initial value problem
\[
\partial_t U(s, t) - U(s, t)A(t) = 0 \quad U(s, s) = 1
\]
\[\text{on } C([0, 1]^2, \mathfrak{sp}^{2n}(2n)).\]
6. If $A(t) = A$ does not depend on $t$, then $S(r) = \exp(rA)$ solves equation (59) with $U(s, t) = S(t - s)$.
7. For all $s, t \in [0, 1]$:
\[
\|U(s, t)\|_\infty \leq \exp\left(\int_s^t \|A(\tau)\|_1 \, d\tau\right).
\]
8. $U(s, t) \in \mathfrak{sp}(2n)$ for all $t, s \in [0, 1]$ and $\gamma(t) = U(0, t)$ fulfills equation (9) with $\gamma(0) = 1$.

Proof. The proof of this (except for the part about $U(s, t) \in \mathfrak{sp}(2n)$) can be found in [Son98] (theorem 55 and lemma C.4.1) for the transposed differential equation $\dot{x}(t) = A(t)x(t)$.

For the last part, note that since $U(s, s) = 1 \in \mathfrak{sp}(2n)$, we have $U(s, s)^T J U(s, s) = J$.

We can now calculate almost everywhere:
\[
\partial_t (U(t, s)^T J U(t, s)) = -U(t, s)^T (A^T(t)J - JA(t)) U(t, s) = 0
\]
since $A(t) \in \mathfrak{sp}(2n)$ and therefore $A^T(t)J - JA(t) = 0$.

But this implies $U(t, s)^T J U(t, s) = J$, hence $U$ is symplectic. Obviously, $U(0, t)$ solves equation (9).

We will also need another well-known lemma from functional analysis:

Lemma A.2. Let $A : [0, 1] \rightarrow \mathfrak{sp}(2n)$, $A \in L^\infty([0, 1], \mathbb{R}^{2n \times 2n})$. Then $A$ can be approximated in $\|\cdot\|_1$-norm by step-functions, which we can assume to map to $\mathfrak{sp}(2n)$ without loss of generality.

The approximation by step-function can be found e.g. in [Rud87] (chapter 2, exercise 24).
Appendix B. The Cayley trick for matrices

In this appendix, we give an introduction to the Cayley-transform. The definition and properties needed in the main text are summarised by the following proposition:

Proposition B.1. Define the Cayley transform and its inverse via:
\[
\begin{align*}
\mathcal{C} : \{ H \in \mathbb{R}^{n \times n} | \text{spec}(H) \cap \{ +1 \} = \emptyset \} & \rightarrow \mathbb{R}^{n \times n} \\
H & \mapsto \frac{1 + H}{1 - H},
\end{align*}
\]
\[
\begin{align*}
\mathcal{C}^{-1} : \{ S \in \mathbb{R}^{n \times n} | \text{spec}(H) \cap \{ -1 \} = \emptyset \} & \rightarrow \mathbb{R}^{n \times n} \\
S & \mapsto \frac{S - 1}{S + 1}
\end{align*}
\]
\(\mathcal{C}\) is a diffeomorphism onto its image with inverse \(\mathcal{C}^{-1}\). Furthermore, it has the following properties:

1. \(\mathcal{C}\) is operator monotone and operator convex on matrices \(A\) with \(\text{spec}(A) \subset (-1, 1)\).
2. \(\mathcal{C}^{-1}\) is operator monotone and operator concave on matrices \(A\) with \(\text{spec}(A) \subset (-1, \infty)\).
3. \(\mathcal{C} : \mathbb{R} \rightarrow \mathbb{R}\) with \(\mathcal{C}(x) = (1 + x)/(1 - x)\) is log-convex on \([0, 1]\).
4. For \(n = 2m\) even, \(H \in \mathbb{R}^{2m \times 2m}\) and \(H \in \mathcal{H}\) if and only if \(\mathcal{C}(H) \in \text{Sp}(2m, \mathbb{R})\) and \(\mathcal{C}(H) \geq iJ\).

where \(\mathcal{H}\) is defined via:
\[
\mathcal{H} = \left\{ H = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \in \mathbb{R}^{2m \times 2m} \left| A^T = A, B^T = B, \text{spec}(H) \subset (-1, 1), \right. \right\}
\]

The definition and the fact that this maps the upper half plane of positive definite matrices to matrices inside the unit circle is present in [AG88] (I.4.2) and [MS98] (proposition 2.51, proof 2). Since no proof is given in the references and they do not cover the whole proposition, we provide them here.

We start with well-definedness:

Lemma B.2. \(\mathcal{C}\) and \(\mathcal{C}^{-1}\) are well-defined and inverses of each other. Moreover, \(\mathcal{C}\) is a diffeomorphism onto its image \(\text{dom}(\mathcal{C}^{-1})\).

Proof. If \(\text{spec}(H) \cap \{ +1 \} = \emptyset\), then \(1 - H\) is invertible and \(H \mapsto (1 + H)/(1 - H)\) is well-defined, as \(\det(1 + H, 1 - H) = 0\). Now let \(H \in \mathbb{R}^{m \times m}\) be such that \(\text{spec}(H) \cap \{ +1 \} = \emptyset\). We will show that \(\mathcal{C}(H)\) contains no eigenvalue \(-1\). To see this, let
\[
H = T \bigoplus_{i} J(n_i, \lambda_i) T^{-1}
\]
be the Jordan normal form with block sizes \(n_i\) and eigenvalues \(\lambda_i\). Let us here consider the complex Jordan decomposition, i.e. \(\lambda_i\) are allowed to be complex. Then:
\[
1 + H = T \bigoplus_{i} J(n_i, 1 + \lambda_i) T^{-1}, \quad 1 - H = T \bigoplus_{i} J(n_i, 1 - \lambda_i) T^{-1}
\]
and thus
\[ C(H) = T \bigoplus \sum_{i} J(n_i, 1 + \lambda_i) \cdot J(n_i, 1 - \lambda_i)^{-1} T^{-1}. \]

For the inverse of the Jordan blocks, we can use the well-known formula:
\[
\left( \begin{array}{cccc}
1 - \lambda_i & 1 & \cdots & 0 \\
0 & 1 - \lambda_i & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 - \lambda_i
\end{array} \right)^{-1} = \left( \begin{array}{cccc}
\frac{1}{1 - \lambda_i} & -1 & \cdots & \frac{(-1)^{n-1}}{(1 - \lambda_i)^{n-1}} \\
0 & \frac{1}{1 - \lambda_i} & \cdots & \frac{(-1)^{n-2}}{(1 - \lambda_i)^{n-2}} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{1 - \lambda_i}
\end{array} \right).
\]

In particular, this is still upper triangular. Then \( J(n_i, 1 + \lambda_i) J(n_i, 1 - \lambda_i)^{-1} \) is still upper triangular with diagonal entries \((1 + \lambda_i)/(1 - \lambda_i)\). Since \((1 + \lambda_i)/(1 - \lambda_i) \neq -1\) for all \(\lambda_i \in \mathbb{C}\), we find that \( J(n_i, 1 + \lambda_i) J(n_i, 1 - \lambda_i)^{-1} \) cannot have eigenvalue \(-1\) for any \(i\), hence \( \text{spec}(C(H)) \cap \{-1\} = \emptyset \).

Finally, we observe:
\[
C^{-1}C(H) = \frac{1 + H}{1 - H} - I = \frac{1 + H - 1 + H}{1 + H + 1 - H} = H.
\]

Moreover, set \( f_1(A) = -2A - I \) for all matrices \( A \in \mathbb{R}^{m \times m} \), \( f_2(A) = A^{-1} \) for all invertible matrices \( A \in \mathbb{R}^{m \times m} \) and \( f_3(A) = A - I \) for all matrices \( A \in \mathbb{R}^{m \times m} \). Then we have
\[
f_1 \circ f_2 \circ f_3(H) = f_1 \circ f_2(H - I) = f_1\left(\frac{1}{H - I}\right) = -\frac{2}{H - I} - I = C(H).
\]

Since \( f_i \) are differentiable for all \( i = 1, 2, 3 \), we have that \( C \) is invertible.

The same considerations with a few signs reversed also lead us to conclude that \( C^{-1} \) is well-defined and indeed the inverse of \( C \). We can similarly decompose \( C^{-1} \) to show that it is differentiable, making \( C \) a diffeomorphism. Here, we define \( g_1(A) = 2A + I \) for all \( A \in \mathbb{R}^{m \times m} \), \( g_2(A) = A^{-1} \) for all invertible \( A \in \mathbb{R}^{m \times m} \) and \( g_3(A) = A + I \) for all \( A \in \mathbb{R}^{m \times m} \). A quick calculation shows
\[
g_1 \circ g_2 \circ g_3(S) = C^{-1}(S).
\]

Denote by \( \mathcal{H} \) the set
\[
\mathcal{H} := \left\{ H = \begin{pmatrix} A & B \\ B & A \end{pmatrix} \middle| A \in \mathbb{R}^{2n \times 2n}, A^T = A, B^T = B, -1 < H < 1 \right\}
\]
where \( H < I \) means that \( I - H \) is positive definite (not just positive semidefinite). We can then prove the Cayley trick:

**Proposition B.3.** Let \( H \in \mathbb{R}^{2n \times 2n} \). Then \( H \in \mathcal{H} \iff (C(H)) \in \text{Sp}(2n) \cap \text{C}(H) \cong iJ) \).

**Proof.** Note that for \( H \in \mathcal{H} \), \( 1 \notin \text{spec}(H) \), hence \( C(H) \) is always well-defined. \( C(H) = (I + H)(I - H)^{-1} \geq 0 \), since \( 1 + H \geq 0 \) and \( (I - H)^{-1} \geq 0 \) as \(-1 < H < 1\). Observe:
Then we can calculate:

\[
(1 + H) \cdot (1 - H)^{-1} J = -(1 + H) \cdot (J (1 - H))^{-1} = -(1 + H) \cdot ((1 + H) J)^{-1} = (1 + H) (1 + H)^{-1} = J (1 - H) \cdot (1 + H)^{-1},
\]

hence \( C(H) J = J C(H)^{-1} \) and as \( C(H) \) is Hermitian, we have \( C(H)^T J C(H) = J \) and \( C(H) \) is symplectic. Via corollary 2.10, as \( C(H) \) is symplectic and positive definite, we can conclude that \( C(H) \geq iJ \).

Conversely, let \( S \in Sp(2n) \) and \( S \geq iJ \). Then \( S \geq -iJ \) by complex conjugation and \( S \geq 0 \) after averaging the two inequalities. Since any element of \( Sp(2n) \) is invertible, this implies \( S > 0 \). From this we obtain:

\[
\begin{align*}
\frac{S - 1}{S + 1} &> -1 \quad \text{as } S + 1 > 1 , \\
\frac{S - 1}{S + 1} &< 1 \quad \text{always.}
\end{align*}
\]

Write \( (S - I) \cdot (S + I)^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \). As \( S \) is Hermitian, \( A^* = A \) and \( C = B^* , D^* = D \).

We have on the one hand

\[
\begin{align*}
\frac{S - 1}{S + 1} J &= (S - I) \cdot (-S^{-T}J - J)^{-1} = (S - 1)(-S^{-1}(S^{-T} + 1)^{-1} \\
&= (SJ - I) \cdot (S^{-T} + 1)^{-1} = J (S^{-T} - I) S^{-T} (S^{-T} + 1)^{-1} \\
&= -J \frac{S - 1}{S + 1}
\end{align*}
\]

and on the other hand

\[
\begin{pmatrix} A & B \\ B^T & D \end{pmatrix} J = \begin{pmatrix} -B & A \\ -D & B^T \end{pmatrix},
\]

\[ -J \begin{pmatrix} A & B \\ B^T & D \end{pmatrix} = \begin{pmatrix} -B^T & -D \\ A & B \end{pmatrix} \]

Put together this implies \( B = B^T \) and \( D = -A \), hence \( C^{-1}(S) \in \mathcal{H} \), which is what we claimed.

**Proposition B.4.** The Cayley transform \( C \) is operator monotone and operator convex on the set of \( A = A^T \in \mathbb{R}^{m \times m} \) with \( \text{spec(A)} \subset (-1, 1) \). \( C^{-1} \) is operator monotone and operator concave on the set of \( A = A^T \in \mathbb{R}^{m \times m} \) with \( \text{spec(A)} \subset (-1, \infty) \).

**Proof.** Recall equation (64) and the definition of \( f_1 , f_2 , f_3 , f_4 \) and \( f_5 \) are affine and thus for all \( X \geq Y : f_1 (X) \geq f_1 (Y) \) and \( f_5 (X) \leq f_5 (Y) \). For \( X \geq Y \geq 0 \), we also have \( f_2 (Y) \geq f_2 (X) \geq 0 \) since matrix inversion is antimonotone. Now let \( -1 \leq Y \leq X \leq 1 \), then \(-21 \leq f_4 (Y) \leq f_4 (X) \leq 0 \) and \(-1/21 \geq f_2 \circ f_4 (Y) \geq f_2 \circ f_4 (X) \geq 0 \) and finally \( C(Y) \geq C(Y) \geq 0 \), proving monotonicity of \( C \). Similarly, one can prove that \( C^{-1} \) is monotonous using equation (65).

For the convexity of \( C \), we note that since \( f_1 , f_2 \) are affine they are both convex and concave. It is well-known that \( 1/x \) is operator convex for positive definite and operator concave for negative definite matrices (to prove this, consider convexity/concavity of the
functions $f_{\psi}(x, X^{-1})$ for all $\psi$). It follows that for $-1 \leq H \leq 1$ we have $f_{\psi}(x) \leq 0$, hence $f_{\psi} \circ f_{\psi}$ is operator concave on $-1 \leq H \leq 1$. As $f_{\psi}(A) = -2A - 1$, this implies that $C = f_{\psi} \circ f_{\psi}$ is operator convex.

For the concavity of $C^{-1}$, recall equation (65) and the definitions of $g_1$, $g_2$, $g_3$. Then, given $-1 \leq X$, we have $g_3(X)$ is positive definite and concave as an affine map. $g_2$ is concave on positive definite matrices, as $1/x$ is convex and $-1/x$ is order-reversing, hence $-1/x$ is concave on positive definite matrices. Since $g_1$ is concave as an affine map, $g_1 \circ g_2 \circ g_3 = C^{-1}$ is operator concave for all $-1 \leq X$.

**Lemma B.5.** $C : \mathbb{R} \to \mathbb{R}$ is log-convex on $[0, 1)$.

**Proof.** We need to see that the function $h(x) = \log \frac{1 + x}{1 - x}$ is convex for $x \in [0, 1)$. Since $h$ is differentiable on $(0, 1)$, this is true iff the second derivative is non-negative:

$$h''(x) = \frac{4x}{(1 - x^2)^2}$$

is clearly positive on $(0, 1)$ and $h$ is therefore log-convex.

**Appendix C. Continuity of set-valued functions**

Here, we provide some definitions and lemmata from set-valued analysis for the reader’s convenience. This branch of mathematics deals with functions $f : X \to 2^Y$ where $X$ and $Y$ are topological spaces and $2^Y$ denotes the power set of $Y$.

In order to state the results interesting to us we define:

**Definition C.1.** Let $X, Y \subseteq \mathbb{R}^n \times m$ and $f : X \to 2^Y$ be a set-valued function. Then we say that a function is **upper semicontinuous** (often also called **upper hemicontinuous** to distinguish it from other notions of continuity) at $x_0 \in X$ if for all open neighbourhoods $Q$ of $f(x_0)$ there exists an open neighbourhood $W$ of $x_0$ such that $W \subseteq \{x \in X | f(x) \subseteq Q\}$. Likewise, we call it **lower semicontinuous** (often called **lower hemicontinuous**) at a point $x_0$ if for any open set $V$ intersecting $f(x_0)$, we can find a neighbourhood $U$ of $x_0$ such that $f(x) \cap V \neq \emptyset$ for all $x \in U$.

Note that the definitions are valid in all topological spaces, but we only need the case of finite dimensional normed vector spaces. Using the metric, we can give the following characterisation of upper semicontinuity:

**Lemma C.2.** Let $X, Y \subseteq \mathbb{R}^n \times m$ and $f : X \to 2^Y$ be a set-valued function such that $f(x)$ is compact for all $x$. Then $f$ is upper semicontinuous at $x_0$ if and only if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in X$ with $\|x - x_0\| < \delta$ we have: for all $y \in f(x)$ there exists a $\tilde{y} \in f(x_0)$ such that $\|y - \tilde{y}\| < \varepsilon$.

**Proof.** $\Rightarrow$: Let $f$ be lower semicontinuous at $x_0$. For any $\varepsilon > 0$ the set

$$B(\varepsilon, f(x_0)) = \bigcup_{y \in f(x_0)} \{\tilde{y} \in Y | \|y - \tilde{y}\| < \varepsilon\}$$

is compact. Since $f(x_0)$ is compact, there exists a subsequence $f(x_{n_k})$ that converges to some point $\tilde{y} \in f(x_0)$. For each $\tilde{y}$, there exists a $\tilde{y}_k \in f(x_{n_k})$ such that $\|y - \tilde{y}_k\| < \varepsilon$ for all $y \in f(x)$.

Thus, $\tilde{y}_k \to \tilde{y}$, and $f$ is upper semicontinuous at $x_0$.

$\Rightarrow$: Let $f$ be upper semicontinuous at $x_0$. Then for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in X$ with $\|x - x_0\| < \delta$ we have:

$$f(x) \subseteq \bigcup_{y \in f(x_0)} \{\tilde{y} \in Y | \|y - \tilde{y}\| < \varepsilon\}$$

for all $y \in f(x)$. Therefore, $f$ is lower semicontinuous at $x_0$.
is an open neighbourhood of \( f(x_0) \). Hence there exists an open neighbourhood \( W \) of \( x_0 \), which contains a ball of radius \( \delta > 0 \) such that \( B_{\delta}(x_0) \subseteq W \subseteq \{ x \in X | f(x) \subseteq B(\varepsilon, f(x_0)) \} \).

Clearly this implies the statement.

\(<=>\) Let \( Q \) be a neighbourhood of \( f(x_0) \). Since \( f(x_0) \) is compact this implies that there is a \( \varepsilon > 0 \) such that \( B(\varepsilon, f(x_0)) \subseteq Q \) where this set is defined as in equation (67). If this were not the case, for every \( n \in \mathbb{N} \) there must be a \( y_n \in Y \setminus Q \) such that \( \inf_{x \in f(x_0)} \| y_n - \hat{y} \| < 1/n \).

Since by construction this implies that \( y_n \in B(1, f(x_0)) \), which is compact, a subsequence of these \( y_n \) must converge to \( y \). As \( Y \setminus Q \) is closed as \( Q \) is open, \( y \in Y \setminus Q \). However, \( \inf_{x \in f(x_0)} \| y - \hat{y} \| = 0 \) by construction and since \( f(x_0) \) is compact, the infimum is attained, which implies \( y \in f(x_0) \). This contradicts the fact that \( Q \) is a neighbourhood of \( f(x_0) \).

Hence we know that for any open \( Q \) containing \( f(x_0) \) there exists a \( \varepsilon > 0 \) such that \( B(\varepsilon, f(x_0)) \subseteq Q \). By assumption, this implies that there exists a \( \delta > 0 \) such that \( B_{\delta}(x_0) \subseteq \{ x \in X | f(x) \subseteq B(\varepsilon, f(x_0)) \} \). Since clearly \( \{ x \in X | f(x) \subseteq B(\varepsilon, f(x_0)) \} \subseteq \{ x \in X | f(x) \subseteq Q \} \) we can choose \( W = B_{\delta}(x_0) \) to finish the proof. \( \square \)

This second characterisation is sometimes called upper Hausdorff semicontinuity and it can equally be defined in any metric space. Clearly, the notions can differ for set-valued functions with non-compact values or in spaces which are not finite dimensional. With these two definitions, we can state the following classic result:

**Proposition C.3** ([DR79]). Let \( Y \) be a complete metric space, \( X \) a topological space and \( f : X \to 2^Y \) a compact-valued set-valued function. The following statements are equivalent:

- \( f \) is upper semicontinuous at \( x_0 \).
- for each closed \( K \subseteq X \), \( K \cap f(x_0) \) is upper semicontinuous at \( x_0 \).

An interesting question would be whether the converse is also true. Even if \( f(x) \) is always convex, this need not be the case if \( K \cap f(x_0) \) has empty interior as simple counterexamples can show. In case the interior is non-empty, another classic result guarantees a converse in many cases:

**Proposition C.4** ([Mor75]). Let \( X \) be a compact interval and \( Y \) a normed space. Let \( f : X \to 2^Y \) and \( g : X \to 2^Y \) be two convex-valued set-valued functions. Suppose that \( \text{diam}(f(t) \cap g(t)) < \infty \) and \( f(t) \cap \text{int}(g(t)) \neq \emptyset \) for all \( t \). Then if \( f, g \) are continuous (in the sense above) so is \( f \cap g \).

**Appendix D. Reduction of the set of necessary operations for state preparation**

In this section, we give a justification of why the operations (O0)–(O7) are enough to implement all operations described in section 5. All of this is known albeit scattered throughout the literature, hence we collect it here.

In order to prepare a state, we could start with the vacuum \( \gamma = 1 \) or alternatively a thermal state for some bath \( (\gamma = (1/2 + N)1) \) with photon number \( N \), see e.g. [Oli12]). Of course, we should be able to draw arbitrary ancillary modes of this system, too. The effect of Gaussian noise on the covariance matrix is given in [Lin00]. Since for any \( \gamma \geq 1 \) we can decompose it as \( \gamma = 1 + \gamma_{\text{noise}} \), this implies that the operations (O0)–(O2) are enough to implement all operations 1. and 2.
As with other squeezing measures, passive transformations should not change the squeezing measure, while single-mode-squeezers are not free. The effect of symplectic transformations on the covariance matrix has already been observed in equation (10), hence (O3) and (O4) implement operations (3) and (8).

Since we have the Weyl-system at our disposal, we can also consider its action on a quantum state (translation in phase space). Direct computation shows that it does not affect the covariance matrix. Including it as operation (O5) is beneficial if we consider a convex combination of states. In an experiment, this can be done by creating ensembles of the states of the convex combination and creating another ensemble where the ratio of the different states is that of the convex combination. On the level of covariance matrices, we have the following lemma:

**Lemma D.1.** Let $\rho$ and $\rho'$ be two states with displacement $d^\rho$ and $d'^\rho$ and (centred) covariance matrices $\gamma^\rho$ and $\gamma'^\rho$. For $\lambda \in (0, 1)$, the covariance matrix of $\hat{\rho} \equiv \lambda \rho + (1 - \lambda) \rho'$ is given by:

$$
\gamma^\rho = \lambda \gamma^\rho + (1 - \lambda) \gamma'^\rho + 2\lambda(1 - \lambda)(d^\rho - d'^\rho)(d^\rho - d'^\rho)^T
$$

A proof of this statement can be found in [WW01] (in the proof of proposition 1). Note that for centralised states with $d^\rho = 0$ and $d'^\rho = 0$, a convex combination of states translates to a convex combination of covariance matrices. Since in particular, $2\lambda(1 - \lambda)(d^\rho - d'^\rho)(d^\rho - d'^\rho)^T \succeq 0$, any convex combination of $\rho$ and $\rho'$ is on the level of covariance matrices equivalent to

- centring the states (no change in the covariance matrices),
- taking a convex combination of the states (resulting in a convex combination of covariance matrices),
- performing a Weyl translation to undo the centralization in the first step (no change in the covariance matrix),
- adding noise $2\lambda(1 - \lambda)(d^\rho - d'^\rho)(d^\rho - d'^\rho)^T \succeq 0$.

This implies that the effect of any convex combination of states (operation 4) on the covariance matrix can equivalently be obtained from operations (O2), (O5) and (O6). Finally, we consider measurements. Homodyne detection is the measurement of $Q$ or $P$ in one of the modes, which corresponds to the measurement of an infinitely squeezed pure state in lemma D.2. A broader class of measurements known as heterodyne detection measures arbitrary coherent states [Wee+12]. Let us focus our attention on the even broader class of projections onto Gaussian pure states.

**Lemma D.2.** Let $\rho$ be an $(n + 1)$-mode quantum state with covariance matrix $\gamma$ and $|\gamma_G, d\rangle \langle \gamma_G, d|$ be a pure single-mode Gaussian state with covariance matrix $\gamma_G \in \mathbb{R}^{2 \times 2}$ and displacement $d$. Let

$$
\gamma = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix}, \quad B \in \mathbb{R}^{2 \times 2}
$$

then the selective measurement of $|\gamma_G, d\rangle$ in the last mode results in a change of the covariance matrix of $\rho$ according to:

$$
\gamma' = A - C(B - \gamma_G)MPC^T,
$$

(68)
where $MP$ denotes the Moore–Penrose pseudoinverse. Homodyne detection corresponds to the case where $\gamma_G$ is an infinitely squeezed state.

This can most easily be seen on the level of Wigner functions, as demonstrated in [ESP02, GIC02]. The generalisation to multiple modes is straightforward.

Since the covariance matrix of a Gaussian pure state is a symplectic matrix (see proposition 2.2), using the Euler decomposition we can implement a selective Gaussian measurement by

1. a passive symplectic transformation $S \in K(n + 1)$,
2. a measurement in the Gaussian state $\text{diag}(d, 1/d)$ for some $d \in \mathbb{R}_+$ according to lemma D.2.

A non-selective measurement (forgetting the information obtained from measurement) would then be a convex combination of such projected states. A measurement of a multimode state can be seen as successive measurements of single-mode states since the Gaussian states we measure are diagonal.

For homodyne detection, since an infinitely squeezed single-mode state is given by the covariance matrix $\lim_{d \to \infty} \text{diag}(1/d, d)$, we have

$$\gamma' = \lim_{d \to \infty} (A - C(B - \text{diag}(1/d, d))^{-1}C^T) = A - C(\pi B \pi)^{MP}C^T, \quad (69)$$

where $\pi = \text{diag}(1, 0)$ is a projection and $^{MP}$ denotes the Moore–Penrose pseudoinverse. It has been shown (see [Wee+12] E.2 and E.3 as well as [ESP02, GIC02]) that any (partial or total) Gaussian measurement is a combination of passive transformations, discarding subsystems, projection onto Gaussian states and homodyne detection.

Therefore, we should also allow to discard part of the system, i.e. taking the partial trace. However, this can be expressed as a combination of operations (O1)–(O6) and homodyne detection:

**Lemma D.3.** Given a covariance matrix $\gamma = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix}$ a partial trace on the second system translates to a map $\gamma \mapsto A$. The partial trace can then be implemented by measurements and adding noise.

**Proof.** When measuring the modes $B$, we note that since $C(\pi B \pi)^{MP}C^T \succeq 0$ in equation (69), a partial trace is equivalent to first performing a homodyne detection on the $B$-modes of the system and then adding noise. □

Given the discussion above, lemmas D.2 and D.3 put together imply: on the level of covariance matrices, in order to allow for general Gaussian measurements, it suffices to consider Gaussian measurements of the state $|\gamma_d, 0\rangle\langle 0, \gamma_d|$ with covariance matrix $\gamma_d = \text{diag}(1/d, d)$ for $d \in \mathbb{R}_+ \cup \{+\infty\}$. All Gaussian measurements are then just combinations of these special measurements and operations (O1)–(O6).

**Appendix E. Numerical implementation and documentation**

Here, we provide a short documentation to the programme written in MATLAB, Version R2014a, and used for the numerical computations in section 6. The source code can be found at GitHub https://github.com/Martin-Idel/operationalsqueezing.
The programme tries to minimise the function $f$ defined in equation (28) over the set $\mathcal{H}$. Throughout, suppose we are given a covariance matrix $\gamma$.

Let us first describe the implementation of $f$: as parameterisation of $\mathcal{H}$, we choose the simplest parameterisation such that for matrices with symplectic eigenvalues larger than one, the set of feasible points has non-empty interior: we parameterise $A, B$ via matrix units $E_i, E_{jk}$ with $i \in \{1, \ldots, n\}$, $k \in \{1, \ldots, n-1\}$ and $j < k$, where $(E_i)_{jk} = \delta_{ji}\delta_{ik}$ and $(E_{jk})_{lm} = \delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl}$. This parameterisation might not be very robust, but it is good enough for our purpose. Instead of working with complex parameters, we compute $s_j(A + iB)$ as $\lambda_j^+(H)$ for the matrix

$$H = \begin{pmatrix} A & B \\ B & -A \end{pmatrix}. \quad (70)$$

The evaluation of $f$ is done in function OBJECTIVE.M. Since $f$ is not convex for $(A, B)$ with the corresponding $H$ having eigenvalues $\geq 1$ or $\leq -1$, the function first checks, whether this constraint is satisfied and outputs a value that is $10^7$-times larger than the value of the objective function at the starting point otherwise.

The constraints are implemented in function MAXRESIDUAL.M. Via symmetry, it is enough to check that for any $H$ tested, $\lambda^{-}_1(H) \geq 1$. The second constraint is given by $C^{-1}(\gamma) \geq H$ and this is tested by computing the smallest eigenvalue of the difference.

The function which is most important for users is MINIMUM.M, which takes a covariance matrix $\Gamma \geq iJ$, its dimensions $n$ and a number of options as arguments and outputs the minimum. Note that the programme checks whether the covariance matrix is valid. For the minimisation, we use the MATLAB-based solver SOLVOPT ([KK97], latest version 1.1). SOLVOPT uses a subgradient based method and the method of exact penalization to compute (local) minima. For convex programming, any minimum found by the solver is therefore an absolute minimum. In order to work, the objective function may not be differentiable on a set of measure zero and it is allowed to be non-differentiable at the minimum. Since $f$ is differentiable for all $H$ with non-degenerate eigenvalues, this condition is met. In addition, SOLVOPT needs $f$ to be defined everywhere, as it is not an interior point method. Since $f$ is well-defined but not convex for $H \not\in \mathcal{H}$ and $\text{spec}(H) \cup \{1\} = \emptyset$, we remedy this by changing the output of OBJECTIVE.M to be very large when $H \not\in \mathcal{H}$ as described above. Constraints are handled via the method of exact penalisation. We used SOLVOPT’s algorithm to compute the penalisation functions on its own.

It is possible (and for speed purposes advisable) to implement analytical gradients of both the objective and the constraint functions. Following [Mag85], for diagonalisable matrices $A$ with no eigenvalue multiplicities, the derivative of an eigenvalue $\lambda_i(A)$ is given by:

$$\partial_{A} \lambda_i(A) = v_i(A)^2 \partial_{A} v_i(A), \quad (71)$$

where $v_i(A)$ is the eigenvector corresponding to $\lambda_i(A)$ and $\partial_{A}(A) = \lim_{h \to 0}(A + hE - A)/h = E$. Luckily, if $A$ is not differentiable, this provides at least one subgradient. An easy calculation shows that a subgradient of the objective function $f$ for matrices $H$ with $-1 < H < 1$ in the parameterisation of the matrix units $E_{ij}$ is given by

$$\langle \nabla f \rangle_i = \sum_{j=1}^{n} \frac{\partial_{A} \lambda_j^+(H)}{(1 + \lambda_j(H))(1 - \lambda_j(H))^2} = \sum_{j,k=1}^{n} \frac{v^*_j F(i) v_k}{(1 + \lambda_j(H))(1 - \lambda_j(H))^2} \quad (72)$$

with $F$ being the matrices corresponding to the chosen parameterisation. The gradient of the constraint function is very similar and given by equation (71) for $A = \gamma - H$ or $A = 2I - H$ depending on which constraint is violated. This is implemented in functions OBJECTIVEGRAD.M and MAXRESIDUALGRAD.M.
SOLVOPT needs a starting point. Given $\Gamma$, via Williamson’s theorem, $\Gamma = S^TDS \geq S^T S$, hence $S^T S$ provides a good starting point. The function WILLIAMSON.M computes the Williamson normal form for $\gamma$ and returns $S$, $D$ and $S^T S$, the latter of which is used as starting point. It computes $S$ and $D$ essentially by computing the Schur decomposition of $\Gamma^{-1/2} \Gamma^{-1/2}$ (in the $\sigma$-basis instead of the $J$-basis). $S$ is then given by $S = \gamma^{1/2} \mathbf{K} \mathbf{D}^{-1/2}$ (see the proof of [SCS99]), where $\mathbf{K}$ is the Schur transformation matrix.

A number of comments are in order:

1. All functions use global variables instead of function handles. This is required by the fact that SOLVOPT has not been adapted to the use of function handles. The user should therefore always reset all variables before running the programme.

2. SOLVOPT is not an interior point method, i.e. the results can at times violate constraints. We use the default value for the accuracy of constraints, which is $10^{-8}$ and can be modified by option six. The preparation error should be of the same order than the accuracy of constraints as long as the largest eigenvalue of the minimising symplectic matrix is of order one.

3. For our numerical tests, we used bounds on the minimal step-size and the minimal error in $f$ (SOLVOPT options two and three) of the order $10^{-6}$ and $10^{-8}$, which seemed sufficient.

4. All functions called by SOLVOPT (the functions OBJECTIVE.M, OBJECTIVEGRAD.M, MAXRESIDUAL.M, MAXRESIDUALGRAD.M and XTOH.M) are properly vectorised to ensure maximal speed.

Finally, BOUNDS.M contains all lower- and upper bounds described in section 4.5. The semidefinite programme was solved using CVX (version SDPT3 4.0), a toolbox developed in MATLAB for disciplined convex programming including semidefinite programming [GB08], [GB14]. The third bound is not described in section 4.5—it is an iteration of corollary 4.8 assuming superadditivity, hence in principle it could be violated. If it were violated, this would immediately disprove superadditivity, which has never been observed in our tests.

**Issues and further suggestions:** It occurs sometimes that the algorithm does not converge to a minimum inside or near the feasible set. We believe that this is due to instabilities in the parameterisation and implementation. The behaviour can occur while using numerical as well as analytical subgradients, although it occurs more often with analytical ones. For every example where we could observe a failure with either numerical or analytical subgradients, one other method (using numerical subgradients, using analytical subgradients or a mixture thereof) worked fine. In cases of failure, the routine issued several warnings and the result usually lies below the lower bound. A different type of implementation might lead to an algorithm that is more stable, but we did not pursue this any further. It might also be worth to consider trying to compute the penalty function analytically.

In terms of performance times, the algorithm is generally fast for small numbers of modes. When analytical subgradients are not implemented, the performance bottleneck is given by the functions XTOH.M, which is called most often. When analytical subgradients are provided, the performance is naturally much faster. This is particularly important when the number of modes increases. While for five modes, the calculation is done within seconds, already for ten modes and depending on the matrix, it can take a minute on a usual laptop (the algorithm now takes the most amount of time for eigenvalue computations, which seems unavoidable). For even larger matrices, it might be advisable to switch from using the Matlab function EIG to EIGF, but for our examples this did not lead to a time gain.
References

[AG88] Arnol’d V I and Givental’ A B 1988 Symplectic geometry Dynamical Systems IV (Berlin: Springer)

[And+15] Andersen U L et al 2016 30 years of squeezed light generation Phys. Scr. 91 053001

[ARL14] Adesso G, Ragy S and Lee A R 2014 Continuous variable quantum information: Gaussian states and beyond Open Syst. Inf. Dyn. 21 1440001

[Arv+i-95a] Arvind et al 1995 The real symplectic groups in quantum mechanics and optics Pramana 45 471–97

[Arv+i-95b] Arvind et al 1995 Two-mode quantum systems: invariant classification of squeezing transformations and squeezed states Phys. Rev. A 52 1609–20

[ASB04] Adesso G, Serafini A and Illuminati F 2004 Extremal entanglement and mixedness in continuous variable systems Phys. Rev. A 70 022318

[Bha07] Bhatia R 2007 Positive Definite Matrices (Princeton, NJ: Princeton University Press)

[Bhal96] Bhatia R 1996 Matrix Analysis (Berlin: Springer)

[BL05] Braunstein S L and van Loock P 2005 Quantum information with continuous variables Rev. Mod. Phys. 77 513–77

[Bra05] Braunstein S L 2005 Squeezing as an irreducible resource Phys. Rev. A 71 055801

[BV04] Boyd S and Vandenberghe L 2004 Convex Optimization (New York: Cambridge University Press) ISBN 0521833787

[DR79] Dolecki S and Rolewicz S 1979 Metric characterizations of upper semicontinuity J. Math. Anal. Appl. 69 146–52

[ESP02] Eisert J, Scheel S and Plenio M B 2002 Distilling Gaussian states with Gaussian operations is impossible Phys. Rev. Lett. 89 137903

[Fil13] Filip R 2013 Distillation of quantum squeezing Phys. Rev. A 88 063837

[GB14] Grant M and Boyd S 2014 CVX: Matlab Software for Disciplined Convex Programming, version 2.1 (http://cvxr.com/cvx)

[GB08] Grant M and Boyd S 2008 Graph implementations for nonsmooth convex programs Recent Advances in Learning and Control (Lecture Notes in Control and Information Sciences) ed V Blondel, S Boyd and H Kimura (Springer) pp 95–110

[GIC02] Giedke G and Cirac J I 2002 Characterization of Gaussian operations and distillation of Gaussian states Phys. Rev. A 66 032316

[Gie+i-01] Giedke G et al 2001 Separability properties of three-mode gaussian states Phys. Rev. A 64 052303

[Gos06] de Gosson M A 2006 Symplectic Geometry and Quantum Mechanics (Operator Theory: Advances and Applications/Advances in Partial Differential Equations) (Basel: Birkhäuser) ISBN 9783764375751

[Hee+i-06] Heersink J et al 2006 Distillation of squeezing from non-gaussian quantum states Phys. Rev. Lett. 96 253601

[KK97] Kuttsevich A and Kappel F 1997 SolvOpt: the solver for local nonlinear optimization problems (manual) (Institute for Mathematics, Karl-Franzens University of Graz)

[KL10] Kok P and Lovett B W 2010 Introduction to Optical Quantum Information Processing (Cambridge: Cambridge University Press)

[Kok+i-07] Kok P et al 2007 Linear optical quantum computing with photonic qubits Rev. Mod. Phys. 79 135–74

[Kra+i-03] Kraus B et al 2003 Entanglement generation and Hamiltonian simulation in continuous-variable systems Phys. Rev. A 67 042314

[Lee88] Lee C T 1988 Wehrl’s entropy as a measure of squeezing Opt. Commun. 66 52–4

[LGW13] Lercher D, Giedke G and Wolf M M 2013 Standard super-activation for gaussian channels requires squeezing New J. Phys. 15 123003

[Lin00] Lindblad G 2000 Cloning the quantum oscillator J. Phys. A: Math. Gen. 33 5059

[Lvo15] Lvovsky A I 2015 Squeezed light Photonics Volume 1: Fundamentals of Photonics and Physics ed D Andrews (New York: Wiley) pp 121–164

[Mag85] Magnus J R 1985 On differentiating eigenvalues and eigenvectors Econometric Theor. 1 179–91

[MK08] Mišta L and Korolkova N 2008 Distribution of continuous-variable entanglement by separable gaussian states Phys. Rev. A 77 050302

[J. Phys. A: Math. Theor. 49 (2016) 445304]
[Mor75] Moreau J J 1975 Intersection of moving convex sets in a normed space *Math. Scand.* **36** 159–73

[MS98] McDuff D and Salamon D 1998 *Introduction to Symplectic Topology* (Oxford: Oxford University Press)

[NC00] Nielsen M and Chuang I 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press) (doi:10.1017/CBO9780511976667)

[Oli12] Olivares S 2012 Quantum optics in the phase space *Eur. Phys. J. Spec. Top.* **203** 3–24

[Rec+94] Reck M *et al* 1994 Experimental realization of any discrete unitary operator *Phys. Rev. Lett.* **73** 58–61

[RFP10] Recht B, Fazel M and Parrilo P A 2010 Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization *SIAM Rev.* **52** 471–501

[Roc97] Rockafellar R T 1997 *Convex Analysis* (Princeton, NJ: Princeton University Press) ISBN 9780691015866

[Rud87] Radin W 1987 *Real and Complex Analysis* (Mathematics Series) (New York: McGraw-Hill) ISBN 9780070542341

[SCS99] Simon R, Chaturvedi S and Srinivasan V 1999 Congruences and canonical forms for a positive matrix: application to the Schweinler–Wigner extremum principle *J. Math. Phys.* **40** 3632–42

[SMD94] Simon R, Mukunda N and Dutta B 1994 Quantum-noise matrix for multimode systems: $U(n)$ invariance, squeezing, and normal forms *Phys. Rev. A* **49** 1567–83

[Son98] Sonntag E D 1998 *Mathematical Control Theory: Deterministic Finite Dimensional Systems* 2nd edn (New York: Springer)

[SZ97] Scully M O and Zubairy M S 1997 *Quantum Optics* (Cambridge: Cambridge University Press) ISBN 9780521435956

[Tho76] Thompson R C 1976 Convex and concave functions of singular values of matrix sums *Pac. J. Math.* **66** 285–90

[VB96] Vandenberghe L and Boyd S 1996 Semidefinite programming *SIAM Rev.* **38** 49–95

[Wee+12] Weedbrook C *et al* 2012 Gaussian quantum information *Rev. Mod. Phys.* **84** 621

[WEPl03] Wolf M M, Eisert J and Plenio M B 2003 Entangling power of passive optical elements *Phys. Rev. Lett.* **90** 047904

[Wil36] Williamson J 1936 On the algebraic problem concerning the normal forms of linear dynamical systems *Am. J. Math.* **58** 141–63

[Wol+04] Wolf M M *et al* 2004 Gaussian entanglement of formation *Phys. Rev. A* **69** 052320

[WW01] Werner R F and Wolf M M 2001 Bound entangled Gaussian states *Phys. Rev. Lett.* **86** 3658–61